Math Camp Weeks 3 and 4:
Linear Algebra and Differentiation

University of Pennsylvania

Economics Department[1]

August 25, 2021

[1]These notes are the result of a collective effort by previous instructors of the second part of the math camp, including Xincheng Qiu (2020-2021), Alejandro Sánchez (2017-2019), David Zarruk (2014-2015) and Ju Hu (2012-2013). The majority of the proofs of the differentiation section are adapted from Pugh (2010), *Real Mathematical Analysis* and Rudin (1976), *Principles of Mathematical Analysis*. These notes are always growing and improving with the comments and insights from our students. If you have any suggestions or see any typos, please email Xincheng Qiu at qiux@sas.upenn.edu.

Contents

Part I

Linear Algebra

Chapter 1

Overview of Linear Algebra

1.1 Matrices

Let $n, m \geq 1$ be two integers. A matrix is a two-dimensional array of numbers in $\mathbb{R}$. A vector is an array that has a single column.

$$A := \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{x} := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Sometimes we abbreviate the notation for this matrix by writing it as (a_{ij}), $i = 1, \cdots, m$ and $j = 1, \cdots, n$. We say that it is an m by n matrix, or an $m \times n$ matrix. The matrix has m **rows** and n **columns**. If $m = n$, we call it a **square** matrix. We can define the multiplication of a matrix times a vector. Let A_m be the m^{th} row of the matrix.

$$A\boldsymbol{x} := \begin{pmatrix} A_1\boldsymbol{x} \\ \vdots \\ A_m\boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix}, \quad \boldsymbol{x} \in \mathbb{R}^n$$

The resulting array $A\boldsymbol{x}$ is a vector in $\mathbb{R}^m$. A linear combination of two vectors, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ is defined as

$$\boldsymbol{w} = \alpha\boldsymbol{x} + \beta\boldsymbol{y} := \begin{pmatrix} \alpha x_1 + \beta y_1 \\ \vdots \\ \alpha x_n + \beta y_n \end{pmatrix}$$

where $\alpha, \beta \in \mathbb{R}$ are scalars. The resulting array is also a vector in $\mathbb{R}^n$. Each coordinate is weighted and added separately, that is $w_j = \alpha x_j + \beta x_j$.

Multiplication of two matrices can be defined analogously. Let B be a $n \times K$ matrix whose columns $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K$ are $n \times 1$ vectors.

$$AB := \begin{pmatrix} A_1\boldsymbol{b}_1 & \cdots & A_1\boldsymbol{b}_K \\ \vdots & \ldots & \vdots \\ A_m\boldsymbol{b}_1 & \cdots & A_m\boldsymbol{b}_K \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}b_{j1} & \cdots & \sum_{j=1}^n a_{1j}b_{jK} \\ \vdots & \cdots & \vdots \\ \sum_{j=1}^n a_{mj}b_{j1} & \cdots & \sum_{j=1}^n a_{mj}b_{jK} \end{pmatrix}$$

Each column of the resulting matrix AB is $A\boldsymbol{b}_k$, $1 \leq k \leq K$.

We say that a matrix (or vector) is **zero matrix** (vector) when all its entries are zero.

1.2 Linear Maps and Matrices

We will focus on a particular type of function called a linear map.

Definition 1.2.1. A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is a finite linear map if

$$T(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha T(\boldsymbol{x}) + \beta T(\boldsymbol{y})$$

for all vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and scalars $\alpha, \beta \in \mathbb{R}$.

In this section we lay the groundwork to prove that all finite linear maps have the form $T(\boldsymbol{x}) = A\boldsymbol{x}$, where A is an $m \times n$ matrix. Clearly not all functions are linear maps. For example, given the function $T(x) = x^2, x \in \mathbb{R}$, we know that $T(\alpha x + \beta y) = (\alpha x + \beta y)^2 = \alpha^2 x^2 + 2\alpha\beta xy + \beta^2 y^2$. This is not equal to $\alpha T(x) + \beta T(y) = \alpha x^2 + \beta y^2$ for all values of x, y and α, β.

1.2.1 Extending Definition

First we need to prove an intermediate lemma. Our objective is to show that a linear map applies to multiple linear combinations, not just pairwise combinations.

Lemma 1.2.1. *$T(\boldsymbol{x})$ is a finite linear map if and only if for any finite integer K,*

$$T\left(\sum_{k=1}^{K} \alpha_k \boldsymbol{x}^{(k)}\right) = \sum_{k=1}^{K} \alpha_k T(\boldsymbol{x}^{(k)})$$

for all $\alpha_k \in \mathbb{R}, \boldsymbol{x}^{(k)} \in \mathbb{R}^n$.

Proof. Proving ($\Longleftarrow$) follows straight from the definition of a linear map by taking $K = 2$. Proving ($\Longrightarrow$) is the novel part. Many proofs in linear algebra proceed by induction, so it is helpful to become familiarized with the technique. We know that the results holds for $K = 2$. Assume that it holds for some $k^* \geq 2$, that is:

$$T\left(\sum_{k=1}^{k^*} \alpha_k \boldsymbol{x}^{(k)}\right) = \sum_{k=1}^{k^*} \alpha_k T(\boldsymbol{x}^{(k)}).$$

Now we will prove that it also holds for $k^* + 1$.

$$
\begin{aligned}
T\left(\sum_{k=1}^{k^*+1} \alpha_k \boldsymbol{x}^{(k)}\right) &= T\left(\sum_{k=1}^{k^*} \alpha_k \boldsymbol{x}^{(k)} + \alpha_{k^*+1}\boldsymbol{x}^{(k^*+1)}\right) && \text{By decomposing sum inside function} \\
&= T\left(\sum_{k=1}^{k^*} \alpha_k \boldsymbol{x}^{(k)}\right) + \alpha_{k^*+1} T(\boldsymbol{x}^{(k^*+1)}) && \text{By definition of a linear map} \\
&= \sum_{k=1}^{k^*} \alpha_k T(\boldsymbol{x}^{(k)}) + \alpha_{k^*+1} T(\boldsymbol{x}^{(k^*+1)}) && \text{By hypothesis in the inductive step} \\
&= \sum_{k=1}^{k^*+1} \alpha_k T(\boldsymbol{x}^{(k)}). && \text{Grouping terms}
\end{aligned}
$$

Therefore we have shown that the result holds for any finite integer K. Notice that in the second line we use the definition of a linear map. We plug in weights $\alpha = 1$ and $\beta = \alpha^{k^*+1}$, choosing vectors $\boldsymbol{x} = \sum_{k=1}^{k^*} \alpha_k \boldsymbol{x}^{(k)}$ and $\boldsymbol{y} = \boldsymbol{x}^{(k^*+1)}$ according to our previous definitions. □

1.2.2 Examples of Infinite Linear Maps

The notion of linear map can be extended beyond the Euclidean space, such as function spaces. In this case it is more common to call a linear map an *operator* rather than a function. What are examples of function spaces? The set of polynomials, the set of continuous functions, etc. (all of which have an infinite number of elements). We will only focus on the Euclidean space, but it is worth commenting that function operators emerge in many economic applications such as nonparametric econometrics.

Example 1. Let f, g be two differentiable functions and let f', g' be their derivatives, and let T be the differentiation operator. Let $T(f) = f'$ and let $T(g) = g'$. Then T is a linear map because $T(\alpha f + \beta g)(x) = \alpha f'(x) + \beta g'(x) = \alpha T(f)(x) + \beta T(g)(x)$. This is the known rule that the derivative of a linear combination two functions is a linear combination of the derivatives.

Example 2. Let f, g be two functions and define the integral operator as $T(f) = \int f(x)dx$. The integration operator is a linear map because $T(\alpha f + \beta g) = \int(\alpha f(x) + \beta g(x))dx = \alpha \int f(x)dx + \beta \int g(x)dx = \alpha T(f) + \beta T(g)$.

1.3 Matrices uniquely represent finite linear maps

Linear maps and matrices are the core of linear algebra. We will show that matrices uniquely represent linear maps by breaking down the result into two lemmas (a common strategy for a representation theorem). First, if we are given a matrix A we can show that a function of the form $T(\boldsymbol{x}) = A\boldsymbol{x}$ is indeed a linear map. Second we will show if a function T is a linear map, then we can construct a unique matrix A such that $T(\boldsymbol{x}) = A\boldsymbol{x}$.

Lemma 1.3.1. *$T(\boldsymbol{x}) = A\boldsymbol{x}$ is a linear map.*

Proof. Let $\boldsymbol{w} = \alpha\boldsymbol{x} + \beta\boldsymbol{y}$. We apply the definition of multiplication of a matrix times a vector.

$$T(\boldsymbol{w}) = A\boldsymbol{w} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}w_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}w_j \end{pmatrix}$$

Then we can substitute each coordinate separately and decompose the above expression as a sum of two vectors:

$$A\boldsymbol{w} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}(\alpha x_j + \beta y_j) \\ \vdots \\ \sum_{j=1}^{n} a_{mj}(\alpha x_j + \beta y_j) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\alpha x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}\alpha x_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^{n} a_{1j}\beta y_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}\beta y_j \end{pmatrix}$$

The scalars α, β multiply each element of the respective vector. Therefore, we can pull it out as a common term.

$$A\boldsymbol{w} = \alpha \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{pmatrix} + \beta \begin{pmatrix} \sum_{j=1}^{n} a_{1j}y_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}y_j \end{pmatrix}$$

Each vector satisfies the definition of matrix multiplication, $A\boldsymbol{x}$ and $A\boldsymbol{y}$, respectively. Therefore we have shown that $A\boldsymbol{x}$ is a linear map, since

$$T(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = T(\boldsymbol{w}) = A\boldsymbol{w} = \alpha A\boldsymbol{x} + \beta A\boldsymbol{y} = \alpha T(\boldsymbol{x}) + \beta T(\boldsymbol{y}).$$

□

Now we will show that the set of $m \times n$ matrices provides an exhaustive representation of all linear maps.

Lemma 1.3.2. *For every finite-dimensional linear map T there exists a unique matrix A such that $T(\boldsymbol{x}) = A\boldsymbol{x}$, for all $\boldsymbol{x} \in \mathbb{R}^n$.*

Proof. We will proceed by constructing the matrix A. We start off by proposing a candidate matrix representation $A\boldsymbol{x}$ evaluated at a finite number of points. To complete the proof we need to show that $T(\boldsymbol{x}) = A\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Our candidate points will be the **set of elementary basis vectors** $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n$. The j^{th} vector has a 1 on coordinate j and 0, otherwise. We provide an illustration for $n = 3$, but the proof applies to any dimension.

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Notice that any vector can be represented as a linear combination of elementary basis vectors, because $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n$. We will construct our matrix A as follows. We give an illustration for the case that $m = 2$ and $n = 3$.

$$A = \begin{pmatrix} \uparrow & \cdots & \uparrow \\ T(\boldsymbol{e}_1) & \cdots & T(\boldsymbol{e}_n) \\ \downarrow & \cdots & \downarrow \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$$

For example, $\boldsymbol{a}_j = T(\boldsymbol{e}_j)$ is a an $m \times 1$ vector, containing coordinates $(a_{1j}, \ldots, a_{mj})$. Notice that m is not necessarily equal to n. Now let $\boldsymbol{x}$ be any arbitrary vector in $\mathbb{R}^n$.

$$A\boldsymbol{x} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} = \sum_{j=1}^n x_j\boldsymbol{a}_j = \sum_{j=1}^n x_j T(\boldsymbol{e}_j)$$

Notice that from the extended definition of a linear map,

$$\sum_{j=1}^n x_j T(\boldsymbol{e}_j) = T\left(\sum_{j=1}^n x_j\boldsymbol{e}_j\right) = T(\boldsymbol{x})$$

where $\boldsymbol{x} = \sum_{j=1}^n x_j\boldsymbol{e}_j$. Therefore, $T(\boldsymbol{x})$ can be represented using the matrix A. To prove uniqueness assume that there exists another matrix B such that $T(\boldsymbol{x}) = B\boldsymbol{x}$, with at least one column such that $\boldsymbol{b}_j \neq \boldsymbol{a}_j$. However, that means $\boldsymbol{b}_j \neq T(\boldsymbol{e}_j)$ and therefore $B\boldsymbol{e}_j \neq T(\boldsymbol{e}_j)$, which is a contradiction. □

The above lemma is significant because it means that to understand the properties of

linear maps we can study $m \times n$ matrices without loss of generality.

1.4 Operator norm for linear maps

The euclidean norm of a vector $\boldsymbol{x}$ in $\mathbb{R}^n$ is defined as:

$$||\boldsymbol{x}|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{\boldsymbol{x}^t \boldsymbol{x}}$$

In theoretical analysis it is also convenient to define a norm for a function, the so-called "operator norm". For example, it will allow us to show that the function $T(\boldsymbol{x}) = A\boldsymbol{x}$ is continuous. As we know, continuity is a desirable property for functions because it means that "small" disturbances in $\boldsymbol{x}$ will lead to small disturbances in $A\boldsymbol{x}$, and we can invoke many theorems in real analysis.

A generic challenge in defining an operator norm is that a function can be evaluated at multiple points (in fact, any point in $\mathbb{R}^n$) but we are only interested in summarizing the norm with a single scalar. How do we reduce this dimensionality? There are two generic solutions to this problem: (1) define a weighting scheme to aggregate the points; (2) consider the extremes. We will follow the latter, using a definition sometimes called the "maximum stretch":

$$||T|| := \sup_{\boldsymbol{x} \in \mathbb{R}^n : ||\boldsymbol{x}|| = 1} ||T(\boldsymbol{x})|| \quad \text{(Operator norm)} \tag{1.1}$$

Example:

$$A = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}, \quad ||A\boldsymbol{x}|| = \sqrt{(5x_1)^2 + x_2^2}$$

Different values of $\boldsymbol{x}$ lead to different norms $||A\boldsymbol{x}||$ in the target space. Which is the largest one? If we do not restrict the space of $\boldsymbol{x}$, we could take x_1, x_2 to infinity and therefore the norm would be unbounded. However, if we restrict attention to vectors that have unit norm we obtain a finite supremum. In this case it is possible to verify that $||T|| = 5$. (Try verifying this on your own. A diagram might help!) This is the "maximum" amount we can "stretch" a vector of unit length from the domain to the range.

1.4.1 Cauchy-Schwarz Inequality

We first restrict attention to the case with $n \times 1$ matrices (i.e. vectors). We will prove the well-known Cauchy-Schwarz (CS) inequality, which is of interest in its own right and will make the proof for general matrices considerably simpler. Our objective will be to impose bounds on how much a vector can be "stretched" by a matrix.

Definition 1.4.1. The inner product of two vector $\boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^n$ is the scalar $\boldsymbol{v}^t \boldsymbol{x}$.

The CS inequality states that the absolute value of the inner product is bounded by the multiplication of the norms.

Theorem 1.4.1 (Cauchy-Schwarz Inequality)**.** *Suppose that* $\boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^n$*. Then:*

$$|\boldsymbol{v}^t \boldsymbol{x}| \leq ||\boldsymbol{v}||\ ||\boldsymbol{x}|| \qquad \textit{(Cauchy Schwarz Inequality)}$$

Sometimes the (CS) inequality is given a geometric interpretation. For instance, in two dimensions $\boldsymbol{v}^t \boldsymbol{x} = ||\boldsymbol{v}||\ ||\boldsymbol{x}|| \cos(\theta)$, where θ is the angle between the vectors. The inner product is a measure of how closely aligned two vectors are. When they are parallel, $\theta = 0$, and the equation in the lemma becomes an equality.

The proof only uses the fact that a norm is non-negative (an inequality) which gives rise to the (CS) inequality by using a clever substitution.

Proof. Let $\lambda \in \mathbb{R}$ and construct a vector $\boldsymbol{z} = \boldsymbol{v} - \lambda \boldsymbol{x}$. Since $||\boldsymbol{z}||^2 \geq 0$ and $||\boldsymbol{z}||^2 = (\boldsymbol{v} - \lambda \boldsymbol{x})^t (\boldsymbol{v} - \lambda \boldsymbol{x})$, it follows that

$$\boldsymbol{v}^t \boldsymbol{v} - 2\lambda \boldsymbol{v}^t \boldsymbol{x} + \lambda^2 \boldsymbol{x}^t \boldsymbol{x} \geq 0$$

Assume WLOG that $\boldsymbol{x} \neq 0$. Consider a particular $\lambda = \frac{\boldsymbol{v}^t \boldsymbol{x}}{(\boldsymbol{x}^t \boldsymbol{x})}$,[1] thus

$$\boldsymbol{v}^t \boldsymbol{v} - 2\frac{(\boldsymbol{v}^t \boldsymbol{x})^2}{(\boldsymbol{x}^t \boldsymbol{x})} + \frac{(\boldsymbol{v}^t \boldsymbol{x})^2}{(\boldsymbol{x}^t \boldsymbol{x})} \geq 0$$

We can rearrange the above terms to show that $(\boldsymbol{v}^t \boldsymbol{x})^2 \leq (\boldsymbol{v}^t \boldsymbol{v})(\boldsymbol{x}^t \boldsymbol{x})$. Since $\boldsymbol{v}^t \boldsymbol{x} = ||\boldsymbol{v}^t \boldsymbol{x}||$ (scalar) and $\boldsymbol{v}^t \boldsymbol{v} = ||\boldsymbol{v}||^2, \boldsymbol{x}^t \boldsymbol{x} = ||\boldsymbol{x}||^2$ it follows that $||\boldsymbol{v}^t \boldsymbol{x}||^2 \leq ||\boldsymbol{v}||^2 ||\boldsymbol{x}||^2$. The result follows by taking the square root on both sides. □

1.4.2 Operator Norm Inequality

The operator norm of finite linear maps has two very useful properties, it is finite and we can define a Cauchy-Schwarz inequality.

Lemma 1.4.1. *Let* A *be an* $m \times n$ *matrix. If* $T(\boldsymbol{x}) = A\boldsymbol{x}$*, then*

1. $||T|| < \infty$ *(Finite operator norm)*
2. $||T(\boldsymbol{x})|| \leq ||T||\ ||\boldsymbol{x}||$ *for all* $\boldsymbol{x} \in \mathbb{R}^n$*. (Operator Cauchy-Schwarz Inequality)*

[1] Notice that this value of λ minimizes the quadratic equation on the left-hand-side.

Before we proceed with the proof, notice that the Cauchy-Schwartz inequality is a special case in which $A = \boldsymbol{v}^t$.

In the general case we can also provide a geometric interpretation. The ratio of the norms $||T(\boldsymbol{x})||/||\boldsymbol{x}||$ is the amount that the input vector is "stretched" by the linear map T. Therefore, the operator norm $||T||$ is the maximum amount that a vector can be "stretched". Informally, the first part of the lemma states that finite linear maps cannot "stretch" vectors too much.

Proof. We will show the first property. Let A_j be the j^{th} row vector of the matrix A. Recall that by the definition of matrix multiplication, that if $\boldsymbol{z} = A\boldsymbol{x}$, then $z_i = A_i\boldsymbol{x}$ (a scalar). This means that the norm of $\boldsymbol{z}$ is:

$$\sqrt{\sum_{i=1}^{m} z_i^2} = \sqrt{\sum_{i=1}^{m} ||A_i\boldsymbol{x}||^2}$$

Then we can use the Cauchy-Schwarz Inequality:

$$||T(\boldsymbol{x})|| = \sqrt{\sum_{i=1}^{m} ||A_i\boldsymbol{x}||^2} \leq \sqrt{\sum_{i=1}^{m} ||A_i||^2 \; ||\boldsymbol{x}||^2}$$

Consider the restriction that $||\boldsymbol{x}|| = 1$, then the expression simplifies to $||T(\boldsymbol{x})|| \leq C = \sqrt{\sum_{i=1}^{m} ||A_i||^2}$. The quantity $C < \infty$ because m, n are finite: we know that $||A_i|| < \infty$ and that there is a finite number of finite terms in the sum. This means that the supremum is bounded by a finite quantity.

To show the second part we use the scalar property of the Euclidean norm: $||\alpha\boldsymbol{z}|| = \alpha||\boldsymbol{z}||$, for all $\boldsymbol{z}$ and for all non-negative $\alpha \in \mathbb{R}$ (verify this as an exercise).

If $\boldsymbol{x}$ in nonzero, then we can normalize it in order to ensure that it has unit norm, that is $\tilde{\boldsymbol{x}} = \frac{\boldsymbol{x}}{||\boldsymbol{x}||}$. Then by definition of the operator norm, which is a supremum overall all unit vectors, $||A\tilde{\boldsymbol{x}}|| \leq ||T||$. To complete the proof it suffices to multiply either side by $||\boldsymbol{x}||$ and using the scalar property of the Euclidean norm:

$$||A\boldsymbol{x}|| = ||A\tilde{\boldsymbol{x}}|| \; ||\boldsymbol{x}|| \leq ||T|| \; ||\boldsymbol{x}||, \quad \forall \boldsymbol{x} \in \mathbb{R}^n \backslash \{0\}$$

where the first equality holds by rewriting

$$A\boldsymbol{x} = A\tilde{\boldsymbol{x}} \; ||\boldsymbol{x}||$$

and using the scalar property by taking $\alpha = ||\boldsymbol{x}||$ and $\boldsymbol{z} = A\tilde{\boldsymbol{x}}$.

Finally we need to make sure that the inequality also holds when $\boldsymbol{x}$ is equal to zero. This does indeed hold because $||A(\mathbf{0})|| = 0$ and the right hand side is also zero.

□

Remarks 1: We proved the lemma by only using the fact that the matrix A has finite dimensions. In the remainder of the course we will impose stronger assumptions on the matrices (e.g. invertibility, orthogonality, etc.) and it is important to emphasize that these properties are **not necessary** for the operator norm to be finite.

Remmark 2: Which linear maps do not have a finite operator norm? Some linear maps on functions spaces (though not all) instead of the Euclidean space. When the space has an infinite number of elements the main argument of our proof (finiteness) does not hold.

1.5 Continuity

In order to prove continuity of linear maps we will define a metric from the Euclidean norm.

$$d(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$$

Our objective in this section is to build up on the properties of the operator norm that we define in the previous section, in order to prove continuity of finite linear maps. The intuition of this result is as follows. A finite $||T||$ means that each vector in the domain is not "stretched" too much by the linear map. This means that "small" perturbations in the domain have "small" perturbations in the range.

Definition 1.5.1. A function $T(\boldsymbol{x})$ is continuous at point $\boldsymbol{x} \in \mathbb{R}^n$, if $\forall \epsilon > 0$ there exists $\delta > 0$ such that if $\boldsymbol{y} \in \mathbb{R}^n$ and $d(\boldsymbol{x}, \boldsymbol{y}) < \delta$, then $d(T(\boldsymbol{x}), T(\boldsymbol{y})) < \epsilon$.

We can now show that finite linear maps are continuous. It is worth emphasizing that this result has minimal assumptions, and does not depend on the specific form of the matrix A, just the fact that it has finite dimensions.

Theorem 1.5.1. *If $T(\boldsymbol{x}) = A\boldsymbol{x}$ then T is continuous.*

Proof. Fix $\epsilon > 0$. Notice that $d(T(\boldsymbol{x}), T(\boldsymbol{y})) = ||A\boldsymbol{x} - A\boldsymbol{y}|| = ||A(\boldsymbol{x} - \boldsymbol{y})||$. Define $\boldsymbol{w} = \boldsymbol{x} - \boldsymbol{y}$ and apply the Cauchy-Schwartz inequality in Lemma 1.4.1.

$$||A(\boldsymbol{x} - \boldsymbol{y})|| \leq ||T||\ ||\boldsymbol{x} - \boldsymbol{y}|| < \epsilon$$

Since $||T||$ is finite, then as long as $d(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}|| < \epsilon / ||T||$, the above inequality holds. We complete the proof by picking $\delta = \epsilon / ||T||$. □

1.6 Application: Markov Chains

Consider the following situation. At time 0, there is a population of individuals split between two states (e.g. two cities), with proportions $x_1, x_2 \in [0, 1]$ that up to one, $x_1 + x_2 = 1$. At time 1, some individuals decide to remain in each city (p_{11}, p_{22}, respectively) while others decide to migrate to the neighboring city (p_{12}, p_{21}, respectively). The process continues for T periods. The proportion of the population in each state, x_{it} for $i \in \{1, 2\}, t \in 0, \ldots, T$ forms a **Markov chain**. The term **Markov** refers to the fact that the state at time $t + 1$ only depends on the proportions at t and the transition probabilities, but not periods $t-1, t-2, \ldots$. See Figure 1.1 for an illustration of this process.

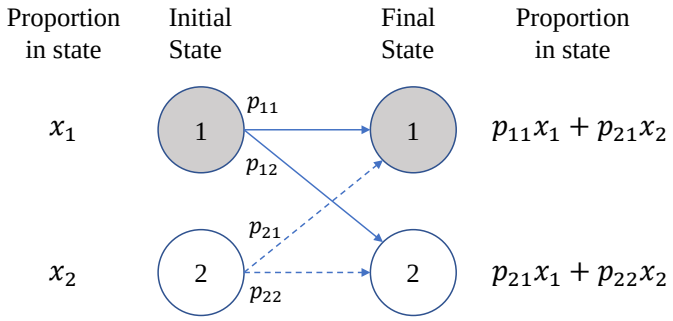

Figure 1.1: A finite markov chain

Linear algebra can help us answer important questions about Markov chains. What is the proportion of individuals in each state at time t? As $T \to \infty$, is there a stationary population in each state? Under what conditions? If there is a stationary state, does it depend on the initial distribution? To develop full answers to these questions we will need to develop more tools throughout the next chapters.

Finite markov chains (i.e. those with a finite number of states) can be represented using a stochastic matrix.

Definition 1.6.1. An $n \times n$ matrix P is a *stochastic matrix* if its entries are non-negative and the sum of the entries in each column adds up to one, $P_{ij} \geq 0, \sum_i P_{ij} = 1$.

Definition 1.6.2. A $n \times 1$ vector $\boldsymbol{x}$ is a *probability vector* if its entries are non-negative and add up to one, $x_i \geq 0, \sum_i \boldsymbol{x}_i = 1$.

It can be verified from the definition that the columns of a stochastic matrix are probability vectors. If P is a stochastic matrix and $\boldsymbol{x}_t$ is a vector of individuals in each state at

time t, then we can represent the process in Figure 1.1 as a law of motion:

$$\boldsymbol{x}_{t+1} = P\boldsymbol{x}_t$$

By recursive substitution we can show that:

$$\boldsymbol{x}_t = P\boldsymbol{x}_{t-1} = P(P\boldsymbol{x}_{t-2}) = \cdots = P^t\boldsymbol{x_0}$$

This makes it easier to analyze this type of markov chains because it means that we only need to understand the properties of the matrix P. We can also show that P^t is also a stochastic matrix.

Lemma 1.6.1. *Assume that P is a stochastic matrix, then*

1. *If the vector has $\boldsymbol{x} \in \mathbb{R}^n$ is a probability vector, then $\boldsymbol{y} = P\boldsymbol{x}$ is also a probability vector.*

2. *P^t is a stochastic matrix.*

Proof. We will prove the first part. By the definition of matrix multiplication $y_i = \sum_{j=1}^n P_{ij}x_j$. Since all the entries of the sum are non-negative, then $y_i \geq 0$. Furthermore, $\sum_{i=1}^n y_i = \sum_{i=1}^n \sum_{j=1}^n P_{ij}x_j$. We can rearrange this sum so that:

$$\sum_{i=1}^{n} y_i = \sum_{j=1} x_j \sum_{i=1}^{n} P_{ij} = \sum_{j=1} x_j = 1$$

That shows that $\boldsymbol{y}$ is a probability vector (entries are non-negative and add up to one).

Now we will prove the second part by induction. First we show the result for $t = 2$. By the definition of the multiplication of two matrices.

$$P^2 = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ Pp_1 & \cdots & Pp_n \\ \downarrow & \cdots & \downarrow \end{bmatrix}$$

By the first part of the lemma, since p_j is a probability vector then so is Pp_j. The columns of P^2 are all probability vectors therefore P^2 is stochastic. Now assume that P^t is stochastic. Using the first part of the lemma $P^t p_j$ is a probability vector, $1 \leq j \leq n$. That means that P^{t+1} is also stochastic. This completes the induction argument, showing that P^t is a stochastic matrix. □

1.7 Properties Appendix: Matrix Transpose

Let A be an $m \times n$ matrix, with entries $[A_{ij}]$. With this notation, define matrix addition as $[(A+B)_{ij}] := [A_{ij} + B_{ij}]$ and matrix multiplication as $[(AC)_{ij}] = [\sum_{l=1}^{n} a_{il}c_{lj}]$ for matrices $B_{m\times n}$ and $C_{n\times k}$.

Definition 1.7.1. The **transpose** of a matrix A is an $n \times m$ matrix with entries $[A_{ji}]$, which we denote A^t.

Definition 1.7.2. A matrix A is said to be symmetric if $A^t = A$.

$$A = \begin{bmatrix} 1 & 5 & 8 \\ 2 & 6 & 0 \end{bmatrix}, \quad A^t = \begin{bmatrix} 1 & 2 \\ 5 & 6 \\ 8 & 0 \end{bmatrix}$$

Lemma 1.7.1. *Let A, B be $m \times n$ matrices, C an $n \times k$ matrix and λ a scalar.*

1. *(Involution Property) $(A^t)^t = A$.*
2. *(Additive Separability) $(A+B)^t = A^t + B^t$.*
3. *(Multiplicative Separability) $(AC)^t = C^t A^t$.*
4. *(Scalar multiplication) $(\lambda A)^t = \lambda A^t$.*
5. *(Transpose of scalar) $A^t = A$ if $m = n = 1$.*
6. *(Bilinear form) $(A+B)^t(A+B) = A^tA + A^tB + B^tA + B^tB$.*

Proof of Properties. .

1. By definition $[A^t_{ij}] = [A_{ji}]$. Applying the definition again, we get $[(A^t)^t_{ij}] = [A^t_{ji}] = [A_{ij}]$.
2. By definition $[(A+B)^t]_{ij} = [(A+B)_{ji}] = [A_{ji} + B_{ji}] = [A^t_{ij} + B^t_{ij}]$.
3. By definition $[AC]_{ij} := [\sum_{l=1}^{n} a_{il}c_{lj}]$, where c_j are the columns of c. Then $[(AC)^t_{ij}] = [(AC)_{ji}] = [\sum_{l=1}^{n} a_{jl}c_{li}] = [\sum_{l=1}^{n} C^t_{il}A^t_{lj}] = [(C^tA^t)_{ij}]$.
4. By definition $[(\lambda A)_{ij}] = [\lambda A_{ij}] = \lambda[A_{ij}]$.
5. If $m = n = 1$, $[A^t_{11}] = [A_{11}]$.

6. First apply additive separability, $(A+B)^t = A^t + B^t$. Then use multiplicative separability of multiplication $(A^t + B^t)(A+B) = A^t(A+B) + B^t(A+B) = A^tA + A^tB + B^tA + B^tB$.

□

1.8 Exercises

1. Suppose that $T(\boldsymbol{x}) = A\boldsymbol{x}$ and that $F(y) = B\boldsymbol{y}$, with $A_{m\times n}$ and $B_{k\times m}$.

   (a) Show that $G = F(T(\boldsymbol{x}))$ is also a linear map.

   (b) Show that $||G|| \leq ||F||\ ||T||$. Is the composite of two linear maps continuous?

   (c) Assume that P is a square matrix. Use part (b) to show that for any non-negative integer t, $||P^t|| \leq ||P||^t$.

   (d) Show that if $\boldsymbol{x}$ is a probability vector, then $||\boldsymbol{x}|| > a$ for some $a > 0$.

   (e) If P is a stochastic matrix, could it be $||P|| < 1$? What would this imply for our migration example if it were true?

2. In this section you will expand some of the details of the proof of the Cauchy-Schwarz inequality. Let $\lambda \in \mathbb{R}, \boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^n$. We know that if $z = \boldsymbol{v} - \lambda\boldsymbol{z}$, $||\boldsymbol{z}|| \geq 0$, then

$$\boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x} \geq 0 \tag{1.2}$$

   (a) Show that the condition in Equation 1.2 is equivalent to:

$$\inf_{\lambda\in\mathbb{R}^n}\{\ \boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x}\ \} \geq 0, \quad \forall\boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^n$$

   (b) Consider the case when $||\boldsymbol{x}|| > 0$. Use the fact that the function is quadratic in λ to show that a minimum exists and that is

$$\frac{\boldsymbol{v}^t\boldsymbol{x}}{\boldsymbol{x}^t\boldsymbol{x}} = \arg\min_{\lambda\in\mathbb{R}^n}\{\ \boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x}\ \}$$

   (c) Show that if $\boldsymbol{v} = \boldsymbol{x}$, then Cauchy-Schwarz attains equality.

Chapter 2

Image and Kernel

Let A be an $m \times n$ matrix and let $T(\boldsymbol{x}) = A\boldsymbol{x}$. There are two main objects that we are interested in concerning linear maps.[1]

$$Im(A) := \{\boldsymbol{z} \in \mathbb{R}^m : \exists \boldsymbol{x} \in \mathbb{R}^n \text{ s.t. } \boldsymbol{z} = A\boldsymbol{x}\}$$

$$Ker(A) := \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \mathbf{0}_{m\times 1}\}$$

which we denote the **image** and **kernel** of the linear map, respectively. Notice that the image is a subset of the range of the function, whereas the kernel is a subset of its domain. Moreover, remember that m is not necessarily equal to n. Therefore, they typically reside in different spaces. However, we will show that they are related in other ways.

Other equivalent terms are used to describe these sets. The kernel is the set of solutions to a **homogenous** system of equations (one where the right hand size is the zero vector). The image of the linear map is also sometimes called the **span** of the columns of A.

2.1 Simple example

The image addresses what information is contained in A, whereas the kernel can help us determine whether its columns contain redundant or unique information. For now we will

[1] It is important to note that the image and kernel are defined with respect to the function T and not the matrix itself. However, because finite linear maps are uniquely represented by matrices by Lemma 1.3.2, we will define $Im(A)$ and $Ker(A)$ rather than $Im(T)$ and $Ker(T)$ in order to streamline the notation.

focus on the image because it is a little easier to analyze. Consider the following matrices

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

In econometrics, A could represent a data matrix with individuals in each row and variables in each column. We say that the two columns in A are **colinear** because one can be expressed as a linear combination of the other one. The usage in econometrics is identical to the one in linear algebra. We can verify that $Im(A) = Im(\tilde{A})$, because

$$A\boldsymbol{x} = \begin{pmatrix} x_1 + x_2 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{A}\boldsymbol{y} = \begin{pmatrix} y_1 \\ 0 \\ 0 \end{pmatrix}$$

where $\boldsymbol{x} \in \mathbb{R}^2$ and $\boldsymbol{y} \in \mathbb{R}^1$. The mapping defined by each matrix has a different domain but their image is the same, a line in $\mathbb{R}^3$ along the first coordinate. The key idea is that since the two columns are identical we can drop one of them and produce a linear map that contains the same information about the image. In the econometrics example, it means that we do not need to include the same variable twice.

On the other hand, the matrices A and $\tilde{A}$ have a different kernel.

$$Ker(A) = \left\{ \begin{bmatrix} \alpha \\ -\alpha \end{bmatrix} : \alpha \in \mathbb{R} \right\}, \quad Ker(\tilde{A}) = \left\{ 0 \right\}$$

The kernel always includes the zero vector because $A_{m\times n}\mathbf{0}_{n\times 1} = \mathbf{0}_{m\times 1}$ regardless of A. As $Ker(\tilde{A})$ shows, when there is no redundant information, the kernel only contains that element (we call this the *trivial kernel*). However, when the vectors are colinear the kernel is non-trivial. As a matter of fact, if the kernel is not trivial it has infinite solutions, as shown in $Ker(A)$.

2.2 Full rank matrices

As with other types of functions, linear maps have a unique solution (if it exists) when they are one-to-one (injective). For finite-dimensional linear maps, we say that the matrix is **full rank**.

Definition 2.2.1. A matrix $A_{m\times n}$ is said to be full rank if $T = A\boldsymbol{x}$ is one-to-one.

It is important to observe that not all full rank matrices are square (see example in the previous section). One of the special things about a linear map is that many of its properties can be obtained from its kernel.

Lemma 2.2.1. *The linear map $T : \mathbb{R}^n \to Im(A)$ is injective if and only if the kernel is trivial, $Ker(A) = \{\mathbf{0}_{n\times 1}\}$.*

Proof. ($\Longrightarrow$) Suppose that T is one-to-one, then $A\boldsymbol{x} = 0$ has a unique solution, if it exists. We know $\mathbf{0}_n$ is a solution since $A\mathbf{0}_n = \mathbf{0}_m$, hence it is the only solution. This means that the kernel exactly contains the zero vector.

($\Longleftarrow$) We will show this by contradiction. Suppose that $Ker(A) = \{0\}$ and that T is not one-to-one. Then there exists $\boldsymbol{x} \neq \boldsymbol{y}$ such that $A\boldsymbol{x} = A\boldsymbol{y}$, which implies that $A(\boldsymbol{x} - \boldsymbol{y}) = 0$. But this implies that $(\boldsymbol{x} - \boldsymbol{y}) \neq 0$ belongs to the kernel, which is a contradiction. □

Lemma 2.2.1 gives a characterization of injective linear maps in terms of the kernel. A kernel that contains a single element, the zero vector $\mathbf{0} \in \mathbb{R}^n$ is sometimes called a **trivial kernel**. This simplifies our task of identifying injective functions considerably. At this stage we are focusing on understanding the high-level properties of the kernel because we are interested in proving existence and constructive results. Later in the course we will discuss practical (numeric) ways to test whether the kernel is trivial.

Corollary 2.2.1. *The following implications follow from Lemma 2.2.1*

1. *If $Ker(A) = \{\mathbf{0}_{n\times 1}\}$, then all its columns are non-zero vectors.*
2. *If $Ker(A) = \{\mathbf{0}_{n\times 1}\}$ and $v \in \mathbb{R}^m$ is not in the image of A, then $Ker(a_1, \ldots, a_n, v) = \{\mathbf{0}_{(n+1)\times 1}\}$.*
3. *$Ker(A) = \{\mathbf{0}_{n\times 1}\}$ if and only if each column vector a_j is non-zero and cannot be expressed as a linear combination of the other vectors, $\{a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n\}$, $1 \leq j \leq n$.*

The Corollary proves equivalent characterizations of trivial kernels.

Proof. We can also prove the following implications of the theorem:

1. Suppose that $Ker(A) = \{\mathbf{0}_{n\times 1}\}$. Suppose by contradiction that some columns are zero, and WLOG, say, that a_1 is a zero vector. Then $\boldsymbol{x} = (\alpha, 0, \ldots, 0)$ is a solution to $A\boldsymbol{x} = \mathbf{0}$, for any $\alpha \in \mathbb{R}$. That means that $K(A) \neq \{\mathbf{0}\}$ because it contains more elements (at least $\boldsymbol{x}$), which is a contradiction.

2. Suppose that $Ker(A) = \{\mathbf{0}_{n\times 1}\}$, $v \in \mathbb{R}^n$ is not in the image of A, and $Ker(a_1, \ldots, a_n, v) \neq \{\mathbf{0}_{(n+1)\times 1}\}$. Then there exists $\beta_1, \ldots, \beta_n, \beta_v$ not all zero, such that.

$$\mathbf{0}_{m\times 1} = \beta_1 a_1 + \ldots + \beta_n a_n + \beta_v v$$

   The case $\beta_v = 0$ and some non-zero $(\beta_1, \ldots, \beta_n)$ is not possible because then $Ker(A) \neq \{0\}$. The case $\beta_v \neq 0$ is also not possible because then $v = -(\beta_1/\beta_v)a_1 - \ldots - (\beta_n/\beta_v)a_n$ and this contradicts the fact that $v \notin Im(A)$. Since this situation contradicts both premises, it follows that $Ker(a_1, \ldots, a_n, v) = \{\mathbf{0}_{(n+1)\times 1}\}$.

3. ($\Longrightarrow$) Suppose that $Ker(A) = \{\mathbf{0}\}$ and by the first part of the lemma, all the column vectors of A have to be non-zero. Suppose that some column can be expressed as a linear combination of the other vectors, WLOG, say, a_1:

$$a_1 = \beta_2 a_2 + \beta_3 a_3 + \ldots + \beta_n a_n.$$

   At least one $\beta_i,\ 2 \leq i \leq n$ is non-zero, otherwise a_1 would zero (which is ruled out by the first part of the lemma). We can rearrange this equation as

$$\mathbf{0}_{m\times 1} = -a_1 + \beta_2 a_2 + \beta_3 a_3 + \ldots + \beta_n a_n$$

   In matrix form that means that $A\boldsymbol{x} = \mathbf{0}$, where $\boldsymbol{x} = (-1, \beta_2, \beta_3, \ldots, \beta_n)$ and at least one $\beta_i \neq 0$. However, this means that $\boldsymbol{x} \in Ker(A)$ and therefore $Ker(A) \neq \{\mathbf{0}_{n\times 1}\}$, a contradiction.

   ($\Longleftarrow$) Construct the matrix A sequentially by adding columns. Since a_1 is non-zero, $A_1 = a_1$ has a trivial kernel. Add each column sequentially. By assumption column a_{k+1} cannot be expressed as a linear combination of the first k columns. Therefore $a_{k+1} \notin Im(a_1, \ldots, a_k)$ and we can apply the second part of the lemma. We can apply this argument sequentially, ensuring that the kernel is trivial at each stage until we have added all the columns.

□

2.2.1 Surjective Maps

However, an injective function does not necessarily have a solution to the system $A\boldsymbol{x} = \boldsymbol{b}$ for $\boldsymbol{b} \in \mathbb{R}^m$. The function is guaranteed a solution if it is both injective and $Im(A) = \mathbb{R}^m$ (it is surjective over the Euclidean space). Otherwise, it only has a solution if $b \in Im(A)$. We prove an important lemma that allows us to assess whether a linear map is surjective or not.

Lemma 2.2.2. *Suppose that $A_{m\times n}$ is full rank and that we have a matrix $W_{m\times K}$, whose columns, $w_1, \ldots, w_K$ are contained in $Im(A)$. It follows that,*

1. *If $K = n$ and W is full rank, then $Im(W) = Im(A)$.*
2. *If $K > n$, then W cannot be full rank.*

The lemma states that all funk rank matrices that span the same space (have the same image) necessarily have the same dimensions. If a full rank matrix spans a space we say that the column vectors are a **basis** for the space. The basis for a space is not generally unique. For example, scaling one of the columns leads to the same image. The **rank** of a space is the number of columns of any basis that spans it. We use the convention that $rank(A) = 0$ if A is the zero matrix.

The second part of the lemma states that the number of columns is always greater than or equal to the rank of the matrix. If a matrix has more columns than its rank then it is necessarily **rank deficient** (not injective). This is a formal proof that a linear system with more unknown than equations cannot have a unique solution, if it exists.

The proof of the lemma is interesting because it uses a **recursive substitution** argument. It starts of with the matrix A and substitutes each vector sequentially until we have have shown that the image spanned by both basis is the same. At each step the image is the same. We will use this type of argument in subsequent proofs because it useful to prove **existence** of bases.

Proof of Lemma 2.2.2. Assume W is full rank. Let $a_1, \ldots, a_n$ be the columns of A. Since w_1 belongs to the image of A,

$$w_1 = \beta_1 a_1 + \cdots + \beta_n a_n$$

There must exist some $1 \leq i \leq n$ such that $\beta_i \neq 0$, otherwise $w_1 = \mathbf{0}$ contradicting that W is full rank. Assume without loss of generality, say it is β_1. Hence

$$a_1 = \frac{1}{\beta_1} w_1 - \frac{\beta_2}{\beta_1} a_2 - \cdots - \frac{\beta_n}{\beta_1} a_n.$$

Construct a new matrix that substitutes a_1 with w_1. This implies that the image of the new matrix is equal to the image of A, that is $Im\{w_1, a_2, \cdots, a_n\} = Im(A)$, because any combination of the original vectors can be written indirectly in terms of $\{w_1, a_2, \cdots, a_n\}$. Since $w_2 \in Im(A)$, it is a linear combination of $w_1, a_2, \cdots, a_n$:

$$w_2 = \gamma_1 w_1 + \beta_2' a_2 + \cdots + \beta_n' a_n.$$

Again, there must exist some $2 \leq i \leq n$ such that $\beta_i' \neq 0$ (using the fact that W is assumed full rank). Without loss of generality, assume it is a_2, then

$$a_2 = \frac{1}{\beta_2'} w_2 - \frac{\gamma_1}{\beta_2'} w_1 - \frac{\beta_3'}{\beta_2'} a_3 - \cdots - \frac{\beta_n'}{\beta_2'} a_n$$

This implies $Im\{w_1, w_2, a_3, \cdots, a_n\} = Im(A)$. Continuing in this fashion, we can show $Im\{w_1, \cdots, w_n\} = Im(A)$. This proves the first part of the lemma.

For the second part, if $K > n$ and assume W is full rank. Consider a submatrix of W with its first n columns $\tilde{W} = (w_1, w_2, ..., w_n)$. From part one we know $Im(\tilde{W}) = Im(A)$. By assumption, $w_{n+1} \in Im(A) = Im(\tilde{W})$ and hence it is a linear combination of $w_1, \cdots, w_n$, which contradicts the assumption that $Ker(W) = \{0\}$ (i.e. full rank). □

Example (Injective but not surjective linear maps).

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The linear map $T(\boldsymbol{x}) = A\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}$ is injective but it is not subjective on $\mathbb{R}^2$ because $\boldsymbol{y} \notin Im(A)$. As an exercise try to show that Lemma 2.2.2 implies that an injective map cannot be surjective ($Im(A) = \mathbb{R}^m$) if A is a non-square full rank matrix. (Hint: Assume that the columns of $W = I_{m \times m}$ are contained in $Im(A)$).

Example (Square matrix that is not injective).

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad \boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

We can verify that $A\boldsymbol{x}_1 = A\boldsymbol{x}_2$, which means that the function $T(\boldsymbol{x}) = A\boldsymbol{x}$ is not injective.

2.2.2 Invertible Matrices

Lemma 2.2.3. *If W is a full rank $n \times n$ matrix then there exists a unique matrix W^{-1} such that $W^{-1}W = I_n$.*

Proof. This proof has two steps:

1. *Show that the linear map $T(x) = W\boldsymbol{x}$ is injective and surjective. That means that we can define an inverse function $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ for every element in $\mathbb{R}^n$.*

   The fact that W is full rank guarantees that the function is injective. Now we will show that full rank square matrix are also surjective. Consider the identity matrix $I_{n\times n}$. We know that $Im(I_{n\times n}) = \mathbb{R}^n$ because every for every vector $\boldsymbol{x} \in \mathbb{R}^n$ can be written as $\boldsymbol{x} = I_{n\times n}\boldsymbol{x}$. Clearly every column in W belongs to $Im(I_{n\times n})$. We can apply Lemma 2.2.2 to show that $Im(W) = \mathbb{R}^n$.

2. *Show that the function T^{-1} is a finite linear map, because then we can apply the representation theorem in Lemma 1.3.2 to show that there exists a matrix W^{-1} such that $T^{-1}(\boldsymbol{y}) = W^{-1}\boldsymbol{y}$.*

   Choose two arbitrary $n \times 1$ vectors $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$. Since T^{-1} is injective and surjective, there exist unique vectors $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ such that $\boldsymbol{y}_1 = W\boldsymbol{x}_1$ and $\boldsymbol{y}_2 = W\boldsymbol{x}_2$. It follows that $\alpha\boldsymbol{y}_1 + \beta\boldsymbol{y}_2 = \alpha W\boldsymbol{x}_1 + \beta W\boldsymbol{x}_2 = W(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2)$ for all $\alpha, \beta \in \mathbb{R}^n$. This implies that $T^{-1}(\alpha\boldsymbol{y}_1 + \beta\boldsymbol{y}_2) = \alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2 = \alpha T^{-1}(\boldsymbol{y}_1) + \beta T^{-1}(\boldsymbol{y}_2)$ (it is a linear map over a finite domain and range). To complete the proof we use Lemma 1.3.2 to show that there exists a unique matrix, denoted by W^{-1}, such that $T^{-1}(\boldsymbol{y}) = W^{-1}\boldsymbol{y}$.

□

2.3 Rank deficient matrices

Lemma 2.3.1. *Let A be a $m \times n$ matrix such that $Ker(A) \neq \{\mathbf{0}\}$. Then (i) there exists an $n \times k$ full rank matrix, Ω, such that $Ker(A) = Im(\Omega)$. (ii) All the matrices Ω with this property have the same dimensions.*

Lemma 2.3.1 states the kernel of **rank deficient** (not full rank) matrices can be represented by a matrix. The proof follows a **basis extension** argument. Start the matrix by choosing a non-zero vector in the kernel. Then sequentially add independent vectors (if they exist). The proof doesn't provide an algorithm for finding such vectors, although such algorithms do exist. The essential part of the proof is to show that there cannot be more than n independent vectors in the kernel, because $Ker(A) \subseteq \mathbb{R}^n$. The second part of the lemma says that every basis for the kernel has the same number of columns, which we call $dim(Ker(A))$.

Proof. Since $Ker(A) \neq \{\mathbf{0}\}$ there exists a non-zero vector such that $\omega_1 \in Ker(A)$. Construct a candidate matrix $\Omega_1 = \omega_1$. First we show that $Im(\Omega_1) \subseteq Ker(A)$. If $\boldsymbol{x} = \alpha\omega_1, \alpha \in \mathbb{R}$ then $\boldsymbol{x} \in Ker(A)$ because $A\boldsymbol{x} = A(\alpha\omega_1) = \alpha(A\omega_1) = \mathbf{0}$. Now suppose that we have constructed a full rank matrix Ω_k with k column vectors whose image is contained in $Ker(A)$, that is:

$$Im(\Omega_k) \subseteq Ker(A), \quad Ker(\Omega_k) = \{\mathbf{0}_{k\times 1}\}$$

If $Im(\Omega_k) = Ker(A)$ then we are done. Otherwise, choose $\omega_{k+1} \in Ker(A)$ such that $\omega_{k+1} \notin Im(\Omega_k)$. By Lemma 2.2.1 then the matrix $\Omega_{k+1} = \{\Omega_k, \omega_{k+1}\}$ (appending a column vector on the right) also has a trivial kernel (is full rank). We also need to show that $Im(\Omega_{k+1}) \subseteq Ker(A)$. Let $\boldsymbol{x} = (\boldsymbol{x}_{1:k}, x_{k+1}) \in \mathbb{R}^k$, where $\boldsymbol{x}_{1:k} = (x_1, \ldots, x_k)$. Since $Im(\Omega_k) \subseteq Ker(A)$, it follows that $A\Omega_{k+1}\boldsymbol{x} = A\Omega_k\boldsymbol{x}_{1:k} + A\omega_{k+1}x_{k+1} = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Therefore, $Im(\Omega_{k+1}) \subseteq Ker(A)$.

This process needs to stop eventually (i.e. $Im(\Omega_k) = Ker(A)$ for some k). Suppose that it doesn't and that we are at a stage where $k \geq n$ and Ω_k is full rank. Since $Im(\Omega_k) \subseteq \mathbb{R}^n = Im(I_n)$ then by Lemma 2.2.2 if $k > n$, Ω_k cannot be full rank, a contradiction. Hence, the process needs to stop for some $k \leq n$.

Now suppose that we choose two full rank matrices Ω_1, Ω_2 such that $Im(\Omega_1) = Im(\Omega_2) = Ker(A)$. They must have the same number of rows because they span the same space. Now, WLOG assume that Ω_1 has strictly more columns than Ω_2. Since the columns of Ω_1 belong to $Im(\Omega_2)$ then Lemma 2.2.2 implies that Ω_1 is not full rank, a contradiction. Therefore, both matrices must have the same number of columns.

□

2.4 Application: Linear Regression

A researcher has access to a database with information about n individuals. There is a vector of outcomes $\boldsymbol{Y} \in \mathbb{R}^n$, each entry represents the outcomes for different individuals. There is a matrix with k explanatory variables $\boldsymbol{X}_{n\times k}$ called the **design matrix**, with $n > k$ (there are more observations than variables). The relationship between the outcome and the explanatory variables is given by:

$$\boldsymbol{Y}_{n\times 1} = \boldsymbol{X}_{n\times k}\beta_{k\times 1} + \epsilon_{n\times 1}$$

where $\epsilon_{n\times 1}$ is a vector of unexplained error terms and $\beta \in \mathbb{R}^k$ is a vector of coefficients. The researcher is interested in estimating the coefficient using the observed data to recover some effect of interest. The researcher has established that he will use the least-square-error criterion to compute the estimator (more details in later chapters) which leads to the following first-order condition.

$$(\boldsymbol{X}^t\boldsymbol{X})\hat{\beta} = \boldsymbol{X}^t Y \tag{2.1}$$

The $k \times k$ square matrix $\boldsymbol{X}^t\boldsymbol{X}$ is called the **gram matrix**. By Lemma 2.2.3 the function $T(\hat{\beta}) = (\boldsymbol{X}^t\boldsymbol{X})\hat{\beta}$ is injective and surjective as long as the gram matrix is full rank. This guarantees that the estimator exists and is unique.

Lemma 2.4.1. *Let $\boldsymbol{X}$ be an $n \times k$ matrix. Then $(\boldsymbol{X}^t\boldsymbol{X})$ is full rank if and only if $\boldsymbol{X}$ is full rank.*

Proof. We will show that $Ker(\boldsymbol{X}^t\boldsymbol{X}) = Ker(\boldsymbol{X})$. Consequently, by Lemma 2.2.1, either both matrices are full rank or neither of them is.

($\Longleftarrow$) Suppose that $\beta \in Ker(\boldsymbol{X})$, then $\boldsymbol{X}\beta = \boldsymbol{0}_{n\times 1}$. That means that $(\boldsymbol{X}^t\boldsymbol{X})\beta = \boldsymbol{0}_{k\times 1}$ and that $\beta \in Ker(\boldsymbol{X}^t\boldsymbol{X})$.

($\Longrightarrow$) Suppose that $\beta \in Ker(\boldsymbol{X}^t\boldsymbol{X})$ then $\boldsymbol{X}^t\boldsymbol{X}\beta = \boldsymbol{0}_{m\times 1}$. This also means that $\beta^t\boldsymbol{X}^t\boldsymbol{X}\beta = 0_{1\times 1} = (\boldsymbol{X}\beta)^t(\boldsymbol{X}\beta) = ||\boldsymbol{X}\beta||^2$. We know that a norm is equal to zero if and only if the vector is zero. Therefore $\boldsymbol{X}\beta = \boldsymbol{0}_{n\times 1}$ and $\beta \in Ker(\boldsymbol{X})$.

□

Lemma 2.4.1 states the gram matrix is full rank if and only if the design matrix is full rank. The proof is interesting because it illustrates that two matrices can have the same kernel even if they have a different number of rows (because the kernel is contained in the domain, which depends only on the number of columns). Empirically the result is interesting because it allows the researcher to asses the rank condition of $\boldsymbol{X}$ rather than the matrix X^tX which can be a more complicated object.

Lemma 2.4.2. *Let $\boldsymbol{X}$ be an $n \times k$ matrix with non-zero entries. (i) If $\boldsymbol{X}$ is not full rank then we can construct a full rank matrix $\boldsymbol{X}^*$ by dropping select columns of $\boldsymbol{X}$, such that $Im(\boldsymbol{X}) = Im(\boldsymbol{X}^*)$. (ii) Suppose that A is a full rank matrix such that $Im(A) = Im(\boldsymbol{X})$. Then A has the same number of columns as $\boldsymbol{X}^*$.*

Proof. By Lemma 2.2.2 if $\boldsymbol{X}$ is not full rank, then at least one of its columns is zero and/or can be expressed as a nonlinear combination of the others. Assume WLOG that it is the last column. Suppose that we construct $\boldsymbol{X}^*$ by dropping that column. Then $\boldsymbol{x}_k = \boldsymbol{X}^*\psi$, where $\boldsymbol{x}_k$ is the last column and ψ is a $(k-1) \times 1$ vector. We show that $Im(\boldsymbol{X}^*) = Im(\boldsymbol{X})$. By definition, $Im(\boldsymbol{X}^*) \subseteq Im(\boldsymbol{X})$ because there are fewer spanning vectors (we can always set the last coefficient to zero). The difficult part is showing that $Im(\boldsymbol{X}) \subseteq Im(\boldsymbol{X}^*)$. Suppose that $\boldsymbol{z} \in Im(\boldsymbol{X})$, then there exists a $\beta = (\beta_{1:(k-1)}, \beta_k) \in \mathbb{R}^k$ such that $\boldsymbol{z} = \boldsymbol{X}\beta$. We can rewrite this as $\boldsymbol{X}\beta = \boldsymbol{X}^*\beta_{1:(k-1)} + \boldsymbol{x}_k\beta_k$ and substitute in $\boldsymbol{x}_k = \boldsymbol{X}^*\psi$,

$$\boldsymbol{X}^*\beta_{1:(k-1)} + \boldsymbol{X}^*\psi\beta_k = \boldsymbol{X}^*(\beta_{1:(k-1)} + \psi\beta_k) \in Im(\boldsymbol{X}^*)$$

This proves that $Im(\boldsymbol{X}^*) = Im(\boldsymbol{X})$. If the matrix $\boldsymbol{X}^*$ is full rank then we are done. If it is not we can repeat the process all over again. The process stops eventually because $\boldsymbol{X}$ is non-zero, meaning that it has at least one non-zero column vector. That means that it is indeed feasible to obtain a full rank matrix by dropping columns (in the extreme case we are just left with one vector). The second part directly follows from Lemma 2.2.2.

□

2.5 Properties Appendix: Block-Partitioned Matrices

Suppose that X is an $m \times n$ matrix. Then the matrix can be represented in block partition form.

$$X = \begin{bmatrix} A_{m_1 \times n_1} & B_{m_1 \times n_2} \\ C_{m_2 \times n_1} & D_{m_2 \times n_2} \end{bmatrix}$$

where $m_1 + m_2 = m$ and $n_1 + n_2 = n$. The matrices A, B, C, D are submatrices of the column X. In general, we could have more partitions of the matrix or less (only A and C, or only A and B). The best way to partition a matrix depends on what the researcher wants to prove about that matrix.

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad \Longrightarrow \quad A = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & 5 \end{bmatrix}, \quad D = \begin{bmatrix} 6 \end{bmatrix}.$$

For example:

$$X = \begin{bmatrix} X_1 & X_2 \\ X_3 & X_4 \end{bmatrix}, \quad Z = \begin{bmatrix} Z_1 & Z_2 \\ Z_3 & Z_4 \end{bmatrix}$$

Transpose

Notice that the transpose changes the position of the blocks.

$$X^t = \begin{bmatrix} X_1^t & X_3^t \\ X_2^t & X_4^t \end{bmatrix}$$

Matrix Multiplication Suppose that $X \in \mathbb{R}^m \times \mathbb{R}^n$ and $Z \in \mathbb{R}^n \times \mathbb{R}^k$. Furthermore suppose that $X_1 \in \mathbb{R}^{m_1} \times \mathbb{R}^{n_1}$ and $Z_1 \in \mathbb{R}^{n_1} \times \mathbb{R}^{k_1}$ (matrices are conformable). Then we can defined block partitioned multiplication.

$$XZ = \begin{bmatrix} X_1 Z_1 + X_2 Z_3 & X_1 Z_2 + X_2 Z_4 \\ X_3 Z_1 + X_4 Z_3 & X_3 Z_2 + X_4 Z_4 \end{bmatrix}$$

Matrix Addition Suppose that $X, Z \in \mathbb{R}^m \times \mathbb{R}^n$. Furthermore suppose that $X_1 \in \mathbb{R}^{m_1} \times \mathbb{R}^{n_1}$ and $Z_1 \in \mathbb{R}^{m_1} \times \mathbb{R}^{n_1}$ (matrices are conformable). Then we can defined block partitioned addition.

$$XZ = \begin{bmatrix} X_1 + Z_1 & X_2 + Z_2 \\ X_3 + Z_3 & X_4 + Z_4 \end{bmatrix}$$

2.6 Exercises

1. Suppose that X is a non-zero $m \times n$ rank deficient matrix. Suppose that we partition its columns $X = [X_1, X_2]$ in such a way that $Im(X_1) = Im(X)$ and X_1 is full rank. The block matrices have n_1, n_2 columns, respectively. This is equivalent to dropping redundant variables in a linear regression.

   (a) Show that Equation 2.1 can be written in block-partitioned form as:

   $$\begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix} \beta = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   (b) Suppose that $\hat{\beta}_1 = (X_1^t X_1)^{-1}(X_1^t Y)$. Construct a vector $\beta^* = \begin{bmatrix} \hat{\beta}_1 \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}$. Show that β^* is a solution to Equation 2.1 if and only if $X_2^t X_1 \hat{\beta}_1 = X_2^t Y$.

   (c) Verify that the columns of X_2 belong in $Im(X_1)$. Use this fact to show that $X_2^t X_1 \hat{\beta}_1 = X_2^t Y$.

   (d) Consider the data matrix,

   $$X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

   Construct $X^t X$ and $X^t Y$. Now partition the matrix into X_1, X_2 and compute β^*. Verify that the results that you proved above are true for the following cases:

   (i) Construct X_1 using columns 1 and 2.

   (ii) Construct X_1 using columns 1 and 3.

   (e) Is β^* the same in both exercises? How can we interpret the result?

   You can use the fact that the inverse of a 2×2 matrix is given by:

   $$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \implies A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Chapter 3

Orthogonality

In the previous chapter we studied the (full) rank of matrices, which determines whether linear systems of equations have unique solutions, if they exist. Unfortunately, there are many practical cases where exact solutions do not exist because the matrix is not surjective over the Euclidean space. In this section, we will explore a novel property, **orthogonality**, which will address the optimality of approximate solutions. The main takeaway of this chapter is that linear systems of equations always have an "approximate" solution which is unique if and only if the matrix is full rank. Therefore, the rank of the matrix continues to play a central (and coherent) role in establishing uniqueness, even in this novel setting.

The second takeaway of this chapter is that the residuals of approximate solutions can be characterized in terms of **projection matrices**, which are square matrices with well-defined properties. In the application we combine projection matrices with the concept of block-partitioning in order to gain new insights about multivariate linear regressions.

Our analysis starts with the definition of orthogonality.

Definition 3.0.1. A pair of vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ is **orthogonal** ($\boldsymbol{x} \perp \boldsymbol{y}$) if $\boldsymbol{x}^t\boldsymbol{y} = 0$. If in addition, they have unit norm, $||\boldsymbol{x}|| = ||\boldsymbol{y}|| = 1$, then they are **orthonormal**.

In $\mathbb{R}^2$ two vectors are orthogonal to each other if they are perpendicular to each other. Orthogonality is the concept that generalizes this notion to higher dimensions.

3.1 Vector Orthogonalization

In this section we address an issue of "orthogonalizing" a vector: transforming it so that it is orthogonal to every vector in an auxiliary matrix. We deliberately use the same terminology as the linear regression example in Section 2.4. We focus on full rank matrices first because the results are constructive and will help us acquire intuition for proving similar existence results in the general case.

Lemma 3.1.1. ***(Vector Orthogonalization)*** *Let $\boldsymbol{X}$ be an $m \times n$ full rank matrix and let $\boldsymbol{y} \in \mathbb{R}^m$. Define a vector of coefficients $\beta^* = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y} \in \mathbb{R}^n$ and define the residual $\boldsymbol{\varepsilon} = \boldsymbol{y} - \boldsymbol{X}\beta^* \in \mathbb{R}^m$. Then*

1. $\boldsymbol{\varepsilon}^t\boldsymbol{X} = \boldsymbol{0}_{1\times n}$.
2. $Im(\boldsymbol{X}, \boldsymbol{\varepsilon}) = Im(\boldsymbol{X}, \boldsymbol{y})$.

The vector ε is sometimes known as the **residual**, the difference between the original vector $\boldsymbol{y}$ and a vector $\boldsymbol{X}\beta^*$, the **projection** onto the image of $\boldsymbol{X}$. Notice that the vector β^* is the solution to the system of equations for a linear regression in Section 2.4. Part 1 of the lemma states that the resulting vector $\boldsymbol{\varepsilon}$ is pairwise orthogonal to all the columns of A. Part 2 states that if a matrix has columns $(\boldsymbol{X}, \boldsymbol{y})$ then we can substitute the last column with an "orthogonalized" residual and still span the same space. Intuitively, this says there is no loss of information in "projecting out" its component in $\boldsymbol{X}$. This will be important for our theoretical analyses.

Proof. Since $\boldsymbol{X}$ is full rank, by Lemmas 2.4.1 and 2.2.3 $\boldsymbol{X}^t\boldsymbol{X}$ is invertible. Therefore β^* exists and is well-defined.

(i) By definition $\boldsymbol{\varepsilon}^t\boldsymbol{X} = (\boldsymbol{y}^t - \beta^{*t}\boldsymbol{X}^t)\boldsymbol{X}$. By definition of $\beta^* = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y}$ we can rewrite this as $\boldsymbol{y}^t(I - \boldsymbol{X}(\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t)\boldsymbol{X}$. By expanding out the terms we can show that this is equal to $\boldsymbol{y}^t(\boldsymbol{X} - \boldsymbol{X}) = \boldsymbol{0}_{1\times n}$.

(ii) Since $\boldsymbol{\varepsilon} = -\boldsymbol{X}\beta^* + \boldsymbol{y}$ it follows that $\boldsymbol{\varepsilon} \in Im(\boldsymbol{X}, \boldsymbol{y})$. Similarly, since $\boldsymbol{y} = \boldsymbol{X}\beta^* + \boldsymbol{\varepsilon}$ it follows that $\boldsymbol{y} \in Im(\boldsymbol{X}, \boldsymbol{\varepsilon})$. That means that any vector linear combination of $\boldsymbol{X}$ and $\boldsymbol{y}$ can be expressed indirectly in terms of $\boldsymbol{X}$ and $\boldsymbol{\varepsilon}$, and vice versa. Therefore, $Im(\boldsymbol{X}, \boldsymbol{\varepsilon}) = Im(\boldsymbol{X}, \boldsymbol{y})$.

□

3.1.1 Optimality of Approximate Solutions

In cases where the system of equations $\boldsymbol{y} = \boldsymbol{X}\beta$ does not have a solution we can define an optimality criterion based on the squared norm of the residual vector (also known as the residual sum of squares or SSR).

$$\text{SSR}(\beta) = ||\boldsymbol{y} - \boldsymbol{X}\beta||^2$$

We will prove that a unique minimizer exists when $\boldsymbol{X}$ is full rank. We can also show that if $\boldsymbol{X}$ is not full rank a (necessarily non-unique) minimizer exists but in order to do so, we need results from the next sections. We start off with the case of full rank matrices because they can help us develop the intuition of what are the key elements that we need for the general case. Furthermore, the full rank case is of interest in its own right for econometric applications.

Define $\hat{\boldsymbol{y}} = \boldsymbol{X}\beta^*$ as our candidate prediction vector using the coefficients $\beta^* = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y}$ that we outlined in the previous section. Then we can decompose vector $\boldsymbol{y}$ into two components that are orthogonal: (i) one that is linearly spanned (modeled) by the regressors $\boldsymbol{X}$ and (ii) the other that cannot be linearly modeled by $\boldsymbol{X}$, a residual $\boldsymbol{\varepsilon} = \boldsymbol{y} - \hat{\boldsymbol{y}}$.

$$\boldsymbol{y} = \underbrace{\hat{\boldsymbol{y}}}_{Projection} + \underbrace{(\boldsymbol{y} - \hat{\boldsymbol{y}})}_{OrthogonalProjection}$$

Lemma 3.1.2. *Let $\boldsymbol{X}$ be an $m \times n$ full rank matrix and let $\boldsymbol{y} \in \mathbb{R}^m$. Define $\boldsymbol{\varepsilon}$ and β^* as in Lemma 3.1.1. Then*

1. $SSR(\beta) = \boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon} + (\beta^* - \beta)^t\boldsymbol{X}^t\boldsymbol{X}(\beta^* - \beta)$.
2. *β^* is the unique minimizer of $SSR(\beta)$.*

Proof. We can rewrite the $\boldsymbol{y} - \boldsymbol{X}\beta$ by adding and subtracting $\boldsymbol{X}\beta^*$. This leads to an expression, $(\boldsymbol{y} - \boldsymbol{X}\beta^*) + (\boldsymbol{X}\beta^* - \boldsymbol{X}\beta)$. The first term does not depend on β (the object over which we want to minimize), which we called $\boldsymbol{\varepsilon}$. We can further group the matrix in the second term. Then we can rewrite the expression as $\boldsymbol{y} - \boldsymbol{X}\beta = \boldsymbol{\varepsilon} - \boldsymbol{X}(\beta^* - \beta)$. We can use this to rewrite an interpretable expression for $SRR(\beta)$.

$$
\begin{aligned}
\text{SSR}(\beta) &= ||\boldsymbol{\varepsilon} + \boldsymbol{X}(\beta^* - \beta)||^2 && \text{Plug in } \boldsymbol{y} - \boldsymbol{X}\beta \\
&= (\boldsymbol{\varepsilon} + \boldsymbol{X}(\beta^* - \beta))^t(\varepsilon + \boldsymbol{X}(\beta^* - \beta)) && \text{Rewriting using transpose} \\
&= \boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon} + \boldsymbol{\varepsilon}^t\boldsymbol{X}(\beta^* - \beta) + (\beta^* - \beta)^t\boldsymbol{X}^t\boldsymbol{\varepsilon} + (\beta^* - \beta)^t\boldsymbol{X}^t\boldsymbol{X}(\beta^* - \beta) && \text{Expanding} \\
&= \boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon} - 2\boldsymbol{\varepsilon}^t\boldsymbol{X}(\beta^* - \beta) + (\beta^* - \beta)^t\boldsymbol{X}^t\boldsymbol{X}(\beta^* - \beta) && \text{Grouping terms} \\
&= \boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon} + (\beta^* - \beta)^t\boldsymbol{X}^t\boldsymbol{X}(\beta^* - \beta) && \text{Applying Lemma 3.1.1}
\end{aligned}
$$

Notice that $\boldsymbol{\varepsilon}^t\boldsymbol{X}(\beta^* - \beta)$ is a scalar. Therefore it is equal to its transpose, using Lemma 1.7.1. Therefore, from lines 3 to 4 we use the fact that $\boldsymbol{\varepsilon}^t\boldsymbol{X}(\beta^* - \beta) = (\beta^* - \beta)^t\boldsymbol{X}^t\boldsymbol{\varepsilon}$. From lines 4 to 5 we apply the result in Lemma 3.1.1 stating that $\varepsilon^t\boldsymbol{X} = \boldsymbol{0}_{1\times n}$.

The term $\varepsilon^t\varepsilon$ is an error component that does not depend on the choice of β. It captures the lack of fit of the model overall, even if we choose an optimal approximate solution. The only terms that matters for the optimization is the second term, which can be rewritten as:

$$
\begin{aligned}
(\beta^* - \beta)^t\boldsymbol{X}^T\boldsymbol{X}(\beta^* - \beta) &= (\boldsymbol{X}(\beta^* - \beta))^t(\boldsymbol{X}(\beta^* - \beta)) \\
&= ||\boldsymbol{X}(\beta^* - \beta)||^2
\end{aligned}
$$

The norm $||\boldsymbol{X}(\beta^* - \beta)||^2$ is equal to zero if and only if $\boldsymbol{X}(\beta^* - \beta) = \boldsymbol{0}_{m\times 1}$. Since $\boldsymbol{X}$ is full rank, there is a unique solution $\beta = \beta^*$. Therefore, β^* is the unique minimize of $SSR(\beta)$.

□

3.2 Orthogonal Spaces

In linear algebra it is convenient to think of sets that have a particular property. Previously we analyzed two important sets, the image and the kernel. Now we will analyze a third important object called the **orthogonal** set.

$$Orthog(A) = Im(A)^{\perp} := \{\boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{y}^t\boldsymbol{z} = 0,\ \forall \boldsymbol{z} \in Im(A)\}$$

Lemma 3.2.1. *Let A be an $m \times n$ matrix, then $Im(A) \cap Orthog(A) = \{\mathbf{0}_{m\times 1}\}$.*

Proof. Suppose that $\boldsymbol{x} \in \mathbb{R}^m$ is non-zero. Then $\boldsymbol{x}^t\boldsymbol{x} > 0$. If $\boldsymbol{x} \in Orthog(A)$ then $\boldsymbol{x}^t\boldsymbol{a} = 0$ for all $\boldsymbol{a} \in Im(A)$. Furthermore, if $\boldsymbol{x} \in Im(A)$ then $\boldsymbol{x}^t\boldsymbol{x} = 0$, which is a contradiction.

To complete the proof we need to verify that $\mathbf{0}_{m\times 1}$ is indeed part of the intersection. The vector $\mathbf{0} \in Im(A)$ because $A\mathbf{0}_{n\times 1} = \mathbf{0}_{m\times 1}$. It is also part of $Orthog(A)$ because $\mathbf{0}^t_{m\times 1}\boldsymbol{y} = 0$ for any $\boldsymbol{y} \in \mathbb{R}^m$ including those contained in $Im(A)$.

□

3.3 Projection Matrices

Projection matrices arise often in econometrics and in other linear systems with approximate solutions. They allow us to decompose a vector in the Euclidean space into a component that is projected onto the image of a matrix and a component that belongs to its orthogonal complement. Lemma 3.3.1 proves a characterization of projection matrices that is easier to verify in practice, which we will use in subsequent proofs.

Definition 3.3.1. Let A be an $m \times n$ matrix. Then the matrix P is a projection matrix onto $Im(A)$ if for all $\boldsymbol{z} \in \mathbb{R}^m$,

$$P\boldsymbol{z} \in Im(A) \subseteq \mathbb{R}^m, \quad (I - P)\boldsymbol{z} \in Orthog(A) \subseteq \mathbb{R}^m$$

Definition 3.3.2. Let A be an $m \times m$ matrix.

(a) The matrix A is idempotent if $AA = A$.

(b) The matrix A is symmetric if $A^t = A$.

Lemma 3.3.1. *Let A be an $m \times n$ matrix and let P be an $m \times m$ matrix. If $Im(P) = Im(A)$ then P is a projection matrix onto A if and only if P is idempotent and symmetric.*

Proof. ($\Longrightarrow$) Suppose that P is a projection matrix. Then for all $\boldsymbol{z} \in \mathbb{R}^m$, $P\boldsymbol{z} \in Im(A)$ and $(I - P)\boldsymbol{z} \in Orthog(A)$. By definition of the orthogonal set we have $\boldsymbol{z}^t(I - P)^t P\boldsymbol{z}$ regardless of the choice of input vectors. That means that $(I-P)^t P = \mathbf{0}_{m\times m}$. Rearranging the equation we get that $P = P^t P$. The matrix is symmetric because $P^t = (P^t P)^t = P^t P = P$. Using the fact that it is symmetric, we can show that it is also idempotent because $P = P^t P = PP$.

($\Longleftarrow$) Now suppose that $Im(P) = Im(A)$, together with the condition that P is idempotent and symmetric. Since $Im(P) = Im(A)$, then for all $\boldsymbol{z} \in \mathbb{R}^m$, $P\boldsymbol{z} \in Im(A)$. Furthermore, for every $\boldsymbol{a} \in Im(A)$ there exists a $\boldsymbol{x} \in \mathbb{R}^m$ such that $P\boldsymbol{x} = \boldsymbol{a}$. Let $\boldsymbol{z} \in \mathbb{R}^m$, then $\boldsymbol{a}^t(I - P)\boldsymbol{z} = \boldsymbol{x}^t P^t(I - P)\boldsymbol{z} = 0$ since $P^t(1 - P) = \mathbf{0}_{m\times m}$. This follows directly from idempotency and symmetry: $P^t - P^t P = P - PP = P - P = \mathbf{0}_{m\times m}$. That means that $(I - P)\boldsymbol{z}$ is orthogonal to every element $\boldsymbol{a} \in Im(A)$, and therefore $(I - P)\boldsymbol{z} \in Orthog(A)$. □

3.3.1 Computation of Full Rank Matrices

Projection matrices can be computed directly for full rank matrices.

Lemma 3.3.2. *If A is an $m \times n$ full rank matrix then $P = A(A^tA)^{-1}A^t$ is a projection matrix onto $Im(A)$.*

Proof. We will verify that P satisfies the conditions of 3.3.1.

First we show that the matrix is symmetric and idempotent. Using Lemma 3.5.1, $((A^tA)^{-1})^t = ((A^tA)^t)^{-1}$. We also use Lemma 1.7.1 to show that $(A^tA)^t = A^t(A^t)^t = A^tA$. Therefore $P^t = A(A^tA)^{-1}A^t = P$, and therefore our candidate matrix is symmetric. It is also idempotent because $PP = A(A^tA)^{-1}A^tA(A^tA)^{-1}A^t$ which is equal to $A(A^tA)^{-1}A^t = P$ by canceling some of the terms.

Second, we show that $Im(P) = Im(A)$.

(i) $Im(P) \subseteq Im(A)$: Let $\boldsymbol{x} \in \mathbb{R}^m$. Therefore $P\boldsymbol{x} = A\boldsymbol{z}$, where $\boldsymbol{z} = (A^tA)^{-1}A^t\boldsymbol{x} \in \mathbb{R}^n$, is contained in $Im(A)$.

(ii) $Im(A) \subseteq Im(P)$: Suppose that $\boldsymbol{z} \in Im(A) \subseteq \mathbb{R}^m$, then there exists a $\boldsymbol{x} \in \mathbb{R}^n$ such that $A\boldsymbol{x} = \boldsymbol{z}$. Then

$$\begin{aligned} P\boldsymbol{z} &= A(A^tA)^{-1}A^t\boldsymbol{z} && \text{(Substituting definition of } P\text{)} \\ &= A(A^tA)^{-1}A^tA\boldsymbol{x} && \text{(Since } \boldsymbol{z} \in Im(A)\text{)} \\ &= A\boldsymbol{x} && \text{(Cancelling out terms)} \\ &= \boldsymbol{z} && \text{(Plugging-in definition of } \boldsymbol{z}\text{)} \end{aligned}$$

That means that $\boldsymbol{z} \in Im(P)$. Therefore, $Im(A) \subseteq Im(P)$.

To conclude the proof we apply Lemma 3.3.1 to show that our candidate matrix P is a projection matrix onto A.

□

Proof. Here is a more basic proof using the definition for a projection matrix. We will verify two things: $\forall \boldsymbol{z} \in \mathbb{R}^m$, (i) $P\boldsymbol{z} \in Im(A)$, (ii) $(I - P)\boldsymbol{z} \in Orthog(A)$.

(i) Plugging in the expression, we have

$$P\boldsymbol{z} = A(A^tA)^{-1}A^t\boldsymbol{z} = A\boldsymbol{w},$$

where $\boldsymbol{w} := (A^tA)^{-1}A^t\boldsymbol{z}$, and hence $P\boldsymbol{z} \in Im(A)$.

(ii) Now we want to show $(I - P)\boldsymbol{z} \in Orthog(A)$. Consider an arbitrary element in $Im(A)$, $A\boldsymbol{x}$. We have

$$((I - P)\boldsymbol{z})^t(A\boldsymbol{x}) = \boldsymbol{z}^t(I - P)A\boldsymbol{x} = \boldsymbol{z}^t(A - A(A^tA)^{-1}A^tA)\boldsymbol{x} = 0.$$

Combining (i) and (ii), we have shown that $P = A(A^tA)^{-1}A^t$ is indeed a projection matrix onto $Im(A)$. □

3.3.2 Computation of Rank Deficient Matrices

If the matrix A is not full rank we cannot use the formula in Lemma 3.3.2 directly. Fortunately, we can formulate a more general theorem with minor modifications.

Theorem 3.3.1. *Let A be an $m \times n$ matrix.*

(a) If A is the zero matrix, $P = \mathbf{0}_{m\times m}$ is a projection matrix onto $Im(A)$.

(b) If A is a non-zero matrix, then there exists an $m \times k$ full rank matrix B such that $Im(B) = Im(A)$. Furthermore, for any B with this property, $P = B(B^tB)^{-1}B^t$ is a projection matrix onto $Im(A)$.

The second part of Theorem 3.3.1 is particularly interesting because it says that we can construct projection matrices in a simple way from rank deficient matrices. It suffices to construct a full rank matrix that spans the same space. One simple alternative is to drop certain columns that are linear combinations of the others.

Proof. (a) If A is the zero matrix, then $Im(A) = \{\mathbf{0}_{m\times 1}\}$. If $P = \mathbf{0}_{m\times m}$ then it is (i) Symmetric, $P^t = \mathbf{0}_{m\times m} = P$, (ii) Idempotent, $PP = \mathbf{0} = P$, and (iii) $Im(P) = \{\mathbf{0}\} = Im(A)$. Therefore, using Lemma 3.3.1 we show that P is a projection matrix unto $Im(A)$.

(b) If A is a non-zero matrix, then by Lemma 2.4.2 there exists a full rank matrix B such that $Im(B) = Im(A)$. The intuition is that we can always drop certain columns to make a matrix full rank and still span the same space.

Now choose an arbitrary B that satisfies this property. Using Lemma 3.3.2, P is a projection matrix onto $Im(B)$ with the property that $P\boldsymbol{z} \in Im(B)$ and $(I - P)\boldsymbol{z} \in Orthog(B)$. We know that $Im(B) = Im(A)$. To complete the proof we just need to show that $Orthog(B) = Orthog(A)$. If $\boldsymbol{w} \in Orthog(B) \subseteq \mathbb{R}^m$ then for any $\boldsymbol{b} \in Im(B)$ we have $\boldsymbol{w}^t\boldsymbol{b} = 0$. Since $Im(B) = Im(A)$ then $\boldsymbol{w} \in Orthog(A)$. We can use a similar argument to show that if $\boldsymbol{w} \in Orthog(A)$ then it also belongs in $Orthog(B)$. Therefore, P is a projection matrix onto $Im(A)$.

□

3.3.3 Uniqueness of Projection Matrices

Lemma 3.3.3. *Let A be an $m \times n$ matrix.*

(a) For each $\boldsymbol{x} \in \mathbb{R}^m$ there exist unique vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^m$ such that (i) $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b}$. (ii) $\boldsymbol{a} \in Im(A)$ and $\boldsymbol{b} \in Orthog(A)$.

(b) If P is a projection matrix onto $Im(A)$ then it is the unique.

Proof. (a) First we prove that such vectors exist. Define P as in Theorem 3.3.1. Then for $\boldsymbol{x} \in \mathbb{R}^m$ define $\boldsymbol{a} = P\boldsymbol{x} \in Im(A)$ and $\boldsymbol{b} = (I - P)\boldsymbol{x} \in Orthog(A)$. We can verify that $\boldsymbol{a} + \boldsymbol{b} = P\boldsymbol{x} + (I - P)\boldsymbol{x} = \boldsymbol{x}$. This shows that such vectors exits.

To prove uniqueness, assume that there exist alternative vectors $\boldsymbol{a}', \boldsymbol{b}'$ with the same properties. Then $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b} = \boldsymbol{a}' + \boldsymbol{b}'$. We can rewrite this as $\boldsymbol{a} - \boldsymbol{a}' = \boldsymbol{b}' - \boldsymbol{b}$. Since $\boldsymbol{a}, \boldsymbol{a}' \in Im(A)$ the left-hand side belongs in $Im(A)$. Since, $\boldsymbol{b}, \boldsymbol{b}' \in Orthog(A)$ the right-hand side belongs in $Orthog(A)$. However, by Lemma 3.2.1, $Im(A) \cap Orthog(A) = \{\boldsymbol{0}\}$. That means that $\boldsymbol{a} = \boldsymbol{a}'$ and $\boldsymbol{b} = \boldsymbol{b}'$.

(b) Suppose that there exist two matrices P, P' such that for all $\boldsymbol{x} \in \mathbb{R}^m$, then $P\boldsymbol{x}, P'\boldsymbol{x} \in Im(A)$ and $(I - P)\boldsymbol{x}, (I - P')\boldsymbol{x} \in Orthog(A)$. Then by the first part of the lemma $P\boldsymbol{x} = P'\boldsymbol{x}$ and $(I - P)\boldsymbol{x} = (I - P')\boldsymbol{x}$. Since $\boldsymbol{x}$ is arbitrary, then $P = P'$. To prove this set $\boldsymbol{x} = \boldsymbol{e}_j$ (an elementary basis vector) and use the fact that $P\boldsymbol{e}_j = p_j = p'_j = P'\boldsymbol{e}_j$ for $j \in \{1, \ldots, m\}$, where p_j, p'_j are the j^{th} columns of P, P', respectively.

□

3.3.4 Optimality and Non-Unique Approximate Solutions

Corollary 3.3.1. *Let $\boldsymbol{X}$ be an $m \times n$ matrix. The residual sum of squares $SSR(\beta)$ always has a minimizer. It is unique if and only if $\boldsymbol{X}$ is full rank.*

Proof. We can rewrite the problem as:

$$\min_{\beta} ||\boldsymbol{y} - \boldsymbol{X}\beta||^2 = \min_{z \in Im(\boldsymbol{X})} ||\boldsymbol{y} - \boldsymbol{z}||^2$$

Then we can use Lemma 3.3.3 to rewrite decompose $\boldsymbol{y}$ into its projections, $\boldsymbol{a} \in Im(\boldsymbol{X})$ and $\boldsymbol{b} \in Orthog(\boldsymbol{X})$ such that $\boldsymbol{y} = \boldsymbol{a} + \boldsymbol{b}$. We can expand the equation as:

$$\begin{aligned} ||\boldsymbol{y} - \boldsymbol{z}||^2 &= ||\boldsymbol{b} + (\boldsymbol{a} - \boldsymbol{z})||^2 \\ &= \boldsymbol{b}^t\boldsymbol{b} + 2\boldsymbol{b}^t(\boldsymbol{a} - \boldsymbol{z}) + (\boldsymbol{a} - \boldsymbol{z})^t(\boldsymbol{a} - \boldsymbol{z}) \\ &= \boldsymbol{b}^t\boldsymbol{b} + (\boldsymbol{a} - \boldsymbol{z})^t(\boldsymbol{a} - \boldsymbol{z}) \\ &= \boldsymbol{b}^t\boldsymbol{b} + ||\boldsymbol{a} - \boldsymbol{z}||^2 \end{aligned}$$

Therefore the minimizer is $\boldsymbol{z} = \boldsymbol{a}$. The system $\boldsymbol{a} = \boldsymbol{X}\beta$ has a unique solution if and only if $\boldsymbol{X}$ is full rank.

□

3.4 Application: Detrending Data

A researcher has access to a database with information about a series of an economic variable over T time periods. There is a vector of outcomes $\boldsymbol{Y} \in \mathbb{R}^T$, with outcomes for different time periods. There are two sets of regressors: (i) A $T \times k_1$ matrix X_1 which contain the main variables an interest and (ii) a $T \times k_2$ of control variables. For example, the control matrix could include a single variable with a trend $X_{2t} = t$, where $t \in \{1, \ldots, t\}$ where t is the time period of each observation.[1] To ensure that our results have unique solutions, we assume that the joint design matrix $X = [X_1, X_2]$ is full rank.

The researcher is debating between different regression specifications, which both use the least-squares optimality criterion analyzed in Lemma 3.1.2. We use $\hat{\beta}, \hat{\psi}_1$ and $\hat{\psi}_Y$ to denote the solution to the optimality criterion and non-hat symbols to denote the "generating model". We are only interested in these solutions, and just provide the generating model for context.

Example 1 (Additional Trend Regressor)**.** *A regression with both the main variables and the controls, with associated parameters* $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$ *where* $\beta_1 \in \mathbb{R}^{k_1}$ *and* $\beta_2 \in \mathbb{R}^{k_2}$*,*

$$\boldsymbol{Y} = \boldsymbol{X}_1\beta_1 + \boldsymbol{X}_2\beta_2 + \boldsymbol{e} \quad \Longrightarrow \quad (\boldsymbol{X}^t\boldsymbol{X})\hat{\beta} = \boldsymbol{X}^t\boldsymbol{Y} \tag{3.1}$$

Example 2 (Detrended Regression)**.** *Two-step procedure:*

1. *First detrend the regressors and the outcome by running two* ***auxiliary*** *regressions (i)* $\boldsymbol{X}_1$ *on* $\boldsymbol{X}_2$ *and (ii)* $\boldsymbol{Y}$ *on* $\boldsymbol{X}_2$*. Let* $\psi_1, \psi_Y \in \mathbb{R}^{k_2}$ *be vectors that solve,*

$$\boldsymbol{X}_1 = \boldsymbol{X}_2\psi_1 + \boldsymbol{u}_1 \quad \Longrightarrow \quad (\boldsymbol{X}_2^t\boldsymbol{X}_2)\hat{\psi}_1 = \boldsymbol{X}_2^t\boldsymbol{X}_1$$

$$\boldsymbol{Y} = \boldsymbol{X}_2\psi_Y + \boldsymbol{u}_Y \quad \Longrightarrow \quad (\boldsymbol{X}_2^t\boldsymbol{X}_2)\hat{\psi}_Y = \boldsymbol{X}_2^t\boldsymbol{Y}$$

   Compute the detrended variables (residuals) (i) Main Regressors: $\hat{\boldsymbol{U}}_1 = \boldsymbol{X}_1 - \boldsymbol{X}_2\hat{\psi}_1$ *and (ii) Outcome:* $\hat{\boldsymbol{U}}_Y = \boldsymbol{Y} - \boldsymbol{X}_2\hat{\psi}_Y$*.*

2. *Run a second regression using the detrended variables:*

$$\hat{\boldsymbol{U}}_Y = \hat{\boldsymbol{U}}_1\beta_1 + \tilde{\boldsymbol{e}} \quad \Longrightarrow (\hat{\boldsymbol{U}}_1^t\hat{\boldsymbol{U}}_1)\tilde{\beta}_1 = \hat{\boldsymbol{U}}_1^t\hat{\boldsymbol{U}}_Y \tag{3.2}$$

[1] The trend is just an example to give economic context, for our algebra results the matrix X_2 is unrestricted. Another meaningful example is the "fixed effects model" in panel data. In that case the researcher has access to information from multiple time periods and individuals. She includes a dummy variable for each individual (capturing an individual-specific effect across time periods);

3.4.1 Projections of Block-Partitioned Matrices

Define the projection matrices $P := \boldsymbol{X}(\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}$ and $P_2 := \boldsymbol{X}_2(\boldsymbol{X}_2^t\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2$. Define the **residual-making matrices** $M := I - P$ and $M_2 := I - P_2$. We can rewrite the residuals of the regression specifications more succinctly in terms of the projection matrices. Suppose that we substitute in the definition of $\hat{\psi}_1$ and $\hat{\psi}_Y$.

$$\begin{aligned} \hat{\boldsymbol{U}}_1 = \boldsymbol{X}_1 - \boldsymbol{X}_2(\boldsymbol{X}_2^t\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2\boldsymbol{X}_1 \quad &\Longrightarrow \quad \hat{\boldsymbol{U}}_1 = M_2\boldsymbol{X}_1 \\ \hat{\boldsymbol{U}}_Y = \boldsymbol{Y} - \boldsymbol{X}_2(\boldsymbol{X}_2^t\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2\boldsymbol{Y} \quad &\Longrightarrow \quad \hat{\boldsymbol{U}}_Y = M_2\boldsymbol{Y} \end{aligned}$$

This means that $\tilde{\beta}$ can be written as $\tilde{\beta} = (\boldsymbol{X}_1^t M_2^t M_2 \boldsymbol{X}_1)^{-1}(\boldsymbol{X}_1^t M_2^t M_2 \boldsymbol{Y})$. In the exercises you will prove that $\tilde{\beta}_1 = \hat{\beta}_1$. The conclusion is that both regression specifications (adding trend as regressor or detrending) yield numerically the same estimator of the main effects. As an input you will need the following lemma.

Lemma 3.4.1. *The matrices P, P_2, M, M_2 satisfy the following properties.*

1. *P, P_2, M, M_2 are idempotent and symmetric.*
2. *$PP_2 = P_2$ and $MP_2 = \mathbf{0}_{T\times T}$.*

Proof. The result has two parts:

1. By Lemma 3.3.2 we know that P, P_1 are idempotent and symmetric. We show that M is also idempotent and symmetric. The proof is analogous for M_2.
   (a) (Idempotency) By definition $MM = (I - P)(I - P)$. We can construct expand out the sum as $I - 2P + PP$. Since P is idempotent, $PP = P$ and the expression simplifies to $MM = I - P = M$.
   (b) (Symmetry) By definition $M^t = (I - P)^t$, which we can expand as $(I - P^t)$. Since P is symmetric, $P^t = P$ and $M^t = I - P = M$.
2. By construction $\boldsymbol{X}_2 \in Im(\boldsymbol{X})$ (each column vector is in the image). Since P is a projection matrix onto $Im(\boldsymbol{X})$, then $P\boldsymbol{X}_2 = \boldsymbol{X}_2$. We can plug-in the definition of P_2 so that $PP_2 = P\boldsymbol{X}_2(\boldsymbol{X}_2^t\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2$. Substituting $P\boldsymbol{X}_2 = \boldsymbol{X}_2$ then $PP_2 = \boldsymbol{X}_2(\boldsymbol{X}_2^t\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2 = P_2$. The second result follows by plugging-in the definition: $MP_2 = (I - P)P_2 = P_2 - PP_2$ which is equal to $MP_2 = P_2 - P_2 = \mathbf{0}_{T\times T}$.

□

3.5 Properties Appendix: Inverse of the Transpose

Lemma 3.5.1. *Let A be a full rank $m \times m$ matrix. Then A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$.*

Proof. If A is full rank then there exists a matrix A^{-1} such that $A^{-1}A = I$. Transpose on both sides, $A^t(A^{-1})^t = I^t = I$. Then $(A^{-1})^t$ is the inverse of A^t. □

3.6 Exercises

1. In this exercise you will prove a version of the Frisch-Waugh-Lovell Theorem (Greene, 2012) in the detrending example.

   (a) Prove that $\tilde{\beta}_1 = (\boldsymbol{X}_1^t M_2 \boldsymbol{X}_1)^{-1}(\boldsymbol{X}_1^t M_2 \boldsymbol{Y})$.

   (b) Show that the system in Equation 3.1 can be written in block-partition form as:

   $$\begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   (c) Show that second row can be rewritten as $\hat{\beta}_2 = (\boldsymbol{X}_2^t \boldsymbol{X}_2)^{-1}(\boldsymbol{X}_2^t \boldsymbol{Y} - \boldsymbol{X}_2^t \boldsymbol{X}_1 \hat{\beta}_1)$.

   (d) Plug the above result into the first row of equations and show that $(\boldsymbol{X}_1^t M_2 \boldsymbol{X}_1)\hat{\beta}_1 = (\boldsymbol{X}_1^t M_2 \boldsymbol{Y})$. Conclude that $\hat{\beta}_1 = \tilde{\beta}_1$.

2. In the detrending example:

   (a) Show that $\boldsymbol{X}$ full rank implies that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are full rank.

   (Hint: Prove by contradiction)

   (b) Define $B = M_2 \boldsymbol{X}_1$. Show that replacing $\boldsymbol{X}_1$ with the matrix B does not change the image, i.e. $Im(\boldsymbol{X}_1, \boldsymbol{X}_2) = Im(B, \boldsymbol{X}_2)$.

   (Hint: Modify Lemma 3.1.1)

   (c) Show that if $\boldsymbol{X}$ is full rank then $(\boldsymbol{X}_1^t M_2 \boldsymbol{X}_1)$ is full rank. (Hint: Review Linear Regression Section)

Chapter 4

Convex Sets (I): Hyperplanes

4.1 Convex Sets

Convex sets feature prominently in microeconomic theory. For example they are used to represent consumers' preference for **diverse** bundles of goods. They capture the idea that consumers prefer to consume a balanced amount (convex combination) of two goods rather than have too much of a single one. Similarly, budget sets can be expressed as a particular type of convex sets: a hyperplane. If two allocations are within a person's budget then a combination of them (rearranging the proportions) is also in her budget. In this case convex sets capture the **feasibility** of an allocation.

Therefore it is of central importance to microeconomic theorists to understand how goods are allocated given a budget set (hyperplane) and a set of preferences (a convex set). The main theorems that we develop in this chapter are about the **existence** of hyperplanes. In economics the hyperplane theorems have wide applicability in proving the **existence** of equilibria and optimal solutions.

In this chapter we focus on proving a set of basic properties of convex sets. We revisit three types of sets from real analysis (the interior, the boundary and the complement of the closure). We will prove a hyperplane theorem for each case. The statement is relatively similar (with minor differences in the assumptions and the results), so it useful to know how and why each assumption is used.

Definition 4.1.1. A set $\mathcal{X} \subseteq \mathbb{R}^n$ is **convex** if

$$\lambda x + (1-\lambda)x' \in \mathcal{X} \quad \forall x, x' \in \mathcal{X}, \quad \forall \lambda \in [0,1]$$

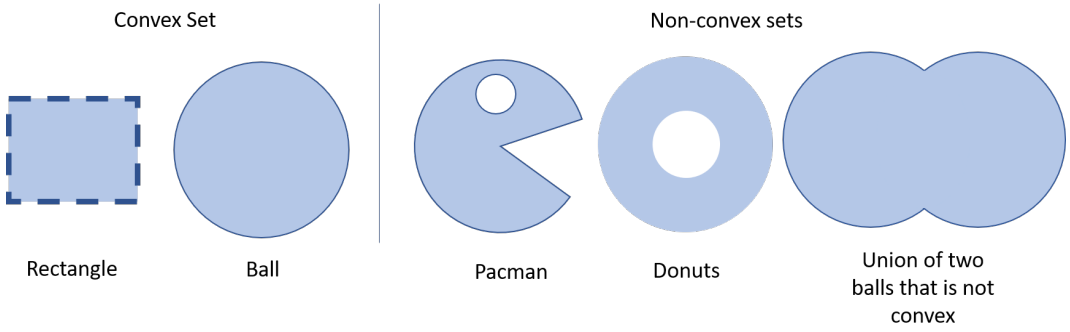

Figure 4.1: Examples

Definition 4.1.2. The vector $x_\lambda \in \mathbb{R}^n$ is called a convex combination of vectors $x_1, \ldots, x_K$, if $x_\lambda = \sum_{k=1}^K \lambda_k x_k$ and $\lambda_k \in [0,1], \quad \sum_k \lambda_k = 1$.

Lemma 4.1.1. *A set $\mathcal{X} \subseteq \mathbb{R}^n$ is convex if and only if every convex combination x_λ of K vectors $x_1, \ldots, x_K \in \mathcal{X}$ is contained in the set, i.e. $x_\lambda \in \mathcal{X}$, for any positive integer K.*

Proof. ($\Longleftarrow$) Set $K = 2$ and it follows by definition.

($\Longrightarrow$) We will prove this direction by induction. If $K = 1$ then $x_\lambda = x$ which belongs to $\mathcal{X}$ by definition. Now, suppose that it holds for some finite K. Our objective is to show it holds for $K + 1$. Let $\lambda_1, \ldots, \lambda_{K+1}$ be scalar in the unit interval that add up to one. If $\lambda_{K+1} = 1$ then we are done because x_λ^{K+1} is a convex combination of a single vector. Otherwise, assume that $0 \leq \lambda_{K+1} < 1$. We can rewrite the convex combination as:

$$x_\lambda^{K+1} = (1 - \lambda_{K+1}) \left(\frac{1}{1 - \lambda_{K+1}} \sum_{k=1}^{K} \lambda_k x_k \right) + \lambda_{K+1} x_{K+1}$$

Define $x_\lambda^K := \frac{1}{1-\lambda_{k+1}} \sum_{k=1}^K \lambda_k x_k$ belongs to $\mathcal{X}$. We can show that $\frac{\lambda_k}{1-\lambda_{K+1}} \geq 0$ and that $\frac{1}{1-\lambda_{K+1}} \sum_{k=1}^K \lambda_k = 1$. Therefore, it follows that x_λ^K is a convex combination of K vectors and by the induction step, it belongs to $\mathcal{X}$. Finally to complete the proof, $x_\lambda^{K+1} = (1 - \lambda_{K+1}) x_\lambda^K + \lambda_{K+1} x_{K+1}$ which is contained in $\mathcal{X}$ by the definition of a convex set.

□

4.2 Hyperplanes

Hyperplanes are useful building blocks to characterize certain sets in economic theory. For example, suppose that there are n goods with prices p_j, $1 \leq j \leq n$. The consumer decides to purchase a quantity x_j of each good. Then her total expenditure can be expressed as $p^t x = \sum_{j=1} p_j x_j$. Moreover, total expenditure needs to be less than or equal to her level of wealth. If $p^t x = w$ then she is spending all her budget, but the feasible set[1] is characterized by $p^t x \leq w$. We will prove a set of existence results for the existence of separating hyperplanes.

Definition 4.2.1. Let $p \in \mathbb{R}^n \backslash \{0_{n\times 1}\}$ and $w \in \mathbb{R}$. The set

$$H(p, w) = \{x \in \mathbb{R}^n : p^t x = w\}$$

Definition 4.2.2. Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^n$. Then

1. $\mathcal{X}$ and $\mathcal{Y}$ are weakly separated by $H(p, w)$ if

$$\begin{matrix} p^t x \geq w \\ w \geq p^t y \end{matrix}, \quad \forall x \in \mathcal{X},\ \forall y \in \mathcal{Y}$$

2. The sets are separated if one inequality is weak but the other is not.

3. The sets are strictly separated if both inequalities are strict.

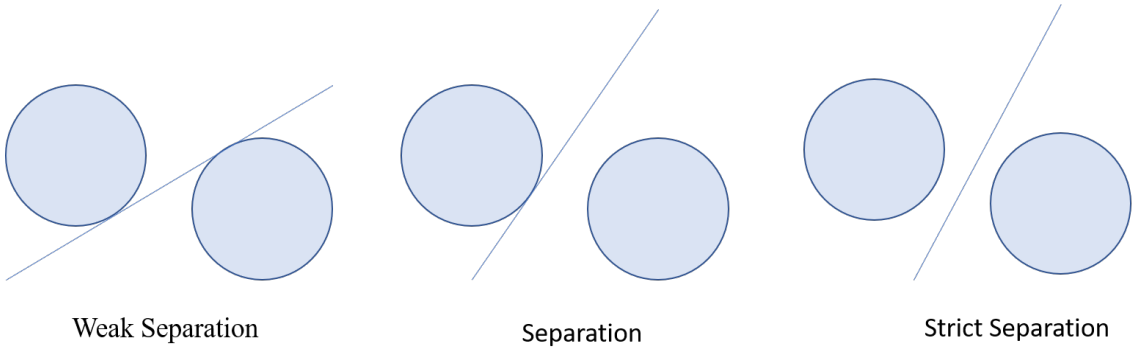

Figure 4.2: Types of Separation of Convex Sets

[1] In practical models there is also a constraint that the quantities consumed need to be non-negative.

4.3 Separating Points from Convex Sets

4.3.1 Topology of Convex Sets

We will focus our attention on the Euclidean space, which is a complete metric space. We define $B(x, \epsilon)$ as an open ball with center x and radius ϵ.

Definition 4.3.1. A point $x \in \mathcal{X} \subseteq \mathbb{R}^n$ is an **interior point** if there exists an $\epsilon > 0$ s.t. $B(x, \epsilon) \subseteq \mathcal{X}$. We define $int(\mathcal{X})$ as the set of all interior points of $\mathcal{X}$, which we call the **interior** of the set.

Definition 4.3.2. A point $x \in \mathcal{X} \subseteq \mathbb{R}^n$ is a **limit point** if there exists a sequence $x_k \in \mathcal{X}$ such that $x_k \to x$. We define $\overline{X}$ the set of all limit points of $\mathcal{X}$, which we call the **closure** of the set.

Definition 4.3.3. The **boundary** of a set $\mathcal{X} \subseteq \mathbb{R}^n$ is defined as $\partial\mathcal{X} := \overline{\mathcal{X}} \setminus int(\mathcal{X})$.

A few relations between the definitions hold for all sets. For example, $int(\mathcal{X}) \subseteq \mathcal{X}$ every interior point is contained in the set. Furthermore, $\mathcal{X} \subseteq \overline{\mathcal{X}}$ because if $x \in \mathcal{X}$ we can always define a sequence $x_k = x$. It also follows that a set is open if all its points are interior and closed if it contains all its limit points. It can be shown that if X is open then $X = int(X)$ and if X is closed then $X = \overline{X}$. See (Rudin et al., 1964) for more details.

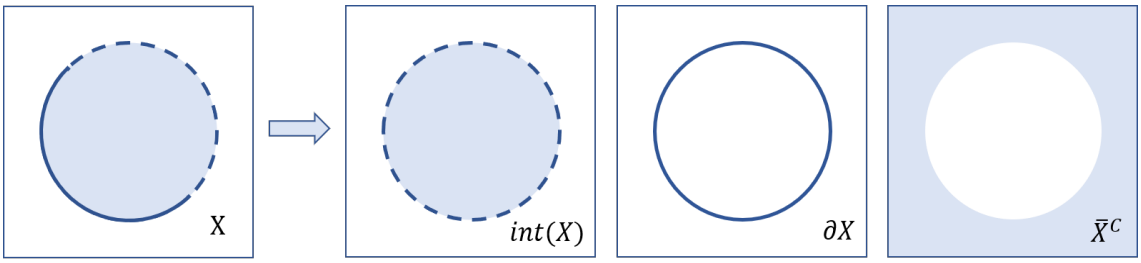

Figure 4.3: Example of the interior, boundary and closure.

By construction the sets $int(\mathcal{X})$, $\partial\mathcal{X}$ and $\overline{\mathcal{X}}^C$ (the complement of the closure) are mutually disjoint and their union is equal to the entire Euclidean space, $\mathbb{R}^n = int(\mathcal{X}) \cup \partial\mathcal{X} \cup \overline{\mathcal{X}}^C$. That means that the three sets comprise an exhaustive list of cases that we will explore in our hyperplane theorems.

Lemma 4.3.1. *The interior $int(\mathcal{X})$ and closure $\overline{\mathcal{X}}$ of a convex set $\mathcal{X}$ are also convex.*

Lemma 4.3.2. *If $\mathcal{X}$ is a convex set, pick $x \in \overline{\mathcal{X}}$ and $y \in int(\mathcal{X})$, then $\lambda x + (1-\lambda)y \in int(x), \forall\lambda \in (0, 1)$.*

Lemma 4.3.3 (Topological Equivalences on Convex Sets). *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a convex set, then (i) $int(\mathcal{X}) = int(\overline{\mathcal{X}})$ and (ii) $\partial\mathcal{X} = \partial\overline{\mathcal{X}}$.*

Proof. (i) Since $\mathcal{X} \subseteq \overline{\mathcal{X}}$, the direction $int(\mathcal{X}) \subseteq int(\overline{\mathcal{X}})$ is straightforward. Let us prove the other direction $int(\overline{\mathcal{X}}) \subseteq int(\mathcal{X})$. Pick $z \in int(\overline{\mathcal{X}})$. By definition, there exists ϵ such that $B(z, \epsilon) \subset \overline{\mathcal{X}}$. We want to show that $z \in int(\mathcal{X})$. From the previous lemma we know that, if we could write $z = \lambda x + (1-\lambda)y$, for some $x \in \overline{\mathcal{X}}, y \in int(\mathcal{X}), \lambda \in (0,1)$, then we are done. So we are going to construct such a convex combination representation for z. Equivalently, we are looking for $x = \frac{1}{\lambda}z - \frac{1-\lambda}{\lambda}y$ with the above restrictions.

Pick $y \in int(\mathcal{X})$. We want to pick λ to guarantee that $x \in \overline{\mathcal{X}}$. Notice that $||x - z|| = \frac{1-\lambda}{\lambda}||z - y||$. Let us pick $\lambda = \frac{1}{1+\epsilon/(2||z-y||)}$, then $||x - z|| = \frac{\epsilon}{2} < \epsilon$, and hence $x \in B(z, \epsilon)$ and therefore $x \in \overline{\mathcal{X}}$. Note that now we have constructed a $\lambda \in (0,1)$ together with a point $y \in int(\mathcal{X})$ and a point $x \in \overline{\mathcal{X}}$, such that $z = \lambda x + (1-\lambda)y$. By the previous lemma, we have $z \in int(\mathcal{X})$.

(ii) By definition we have $\partial\mathcal{X} = \overline{\mathcal{X}} \backslash int(\mathcal{X})$. Using result in (i), we know it is also equal to $\overline{\mathcal{X}} \backslash int(\overline{\mathcal{X}})$, which by definition is $\partial\overline{\mathcal{X}}$. □

The interior, closure and boundary are particularly useful in convex analysis because they are easier to analyze than other types of sets. Intuitively, Lemma 4.3.3 formalizes the idea that a convex set does not have any holes "inside" the set.

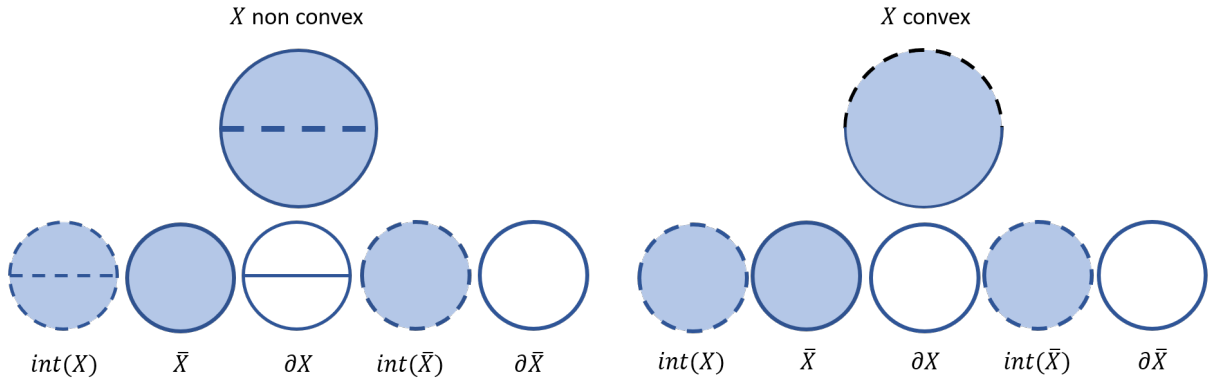

Figure 4.4: Example with convex and non-convex sets.

Here is a numerical counterexample that shows convexity is important in the result. Consider

$$\mathcal{X} = (-1, 0) \cup (0, 1),$$

which is not a convex set. The closure of $\mathcal{X}$ is $\overline{\mathcal{X}} = [-1, 1]$. But $int(\mathcal{X}) = (-1, 0) \cup (0, 1)$ while $int(\overline{\mathcal{X}}) = (-1, 1)$.

4.3.2 Non-Existence

Lemma 4.3.4. *Suppose that $\mathcal{X} \subseteq \mathbb{R}^n$ and that $d \in int(\mathcal{X})$. Then there does not exist a non-zero vector $p \in \mathbb{R}^n$ such that $p^t x \geq p^t d, \forall x \in \mathcal{X}$.*

Proof. Proof by contradiction. Suppose that exists a non-zero vector that separates $\mathcal{X}$ and d. Construct a new vector $x^* = d - \lambda p$. Then $||x^* - d|| = ||\lambda p|| = |\lambda|||p||$. On the other hand, since the point d is interior there exists an open ball of radius ϵ and center d such that $B_{d,\epsilon} \subseteq \mathcal{X}$. Set $\lambda < \frac{\epsilon}{||p||}$ then $x^* \in B(d, \epsilon) \subseteq \mathcal{X}$.

It follows that $p^t x^* = p^t(d - \lambda p) = p^t d - \lambda p^t p$. The above quantity is strictly less than $p^t d$ because $p^t p = ||p||^2 > 0$ by assumption. Therefore there does not exist a separating hyperplane.

□

We can actually strengthen this result to extend to sets with minor changes.

Corollary 4.3.1. *Suppose that $\mathcal{X}, \mathcal{Z} \subseteq \mathbb{R}^n$ and there is a vector $d \in \mathcal{X} \cap \mathcal{Z}$ such that $d \in int(\mathcal{X})$. Then there does not exist a non-zero vector $p \in \mathbb{R}^m$ such that $p^t x \geq p^t z$ for all $x \in \mathcal{X},\ z \in \mathcal{Z}$.*

Proof. By Lemma 4.3.4 there does not exist a non-zero vector p such that $p^t x \geq p^t d$. Since $d \in \mathcal{Z}$ there does not exists a hyperplane that separates the sets.

□

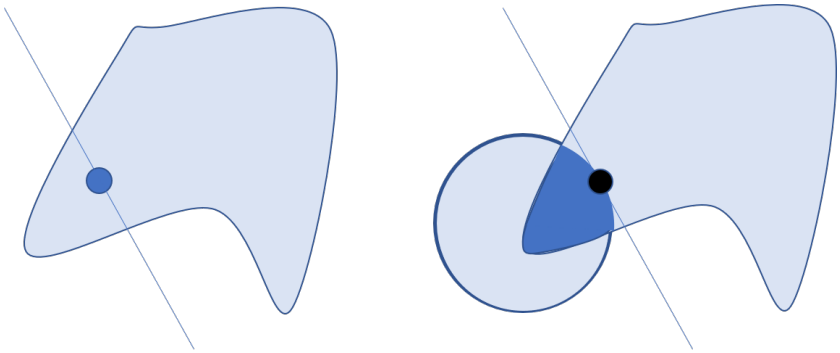

Figure 4.5: Separating hyperplanes do not exist between a point in the interior of the set and the rest of the set. This also holds for two sets, if they intersect at an interior point of one of the sets. This applies in general, not only to convex sets.

4.3.3 Strict Separation

Lemma 4.3.5. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a non-empty and closed and $d \notin \mathcal{X}$. Then there exists a minimizer $x^* = \arg\min_{x\in\mathcal{X}} ||x-d||^2$. Furthermore, the vector $p = x^* - d$ is non-zero.*

Proof. If the set $\mathcal{X}$ is bounded, set $\mathcal{A} = \mathcal{X}$. If not, define a set $\mathcal{A}(r) := \overline{B(d,r)} \cap \mathcal{X}$. Choose $0 < r^* < \infty$ such that $A(r^*) \neq \emptyset$, and set $\mathcal{A} = \mathcal{A}(r^*)$. A finite r^* is guaranteed to exist because $\mathcal{X}$ is **nonempty**. Since $\mathcal{X}$ is **closed** and $\overline{B(d,r^*)}$ is closed and bounded in $\mathbb{R}^n$, then $\mathcal{A}$ is compact.

Define a function $f(x) := ||x-d||^2$, which measures the square distance between d and a point in the set $\mathcal{X}$. The function f is continuous. Since $\mathcal{A}$ is compact and non-empty we can use the extreme value theorem to show that the function has a unique minimizer on the set, $x^* = \arg\min_{x\in\mathcal{A}} f(x)$. This is the closest point to d on the set $\mathcal{A}$. We can also show that this is the closest point in all of $\mathcal{X}$. By definition $f(x) > r^{*2} \geq ||x^* - d||^2$ for all x in the set $\mathcal{A}^C \cap \mathcal{X}$. Therefore $||x-d||^2 \geq ||x^*-d||^2$ for all x contained in $\mathcal{X}$. Defining the set $\mathcal{A}$ is a necessary intermediate step to be able to use the extreme value theorem, which is only stated for compact sets. □

Theorem 4.3.1. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a non-empty, convex set and $d \notin \bar{\mathcal{X}}$. Then there exists a hyperplane $H(p,w)$ that strictly separated $\mathcal{X}$ and $\{d\}$. That is,*

$$p^t d < w \quad \textit{and} \quad w < p^t x, \quad \forall x \in \mathcal{X}$$

Proof. We prove our results initially for $\bar{\mathcal{X}}$, which ensures the existence of a minimum distance point $x^* \in \bar{\mathcal{X}}$ in the preceding lemma, because $\bar{\mathcal{X}}$ is closed. Our results will apply to $\mathcal{X}$ because $\mathcal{X} \subseteq \bar{\mathcal{X}}$.

Define the vector $p := x^* - d$ and the convex combination $x_\lambda = \lambda x^* + (1-\lambda)x$ for some $\lambda \in [0,1], x \in \bar{\mathcal{X}}$. Since the set $\bar{\mathcal{X}}$ is **convex** (see exercise) then $x_\lambda \in \bar{\mathcal{X}}$, which in turn implies that $||x_\lambda - d||^2 \geq ||x^* - d||^2$, $\forall \lambda \in [0,1]$. This inequality will be the basis for the hyperplane separation, which is also an inequality. We can readily verify that $x_\lambda - d = \lambda(x^* - d) + (1-\lambda)(x-d)$, which is equal to $\lambda p + (1-\lambda)(x-d)$. Therefore, we can use Lemma 4.3.5 to rewrite the inequality.

$$\lambda^2||p||^2 + 2\lambda(1-\lambda)p^t(x-d) + (1-\lambda)^2||x-d||^2 \geq ||p||^2, \quad \lambda \in [0,1]$$

We can subtract $||p||^2$ from both sides. Notice that $(\lambda^2 - 1)||p||^2 = -(1+\lambda)(1-\lambda)||p||^2$. We

can rearrange the inequality and divide by $(1-\lambda)$ as

$$-(1+\lambda)||p||^2 + 2\lambda p^t(x-d) + (1-\lambda)||x-d||^2 \geq 0, \quad \forall \lambda \in [0,1)$$

The inequality is not defined for $\lambda = 1$ because then we cannot divide by $(1-\lambda)$. Now suppose that we take a sequence $\lambda_k \in [0,1)$ such that $\lambda_k \to 1$. We use the property that limits preserve weak inequalities[2]:

$$-2||p||^2 + 2p^t(x-d) \geq 0$$

Therefore, $p^t(x-d) \geq ||p||^2 > 0$. That means that $p^t x > p^t d$ for all $x \in \bar{\mathcal{X}}$. This $p^t x > p^t d$ for all $x \in \mathcal{X}$ since $\mathcal{X} \subseteq \bar{\mathcal{X}}$. To complete the proof set $w = p^t d + \frac{||p||^2}{2}$.

□

Corollary 4.3.2. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a non-empty, closed, convex set and $d \notin \mathcal{X}$. Then there exists a hyperplane $H(p,w)$ that strictly separated $\mathcal{X}$ and $\{d\}$. That is,*

$$p^t d < w \quad \textit{and} \quad w < p^t x, \quad \forall x \in \mathcal{X}$$

Proof. If $\mathcal{X}$ is closed, then $\bar{\mathcal{X}} = \mathcal{X}$ and we can apply the strict separating hyperplane theorem directly for the case when $d \notin \mathcal{X}$. □

[2]We can also write our set of inequalities as $g(\lambda) \geq 0$ where g is a quadratic function whose coefficients depend on constants $||p||$ and $||x-d||$. Then $\lim_{\lambda \to 1} g(\lambda) \geq 0$. The left-hand side converges because g is continuous.

4.3.4 Weak Separation

Theorem 4.3.2. *Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ is a non-empty convex set, $d \notin int(\mathcal{X})$. Then there exists a non-zero $p \in \mathbb{R}^n$ such that $p^t x \geq p^t d, \forall x \in \mathcal{X}$.*

Proof. Recall that $int(\mathcal{X})^C = \partial\mathcal{X} \cup \bar{\mathcal{X}}^C$. If $x \in \bar{\mathcal{X}}^C$ then we can apply Theorem 4.3.1 and obtain a strict separating hyperplane. Since strict separation implies weak separation.

Suppose that $d \in \partial\mathcal{X}$. By using Lemma 4.3.3 if $\mathcal{X}$ is a convex set then $\partial\mathcal{X} = \partial\bar{\mathcal{X}}$. That means that $d \in \partial\bar{\mathcal{X}}$. We will initially prove the theorem for $\bar{\mathcal{X}}$. This implies that for any integer n there exists a vector $d_n \notin \bar{\mathcal{X}}$ such that $||d - d_n|| \leq \frac{1}{n}$ (if it didn't exist then x would belong to the interior of $\bar{\mathcal{X}}$). By Theorem 4.3.1, for every integer n there exists a non-zero p_n such that

$$p_n^t(x - d_n) > 0, \quad \forall x \in \bar{\mathcal{X}}$$

Unfortunately, we cannot be sure that p_n converges. Let us transform the vector to ensure that $\tilde{p}_n = p_n/||p_n||$ has unit length. Since $||p_n|| > 0$ then we can divide the inequality on both sides:

$$\tilde{p}_n^t(x - d_n) > 0, \quad \forall x \in \bar{\mathcal{X}}$$

The set of vectors of unit length is compact (closed and bounded). Therefore, there exists a convergent subsequence such that $\tilde{p}_{n_k} \to \tilde{p}$. By construction $\tilde{p}$ is of length one and therefore, non-zero. Furthermore, since d_n converges to d, it follows that every convergent subsequence also converges, including d_{n_k}. We can take limits on both sides

$$\lim_{k\to\infty} \tilde{p}_{n_k}^t(x - d_{n_k}) > 0, \quad \forall x \in \bar{\mathcal{X}}$$
$$\tilde{p}^t(x - d) \geq 0, \quad \forall x \in \bar{\mathcal{X}}$$

The inequality is weak because limits only preserve weak inequalities. To complete the proof notice that $\mathcal{X} \subseteq \bar{\mathcal{X}}$. Therefore,

$$\tilde{p}^t(x - d) \geq 0, \quad \forall x \in \mathcal{X}$$

□

4.4 Separating Two Convex Sets

4.4.1 Operations on Convex Sets

Definition 4.4.1. Let $\mathcal{C}(\mathcal{A}) := \{\mathcal{X}_\alpha \subseteq \mathbb{R}^n : \alpha \in \mathcal{Z}\}$ be an arbitrary (finite or infinite) collection of sets indexed by a $\alpha \in \mathcal{A} \subseteq \mathbb{Z}$. Define the intersection of the sets in the collection as

$$\bigcap_{\alpha\in\mathcal{A}} \mathcal{X}_\alpha := \{x \in \mathbb{R}^n : x \in X_\alpha, \forall\alpha \in \mathcal{A}\}$$

Lemma 4.4.1. *If $\mathcal{C}(\mathcal{A})$ is a collection of convex sets. If $\bigcap_{\alpha\in\mathcal{A}} \mathcal{X}_\alpha$ non-empty then it is convex.*

Proof. Suppose that we choose $x, x' \in \bigcap_{\alpha\in\mathcal{A}} \mathcal{X}_\alpha$ (which is non-empty by assumption). Choose an arbitrary set $X_\alpha \in \mathcal{C}(\mathcal{A})$, then $x, x' \in \mathcal{X}_\alpha$ by definition. Construct $x_\lambda = \lambda x + (1-\lambda)x', \lambda \in [0,1]$. The vector $x_\lambda \in \mathcal{X}_\alpha$ because the set is convex. Since this holds for all X_α then $x_\lambda \in \bigcap_{\alpha\in\mathcal{A}} \mathcal{X}_\alpha$ for all $\lambda \in [0,1]$. This shows that the set is convex. □

Definition 4.4.2. Suppose that we have two sets $A, B \subseteq \mathbb{R}^n$. Addition and subtraction of the sets is defined, respectively, as.

$$A + B = \{z \in \mathbb{R}^n : z = a + b,\ a \in A,\ b \in B\} \qquad \text{Addition of Sets}$$
$$A - B = \{z \in \mathbb{R}^n : z = a - b,\ a \in A,\ b \in B\} \qquad \text{Subtraction of Sets}$$

Lemma 4.4.2. *Suppose that $A, B \subseteq \mathbb{R}^n$ are non-empty convex sets. Then (i) $A + B$ and $A - B$ are convex. (ii) $\mathbf{0}_{n\times 1} \in A - B$ if and only if $A \cap B \neq \emptyset$.*

Proof. (i) Choose two arbitrary elements in $x, x' \in A+B$. Then there exist vectors $a, a' \in A$ and $b, b' \in B$ such that $x = a+b,\ x' = a'+b'$. Since A and B are convex, $a_\lambda := \lambda a+(1-\lambda)a' \in A$ and $b_\lambda := \lambda b+(1-\lambda)b' \in B$. That means that $a_\lambda+b_\lambda = \lambda(a+b)+(1-\lambda)(a'+b') \in A+B$, which means that $A + B$ is convex. The proof for $A - B$ is analogous.

(ii) $\implies$ if $\mathbf{0}_{n\times} \in A - B$ then there exist $a \in A, b \in B$ such that $a - b = 0$. This implies $a = b$ and therefore $A \cap B \neq \emptyset$.

$\impliedby$ Conversely, suppose that $x \in A \cap B$, which is non-empty by assumption. Then $\mathbf{0}_{n\times} = x - x$, which is contained in $A - B$.

□

4.4.2 Weak Separation

Lemma 4.4.3. *Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^n$ be two convex sets. If $\mathcal{X} \cap \mathcal{Y} = \emptyset$ then there exists a non-zero $p \in \mathbb{R}^n$ such that $p^t x \geq p^t y$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$.*

Proof. Define the set $\mathcal{W} := \mathcal{X} - \mathcal{Y}$. Then we can plug-in the definition of the set $\mathcal{W}$ to rewrite the statement of the lemma.

$$p^t(x - y) \geq 0, \quad \forall x \in \mathcal{X}, \forall y \in \mathcal{Y} \quad \iff \quad p^t w \geq 0, \quad \forall w \in \mathcal{W}$$

By Lemma 4.4.2 the set $\mathcal{W}$ is convex and since $\mathcal{X} \cap \mathcal{Y} = \emptyset$ then $\{0_{n\times 1}\} \notin \mathcal{W}$. Since $int(W) \subseteq W$ then $\{0_{n\times 1}\} \notin int(\mathcal{W})$ and we can apply Theorem 4.3.2 to show that there exists non-zero p such that $p^t w \geq p^t 0_{n\times 1} = 0$. This completes the proof. □

4.4.3 Strict Separation

Lemma 4.4.4. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a compact set and let $\mathcal{Y} \subseteq \mathbb{R}^n$ be a closed set. Then the set $\mathcal{X} - \mathcal{Y}$ is closed.*

Proof. Define $\mathcal{W} = \mathcal{X} - \mathcal{Y}$. Then w is a limit point of $\mathcal{W}$ if there exists $w_k \in \mathcal{W}$ such that $w_k \to w$. By definition, there exist $x_k \in \mathcal{X}$ and $y_k \in \mathcal{Y}$ such that $w_k = x_k - y_k$. Since the set $\mathcal{X}$ is compact there exists a convergent subsequence k_s such that $x_{k_s} \to x^* \in \mathcal{X}$. Since w_k converges to w, the subsequence w_{k_s} also converges to w. This implies that the subsequence $y_{k_s} = x_{k_s} - w_{k_s}$ converges to some $y^* = w - x^*$. Since $\mathcal{Y}$ is closed, then $y^* \in \mathcal{Y}$.

Therefore, there exists $x^* \in \mathcal{X}$ and $y^* \in \mathcal{Y}$ such that $w = x^* - y^*$ and the limit point is contained in $\mathcal{W}$. To conclude, this means that the set $\mathcal{W}$ is closed.

□

Theorem 4.4.1. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a non-empty, convex, compact set and let $\mathcal{Y} \subseteq \mathbb{R}^n$ be a non-empty, convex, closed set. If $\mathcal{X} \cap \mathcal{Y} = \emptyset$, then there exists a non-zero $p \in \mathbb{R}^n$ and a scalar $w \in \mathbb{R}$ such that $p^t x > a > p^t y$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.*

Proof. Define the set $\mathcal{W} := \mathcal{X} - \mathcal{Y}$. By definition, we can restate the Lemma as:

$$p^t(x - y) > a, \quad \forall x \in \mathcal{X}, \forall y \in \mathcal{Y} \quad \iff \quad p^t w > a, \quad \forall w \in \mathcal{W}$$

By Lemma 4.4.2 the set $\mathcal{W}$ is convex and since $\mathcal{X} \cap \mathcal{Y} = \emptyset$ then $\{0_{n\times 1}\} \notin \mathcal{W}$. Since W is closed we can apply 4.3.2 to show that there exists a non-zero $p \in \mathbb{R}^n$ and a scalar $a^* \in \mathbb{R}$ such that $p^t w > a^* > 0$.

Let $L_{\mathcal{X}}$ denote the infimum of $p^t x$ over $\mathcal{X}$ and let $U_{\mathcal{Y}}$ be the supremum of $p^t y$ over $\mathcal{Y}$. Then we can plug-in the definition of $\mathcal{W}$ and rewrite the equation as

$$\begin{aligned}
p^t(x - y) \geq a^* \quad \forall x \in \mathcal{X},\ \forall y \in \mathcal{Y} & \\
p^t x \geq a^* + p^t y \quad \forall x \in \mathcal{X},\ \forall y \in \mathcal{Y} & \qquad \text{Rearranging Equation} \\
\inf_{x \in \mathcal{X}} p^t y \geq a^* + p^t y \quad \forall x \in \mathcal{X} & \qquad \text{Finite Infimum because of Finite RHS} \\
L_{\mathcal{X}} \geq a^* + \sup_{y \in \mathcal{Y}} p^t y & \qquad \text{Finite Supremum because of Finite LHS} \\
L_{\mathcal{X}} \geq a^* + U_{\mathcal{Y}} &
\end{aligned}$$

Set $a = \frac{a^*}{2} + U_{\mathcal{Y}}$, a midpoint of $[U_{\mathcal{Y}}, a^* + U_{\mathcal{Y}}]$. Then,

$$p^t x \geq L_{\mathcal{X}} > \ a \ > U_{\mathcal{Y}} \geq p^t y, \quad \forall x \in \mathcal{X},\ \forall y \in \mathcal{Y}$$

□

4.5 Exercises

1. For any $p \in \mathbb{R}^n \backslash \{0\}$ and $a \in \mathbb{R}$, let

$$h(p,a) \equiv \left\{x \in \mathbb{R}^n | p^T x \geq a\right\}$$

be the half space generated by the hyperplane $H(p,a)$. Assume D is a closed subset of $\mathbb{R}^n$. Let E be the intersection of all half spaces that contain D, i.e.

$$E \equiv \bigcap_{h(p,a) \supset D} h(p,a).$$

Prove D is convex if and only if $D = E$. This gives another characterization of convexity. (Hint: separating hyperplane theorem.)

2. Assume $U \subset \mathbb{R}^n$ is convex. Let $x^* \in U$ be a point. Prove the followings are equivalent:

   (a) there is no $x \in U$ such that $x_i > x_i^*$ for all $i = 1, \cdots, n$,

   (b) there exists $\lambda \in \mathbb{R}_+^n \backslash \{0\}$ such that x^* solves

$$\max_{x \in U} \lambda^T x.$$

3. Let D be a nonempty convex subset of $\mathbb{R}^n$. Prove its closure $\overline{D}$ is convex.

Chapter 5

Convex Sets (II): Cones

Some types of optimization problems are unconstrained. For example, in a linear regression the parameters are optimized over all of $\mathbb{R}^n$. However, problems involving resource allocations (in practical or theoretical problems) are bounded by capacity or other types of resource constraints.

Definition 5.0.1. Let A be an $m \times n$ matrix. The **finite cone** spanned by the column vectors $a_j \in \mathbb{R}^n$ is defined as

$$cone(A) := \{z \in \mathbb{R}^m : z = \sum_{j=1}^{n} \lambda_j a_j, \quad \lambda_j \geq 0\}$$

The set of vectors with non-negative entries is denoted by $\mathbb{R}^m_+$. For example, we could replace the restriction in the definition of a cone with $\lambda \in \mathbb{R}^m_+$ instead of $\lambda_j \geq 0$ for all j.

Definition 5.0.2 (Ordering of vectors)**.** Let $a, b \in \mathbb{R}^m$.

(i) (Weak Inequality) We say that $a \geq b$ if $a_i \geq b_i$ for all $i \in \{1, \ldots, m\}$.

(ii) (Strict Inequality I) We say that $a > b$ if $a \geq b$ and $a_{i^*} > b_{i^*}$ for at least one i^*.

(iii) (Strict Inequality II) We say that $a \gg b$ if $a_i > b_i$ for all $i \in \{1, \ldots, m\}$.

5.1 Finite Cones are Convex Sets

Lemma 5.1.1. *Let A be an $m \times n$ matrix. The cone of A is a convex set.*

Proof. Suppose that $b, b' \in Cone(A)$ then there exists a vector $\lambda, \lambda' \in \mathbb{R}^n_+$ such that $A\lambda = b$ and $A\lambda' = b'$. Let $\theta \in [0, 1]$. Define $b_\theta := \theta b + (1-\theta)b' = \theta A\lambda + (1-\theta)A\lambda'$. By linearity this is equal to $A(\theta\lambda + (1-\theta)\lambda')$. We can verify that $\lambda_\theta \geq 0$ because it is the convex combination of the two non-negative vector (each individual entry is non-negative). Therefore, $b_\theta \in Cone(A)$ for all $\theta \in [0, 1]$. That means that $Cone(A)$ is convex.

□

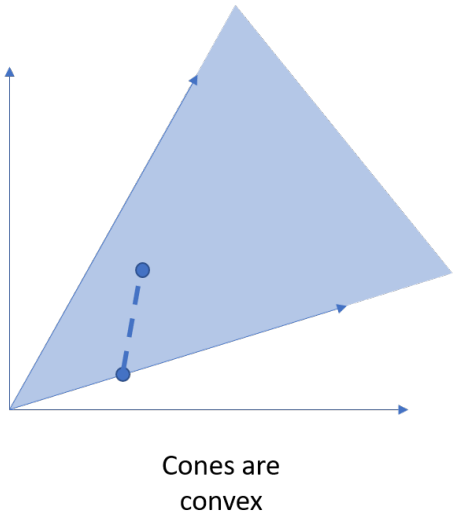

Figure 5.1: Example

5.2 Finite Cones are Closed Sets

5.2.1 Carathéodory's Theorem

Definition 5.2.1. Let A be an $m \times n$ matrix. A column submatrix of A is a matrix that is constructed by selecting some (or all) the columns of A.

Definition 5.2.2. Let A be an $m \times n$ matrix. Define $\mathcal{A} := \{A_k\}$ as the collection of column sub-matrices of A. The collection of full rank column submatrices is denoted by $\mathcal{A}^* \subseteq \mathcal{A}$.

Theorem 5.2.1. *(Carathéodory's Theorem) Let A be a non-zero $m \times n$ matrix, then*

$$Cone(A) = \bigcup_{A_k \in \mathcal{A}^*} Cone(A_k).$$

Proof. We will show the equality of the sets in two steps.

(i) $\bigcup_{A_k \in \mathcal{A}^*} Cone(A_k) \subseteq Cone(A)$: If $x \in \mathbb{R}^n$ belongs to the union of cones, then there exists a full-rank column submatrix A_k such that $x \in Cone(A_k)$. Suppose that A_k is an $m \times p$ matrix. That means that there is a vector with non-zero entries, $\lambda \in \mathbb{R}^p_+$ such that $x = A_k\lambda$. Suppose (WLOG) that A_k is constructed by dropping the last $n-p$ columns of A, which are recorded in the matrix B. Then $x = A_k\lambda_p + B\mathbf{0}_{(n-p)\times 1}$. That means that $x = A\lambda^*$, where λ^* stacks λ and $\mathbf{0}_{(n-p)\times 1}$. Since $\lambda^* \in \mathbb{R}^n_+$, then $x \in Cone(A)$.

(ii) $Cone(A) \subseteq \bigcup_{A_k \in \mathcal{A}^*}$. The first part of the proof involves some preprocessing of the matrix A. We want to discard some easy cases where the relationship holds in order to apply our main proof strategy. (I) Suppose that A is already full rank, then define $A_k = A$ and we are done. (II) Suppose that A is not full rank rank. Let's break this down into two cases: (IIa) First, suppose that $x = \mathbf{0}_{m\times 1}$. Since A is non-zero, we can always construct a full rank column submatrix by setting $A_k = a^*$, where a^* is a non-zero vector. Then we can define $\mathbf{0}_{m\times 1} = A_k\lambda$ where $\lambda = 0$. That means that $\mathbf{0} \in Cone(A_k) \subseteq \bigcup_{A_k \in \mathcal{A}^*} Cone(A_k)$.

(IIb) Now let's consider the case where x is non-zero. Since $x \in Cone(A)$, then $x = \sum_{j=1} \lambda_j a_j$, where $\lambda \in \mathbb{R}^n_+$. If some λ_j are zero then construct a new matrix by dropping some of the columns. For notational simplicity assume that this column submatrix is our new starting point, calling it A. WLOG assume that this column submatrix is not full rank, otherwise we are done.

By Lemma 2.2.1 since A is not full rank, there exists a non-zero vector $\beta \in \mathbb{R}^n$ such that $0 = \sum_{j=1}^n \beta_j a_j$. At least one β_j is non-zero and we can assume without loss of generality

that it is strictly positive (if it isn't just multiply both sides by negative one). We define the following auxiliary quantity,

$$\mu := \max_{j \in \{1,\ldots,n\}} \frac{\beta_j}{\lambda_j}, \quad \mu > 0$$

The quantity μ is positive by construction because we have preprocessed the matrix (dropping certain columns) so that all $\lambda_j > 0$ and there is at least one $\beta_j > 0$. We can rewrite the vector x as:

$$\begin{aligned}
x &= \sum_{j=1} \lambda_j a_j + \mathbf{0}_{m\times 1} && \text{(Add a zero vector)} \\
&= \sum_{j=1} \lambda_j a_j + \frac{1}{\mu} \sum_{j=1} \beta_j a_j && \text{(Because } A \text{ has a non-trivial kernel)} \\
&= \sum_{j=1} \left(\lambda_j - \frac{\beta_j}{\mu} \right) a_j && \text{(Grouping terms)} \\
&= \sum_{j=1} \lambda_j \left(1 - \frac{\left(\frac{\beta_j}{\lambda_j}\right)}{\mu} \right) a_j && \text{(Multiplying and dividing by } \lambda_j > 0) \\
&= \sum_{j=1} \tilde{\lambda}_j a_j && \text{Define } \tilde{\lambda}_j := \lambda_j \left(1 - \frac{\left(\frac{\beta_j}{\lambda_j}\right)}{\mu} \right)
\end{aligned}$$

Since $\mu = \frac{\beta_{j^*}}{\lambda_{j^*}}$ for some j^*, then at least one $\tilde{\lambda}_{j^*} = 0$. Furthermore, $\tilde{\lambda}_j \geq 0, \ \forall j \in \{1, \ldots, n\}$ because $\frac{\beta_j}{\lambda_j} \leq \mu$. Drop all the columns for which $\tilde{\lambda}_j = 0$ (there is at least one column dropped), call this A_k. Then $x \in Cone(A_k)$. If the matrix is full rank then we are done. If not repeat the process until you obtain a full rank matrix A_k. The process has to stop eventually because (i) there are a finite number of columns to start with, (ii) at least one column is dropped at every step if A_k is not full rank and (iii) x is non-zero and A has at least one non-zero column (which rules out the case there all $\tilde{\lambda}$ are zero).

□

5.2.2 Main Result

Theorem 5.2.2. *Let A be an $m \times n$ matrix, then $cone(A)$ is a closed set.*

Proof. By definition, $cone(A)$ is closed if for every sequence $x_s \in cone(A)$ such that $x_s \to x$, then $x \in cone(A)$. If A is a zero matrix, then $cone(A)$ only contains one point (the zero vector) and therefore it is closed. If A is non-zero, by Theorem 5.2.1 each x_s belongs to the cone of at least one full rank submatrix $A_k \in \mathcal{A}^*$. Assign each vector to a submatrix such that $x_s \in cone(A_k)$. Let x_{s_k} denote a subsequence of vectors assigned to matrix k. The total number of submatrices is finite because A has a finite number of columns. This implies that at least one subsequence has an infinite number of elements because of the pidgeonhole principle. WLOG assume that it is x_{s_k}. Because x_s is a convergent sequence, x_{s_k} also converges to x.

Suppose that A_k is an $m \times l$ matrix ($l \leq n$). Since it is full rank and $x_{s_k} \in cone(A_k)$, there exists a unique vector $\lambda_{s_k} \in \mathbb{R}^l_+$ such that $A_k\lambda_{s_k} = x_{s_k}$ (it is an over-identified system). It can be solved by computing $\lambda_{s_k} = (A_k'A_k)^{-1}A_k'x_{s_k} = Bx_{s_k}$. The linear map $T^{-1}(x) = Bx$ is continuous, which means that $\lim_{s_k\to\infty} \lambda_{s_k} = B \lim_{s_k\to\infty} x_{s_k}$ which is equal to $Bx = \lambda^*$. The vector $\lambda^* \in \mathbb{R}^l_+$ because every element in the sequence belongs to $\mathbb{R}_+$ and the set is closed. To complete the proof we need to show that $A_k\lambda^* = x$. The linear map $T(\lambda) := A_k\lambda$ is continuous, which means that $A_k\lambda^* = A_k\left(\lim_{s_k\to\infty} \lambda_{s_k}\right) = \lim_{s_k\to\infty}\left(A_k\lambda_{s_k}\right) = x$.

□

5.3 Farkas' Lemma

Lemma 5.3.1. *Let A be an $m \times n$ matrix and $b \in \mathbb{R}^m$ then exactly one of the following statements is true:*

1. $b \in Cone(A)$.
2. *There exists a non-zero $p \in \mathbb{R}^m$ such that $p^t A \geq \mathbf{0}_{1\times n}$ and $p^t b < 0$.*

Proof. We will prove the result by cases.

(i) Suppose that $b \notin Cone(A)$. By Lemma 5.1.1 the $Cone(A)$ is convex and because of Lemma 5.2.2, it is closed. Therefore, there exists a separating hyperplane such that $p^t x > a > p^t b$, for all $x \in Cone(A)$. We will impose a stronger version of this inequality by using the specific properties of the convex cone. Since $0 \in Cone(A)$, it follows that $p^t 0 = 0 > a$. On the other hand suppose that there exists an $x \in Cone(A)$ such that $p^t x < 0$. It follows that $x^* = \lambda x \in Cone(A)$ for any $\lambda > 0$. However, if λ is chosen sufficiently large then

$$p^t x^* = p^t(\lambda x) = \lambda p^t x < p^t b$$

which violates the hyperplane result. That means that $p^t x \geq 0$ and we can prove the stronger inequality $p^t x \geq 0 > p^t b$, $\forall x \in Cone(A)$. Furthermore, notice that the column $a_j \in Cone(A)$. That means that $p^t a_j \geq 0$ for all $j \in \{1, \ldots, n\}$ or equivalently, $p^t A \geq \mathbf{0}_{1\times n}$.

(ii) Suppose that $b \in Cone(A)$. Assume that there exists a non-zero $p \in \mathbb{R}^m$ such that $p^t A \geq 0$ and $p^t b < 0$. Since $b \in Cone(A)$ then there exists $\lambda \in \mathbb{R}_+$ such that $b = A\lambda$. Then $p^t b = p^t A\lambda$. Since $p^t A$ is a $1 \times n$ non-negative vector and λ is non-negative, then $p^t A\lambda \geq 0_{1\times 1}$. However, this violates the premise that $p^t b < 0$. This completes the proof.

□

5.4 Application: Financial Arbitrage

This section is inspired by Naiman and Scheinerman (2017).

There is a financial market which trades n different assets. The price of one unit of each asset is q_j for $j \in \{1, \dots, n\}$ and we can stack the prices in a vector $q \in \mathbb{R}^n$.

The return of each asset is risky. We will model this uncertainty by assuming that there are m different states of the world. For example, the assets could represent the stock returns of different companies and the states capture cases where the economy is performing well, average or poorly. We will assume that the returns are modeled as an $m \times n$ return matrix, R, with rows representing the state of the world and the columns the return of each company. In this model there is no time (although we can interpret R as a matrix that already discounts future returns).

$$R = \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{m1} & \cdots & R_{mn} \end{bmatrix}, \quad q = \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} \quad \Longrightarrow \quad \Pi = \begin{bmatrix} R_{11} - q_1 & \cdots & R_{1n} - q_n \\ \vdots & \ddots & \vdots \\ R_{m1} - q_1 & \cdots & R_{mn} - q_n \end{bmatrix}$$

The matrix Π represents the profits in each state, net of the assets' initial price. This can be expressed more succinctly in matrix form as $\Pi = R - \mathbf{1}_{n\times 1} q^t$. The unit vector ensures that the initial price is subtracted from the return in each state (because the initial price is paid regardless). Notice that m is not necessarily equal to n, that means that the market could be incomplete ($m > n$) or that there are redundant assets ($m < n$).

The investor decides to invest an amount x_j in each asset. Her total **portfolio** is a vector $x \in \mathbb{R}^n$. If $x_j > 0$ then the investor pays the initial price and receives a return tomorrow (a **long** position). If instead $x_j < 0$ then the investor sells the asset to somebody else today (e.g. stocks to raise capital) and agrees to pay the returns tomorrow (a **short** position). The total returns from a portfolio are $r := \Pi x$.

Definition 5.4.1. If there exists a portfolio vector $x \in \mathbb{R}^n$ such that $\Pi x \gg 0$, then we say that there is an **arbitrage** opportunity in the market.

If an arbitrage opportunity exists then an investor can ensure that she can obtain a strictly positive return in each state. These opportunities arise from a mispricing of the assets in the market. However, what does it mean to "price" the asset correctly, specially in a world with incomplete markets? This is the question that the Arbitrage Theorem attempts to answer.

Theorem 5.4.1 (Arbitrage Theorem)**.** *Let* Π *be an* $m \times n$ *net profits matrix. Then exactly one of the following statements holds.*

(a) There exists an $x \in \mathbb{R}^n$ *such that* $\Pi x \gg \mathbf{0}_{m\times 1}$ *(Arbitrage).*

(b) There exists a probability vector p *such that* $p^t\Pi = \mathbf{0}_{1\times n}$ *(Expected value pricing)*

Recall from Definition 1.6.2 that a probability vector is a vector that has non-negative entries ($\pi \geq 0$) and its entries add up to one ($p^t\mathbf{1}_{n\times 1} = 1_{1\times 1}$). The second condition can be restated as follows

$$p^t\Pi = p^tR - p^t\mathbf{1}_{n\times 1}q^t = \mathbf{0}_{1\times n} \qquad \Longrightarrow \qquad q^t = p^tR$$

The vector p^tR is vector with the average return of each asset (its expected value). Viewed in this way the Arbitration Theorem can be stated as follows: "A financial market does not have arbitrage opportunities if and only if assets are priced according to their expected returns". This is another way of stating theorems that are written as "exactly one condition must hold".

The Arbitrage theorem is an insightful application of Farkas' Lemma. The proof is interesting because it reveals that we can transform many problems involving probabilities into a linear system with cones.

Proof of Arbitrage Theorem. We break down the proof into two cases:

(i) Suppose that condition (b) holds, then there exists a probability vector such that $p^t\Pi = \mathbf{0}_{1\times n}$. If (a) holds then there exists a vector $x^* \in \mathbb{R}^n$ such that $\Pi x^* \gg \mathbf{0}_{m\times 1}$. Therefore if p is an $m \times 1$ non-negative vector with at least some positive entries,, $p^t\Pi x^* > 0$. However, if condition (b) holds then $p^t\Pi = \mathbf{0}_{1\times n}$ which implies that $p^t\Pi x^* = 0$, a contradiction.

(ii) Suppose that condition (b) fails. It is useful to rewrite the condition to rewrite it in terms of cones before negating it so that we can apply our previous results. Condition (b) can be stated as: There exists a p such that:

$$\begin{array}{c} p^t\Pi = \mathbf{0}_{1\times n} \\ p \text{ probability vector} \end{array} \iff \begin{array}{c} p^t\Pi = \mathbf{0} \\ p^t\mathbf{1}_{m\times 1} = 1 \\ p \geq 0 \end{array} \iff \begin{bmatrix} \Pi^t \\ \mathbf{1}_{1\times m} \end{bmatrix} p = \begin{bmatrix} \mathbf{0}_{n\times 1} \\ 1 \end{bmatrix},\ p \geq 0$$

Define $A_{n\times m} := \begin{bmatrix} \Pi^t \\ \mathbf{1}_{1\times m} \end{bmatrix}$ and $b_{n\times 1} := \begin{bmatrix} \mathbf{0}_{n\times 1} \\ 1 \end{bmatrix}$. Therefore the negation of the condition is $b \notin Cone(A)$ (there does not exist such a p). We can then apply Farkas' lemma (using slightly different notation for the vectors): There exists a non-zero vector $s \in \mathbb{R}^{n+1}$ such that $s^tA \geq 0$ and $s^tb < 0$. We can partition the vector as $s = \begin{bmatrix} s_\Pi \\ s_1 \end{bmatrix}$, where $s_\Pi \in \mathbb{R}^n$ and $s_1 \in \mathbb{R}$. In block partitioned form this means that:

$$\begin{array}{c} s^tA \geq \mathbf{0}_{1\times n} \\ s^tb < 0 \end{array} \iff \begin{array}{c} \begin{bmatrix} s_\Pi^t & s_1 \end{bmatrix} \begin{bmatrix} \Pi^t \\ \mathbf{1}_{1\times m} \end{bmatrix} \geq \mathbf{0}_{1\times m} \\ \\ \begin{bmatrix} s_\Pi^t & s_1 \end{bmatrix} \begin{bmatrix} \mathbf{0}_{m\times 1} \\ 1 \end{bmatrix} < 0 \end{array} \iff \begin{array}{c} s_\Pi^t\Pi^t + s_1\mathbf{1}_{1\times m} \geq \mathbf{0}_{1\times n} \\ s_1 < 0 \end{array}$$

Notice that the term s_1 is not transposed because it is a scalar. Therefore $s_\Pi^t\Pi^t \geq -s_1\mathbf{1}_{1\times m}$ which is strictly greater than zero because s_1 is strictly negative. We can transpose the result to show that $\Pi s_\Pi \geq s_1\mathbf{1}_{m\times 1} \gg 0$. Therefore condition (a) has to hold.

□

5.5 Exercises

1. There are several different characterizations of Farkas' Lemma. For example

   Lemma 5.5.1 (Farkas' Lemma V2)**.** *Let A be an $m \times n$ matrix and $b \in \mathbb{R}^m$. Then one and only one is true:*

   (i) *There exists $x \in \mathbb{R}^n$ such that $Ax \leq b$.*

   (ii) *There exists $y \in \mathbb{R}^m$ such that $y \geq \mathbf{0}_{m\times 1}$, $y^t A = \mathbf{0}_{1\times n}$, $y^t b < 0$.*

   In this exercise, you will prove the lemma.

   (a) Define $C = [A, -A, I_{m\times m}] \in \mathbb{R}^m \times \mathbb{R}^{2n+m}$. Show that condition (i) is equivalent to $b \in Cone(C)$ (Hint: Use properties of block-partitioned matrices and define a vector $z \in \mathbb{R}^{2n+m}_+$).

   (b) Show that condition (ii) is equivalent to: There exists $y \in \mathbb{R}^m$ such that $y^t C \geq \mathbf{0}_{1\times(2n+m)}$ and $y^t b < 0$.

   (c) Use the original Farkas' Lemma to prove (Version 2).

2. Consider an alternative restriction on asset prices.

   Definition 5.5.1 (Pricing Restrictions)**.** Suppose that there does not exist an $x \in \mathbb{R}^n$ such that ($q^t x \leq 0$ and $Rx > \mathbf{0}_{m\times 1}$) or such that ($q^t x < 0$ and $Rx \geq \mathbf{0}_{m\times 1}$).

   (a) Write down an economic interpretation of this condition.

   (b) Suppose that there exists a set of portfolio weights $x \in \mathbb{R}^n$ that yield positive returns in every state ($\Pi x \gg 0$). Show that $Rx > \mathbf{1}_{m\times 1} q^t x$. Give a simple example of a return matrix R, a price vector q and a portfolio x where this holds but the conditions in Definition 5.5.1 does not hold.

   (c) Suppose that there exists a probability vector $\alpha \in \mathbb{R}^m$ with **strictly positive** entries which satisfies $\alpha^t \Pi = \mathbf{0}_{1\times n}$. Show that Definition 5.5.1 is satisfied.

Chapter 6

Quadratic Forms

In previous chapters we have limited our attention to linear maps. Now we will focus on a different type of map, a **quadratic form** which generalizes quadratic function $f(x) = ax^2$ to the Euclidean space. This appears in several areas of economics. For example, the variance of a linear combination of random variables can be represented as a quadratic form. Similarly, this type of maps can be used to characterize the derivatives of certain types of function (e.g. convex) that arise frequently in decision theory.

Definition 6.0.1. Let A be an $n \times n$ matrix and let $x \in \mathbb{R}^n$. Then the function $T(x) = x^t Ax$ is a quadratic form in x.

Notice that $T : \mathbb{R}^n \to \mathbb{R}$, which means that the output is a scalar. We will illustrate the definition with an example. The variance formula is a canonical type of a quadratic form. We will not prove why it holds but rather focus on what it implies from the point of view of matrices.

Example 3. *Suppose that $Y_1, \ldots, Y_n$ is a set of random variables,and let $x_1, \ldots, x_n$ be a set of constant weights. Define the weighted average as $\bar{Y}_\omega = \sum_{i=1}^n x_i Y_i$. For example, if $x_i = \frac{1}{n}$ then $\bar{Y}$ is a simple average. Then the usual variance formula is defined as:*

$$Var(\bar{Y}_\omega) := \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j Cov(Y_i, Y_j)$$

If A is a covariance matrix with ij entries equal to $Cov(Y_i, Y_j)$ and x is a vector of weights, then $Var(\bar{Y}) = x^t Ax$.

The example with the variance illustrates that quadratic forms can be written in terms of a double sum, $x^t Ax = \sum_i \sum_j x_i x_j a_{ij}$, which can be verified using the definition of matrix multiplication.

6.1 Positive (Semi) Definite Matrices

The covariance matrix A of a vector of random variables has several interesting properties. It is symmetric because $Cov(Y_i, Y_j) = Cov(Y_j, Y_i)$. Furthermore, since all random variables have non-negative variance, then $Var(Z) \geq 0$, where $Z = \sum_i x_i Y_i$. These properties can be expressed succintly in terms of the covariance matrix.

Definition 6.1.1. Let A be an $n \times n$ matrix. The matrix A is positive semi-definite if it is symmetric and $x^t Ax \geq 0$ for all $x \in \mathbb{R}^n$.

Notice that the definition allows for the possibility that $x^t Ax = 0$. To see an example where the variance is equal to zero, assume that $Z = Y - Y \equiv 0$ where $Var(Y) > 0$. In this case the same random variable is subtracted from itself. In matrix form suppose that $Y_1 = Y$ and $Y_2 = Y$, then $Cov(Y_1, Y_2) = Cov(Y, Y) = Var(Y)$. The resulting covariance matrix is:

$$A = \begin{bmatrix} Var(Y) & Var(Y) \\ Var(Y) & Var(Y) \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \Longrightarrow x^t Ax = 0$$

The reason why the variance of the linear combination is zero in this case was because the two random variables were colinear, which suggests a link between stochastic (random) notions of colinearity and positive semi-definiteness of the covariance matrix. In fact, stochastic linear independence is best captured by a stronger property known as positive definiteness.

Definition 6.1.2. Let A be an $n \times n$ matrix. The matrix A is positive definite if it is symmetric and $x^t Ax > 0$ for all $x \in \mathbb{R}^n \backslash \{\mathbf{0}_{n\times 1}\}$.

The definition of positive definiteness excludes $x = \mathbf{0}$ because $\mathbf{0}^t A\mathbf{0} = 0$ regardless of the properties of the matrix. In the variance example, it says that any linear combination of random variables with non-zero weights has strictly positive variance.

We can also define similar notions of negative (semi) definiteness.

Definition 6.1.3. Let A be an $n \times n$ symmetric matrix.

(i) It is negative semi-definite if $x^t Ax \leq 0$ for all $x \in \mathbb{R}^n$.

(ii) It is negative definite if $x^t Ax < 0$ for all $x \in \mathbb{R}^n \backslash \{\mathbf{0}_{n\times 1}\}$.

Mathematically, a lot of derivations are similar for negative (semi) definite matrices as in the positive (semi) definite case so we will only focus on the latter.

6.1.1 Implications, Examples and Counter Examples

Lemma 6.1.1. *Let A be an $n \times n$ positive semi-definite matrix. Then the diagonal entries are all non-negative. Furthermore, if A is positive definite the diagonals are all strictly positive.*

Proof. Suppose that e_i is an elementary basis vector (includes a one in entry i and zero in all other entries). Then we can verify that $e_i^t A e_i = A_{ii}$. When A is positive semi-definite $e_i^t A e_i \geq 0$ and when it is positive definite $e_i^t A e_i > 0$. □

However, the non-negativity of the diagonals is not sufficient to ensure positive semi-definiteness. Consider the following counter example:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \implies Ax = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \implies x^t A x = -2 < 0$$

More conditions need to be imposed on the off-diagonal elements. It is hard to visualize such restrictions in general because the matrix A can have a lot of elements. Consider a restricted case where the diagonals are all one (a property satisfied by correlation matrices).

$$A = \begin{pmatrix} 1 & r_{XY} & r_{XZ} \\ r_{XY} & 1 & r_{YZ} \\ r_{XZ} & r_{YZ} & 1 \end{pmatrix}$$

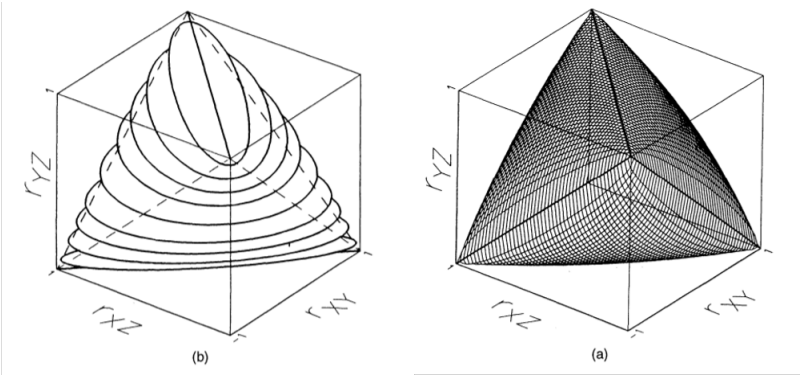

Figure 6.1: Shape of a Correlation Matrix (Rousseeuw and Molenberghs, 1994)

Figure 6.1 shows that the set of positive definite matrices as a convex set (see exercises).

In this case correlation matrices, satisfy a few interesting properties. The off-diagonal elements of the correlation matrix need to be less than or equal to 1 in absolute value, but that is not the only restriction. The valid correlations are inside the solid mesh. Typically correlations are not transitive (X positively correlated to Y and Y positively correlated to Z does not imply that X is positively correlated to Z), unless correlations are close to the boundary of the set.

6.1.2 Cholesky Decomposition

Positive scalars have a well-defined square root. The way to generalize this concept to positive definite matrices is through a Cholesky decomposition.

Definition 6.1.4. An $m \times m$ matrix A has a Cholesky decomposition if there exists a lower triangular, full rank matrix L such that $A = LL^t$.

Lemma 6.1.2. *Let A be an $m \times m$ matrix. The matrix is positive definite if and only if it has a Cholesky decomposition.*

Proof. We prove the lemma in two parts: $\Longleftarrow$ If L is a lower triangular full rank matrix. $\forall x \neq 0, L^t x \neq 0$ (because L is full rank). Then $x^t L L^t x = (L^t x)^t L^t x > 0$.

$\Longrightarrow$ We will use an induction argument.

(i) If $m = 1$, then A is a strictly positive scalar. Set $L = \sqrt{a}_{11}$.

(ii) Let $m \geq 2$. In the induction step suppose that all positive definite $(m-1) \times (m-1)$ matrices have a Cholesky decomposition.

(iii) Now we will show that the result holds for m. Since $m \geq 2$, we can write a matrix $A_{m \times m}$ in terms of blocks a_{11} (1×1 matrix), A_{12} ($1 \times (m-1)$ matrix), A_{21} ($(m-1) \times 1$ matrix), A_{22} ($(m-1) \times (m-1)$ matrix). In block form:

$$A = \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

The proof will have two steps.

(a) Define $S := A_{22} - \frac{1}{a_{11}} A_{21} A_{12}$, a candidate matrix. We will show that S is positive definite. Construct $x = \begin{bmatrix} -\frac{1}{a_{11}} A_{12} y \\ y \end{bmatrix} \in \mathbb{R}^m$ for an arbitrary $y \in \mathbb{R}^{m-1} \backslash \{\mathbf{0}\}$. The resulting x is non-zero.

$$
\begin{aligned}
0 &< x^t A x \\
&= \begin{bmatrix} (-\frac{1}{a_{11}} A_{12} y)^t & y^t \end{bmatrix} \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} -\frac{1}{a_{11}} A_{12} y \\ y \end{bmatrix} \\
&= \begin{bmatrix} (-\frac{1}{a_{11}} A_{12} y)^t & y^t \end{bmatrix} \begin{bmatrix} -\frac{a_{11}}{a_{11}} A_{12} y + A_{12} y \\ -\frac{1}{a_{11}} A_{21} A_{12} y + A_{22} y \end{bmatrix} && \text{(Right Multiplying)} \\
&= \begin{bmatrix} (-\frac{1}{a_{11}} A_{12} y)^t & y^t \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{1}{a_{11}} A_{21} A_{12} y + A_{22} y \end{bmatrix} \\
&= y^t (A_{22} - \frac{1}{a_{11}} A_{21} A_{12}) y && \text{(Left Multiplying)} \\
&= y^t S y && \text{Define } S = A_{22} - \frac{1}{a_{11}} A_{21} A_{12}.
\end{aligned}
$$

Therefore, $y^t S y > 0$ for all $y \in \mathbb{R}^{m-1} \backslash \{\mathbf{0}\}$. By Definition S is positive definite.

(iv) By the induction step, since S is a positive definite matrix of size $(m-1) \times (m-1)$ then it has Cholesky decomposition. That means that there exists a full-rank, lower triangular matrix such that $S = L_S L_S^t$. We will propose a Cholesky decomposition using a guess and verify strategy:

$$
L = \begin{bmatrix} \sqrt{a_{11}} & 0_{1 \times n} \\ \frac{1}{\sqrt{a_{11}}} A_{21} & L_S \end{bmatrix}
$$

To complete the proof we show three things:

(a) The matrix is lower triangular (by construction).

(b) It is full rank. By Lemma 6.1.1, since A is positive definite, then $\sqrt{a}_{11} > 0$. Construct a matrix $B = \begin{bmatrix} 0_{1 \times n} \\ L_S \end{bmatrix}$ which is full rank because L_S is full rank. Then the column $\begin{bmatrix} \sqrt{a}_{11} \\ \frac{1}{\sqrt{a}_{11}} A_{21} \end{bmatrix}$ is not in the image of B (because all columns of B have zero on the first row). By Corollary 2.2.1, the matrix L has to be full rank.

(c) We will show that $A = LL^t$.

$$
\begin{aligned}
LL^t &= \begin{bmatrix} \sqrt{a_{11}} & 0_{1\times n} \\ \frac{1}{\sqrt{a_{11}}}A_{21} & L_S \end{bmatrix} \begin{bmatrix} \sqrt{a_{11}} & \frac{1}{\sqrt{a_{11}}}A_{12} \\ 0_{n\times 1} & L_S^t \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & \frac{1}{a_{11}}A_{21}A_{12} + L_S L_S^t \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & \frac{1}{a_{11}}A_{21}A_{12} + S \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A
\end{aligned}
$$

where the first equality is using the fact that A is symmetric (from positive definiteness) and hence $A_{21}^t = A_{12}$. This completes the proof, therefore all positive definite matrices have a Cholesky Decomposition.

□

6.1.3 Partial Ordering

The covariance matrix of estimators can often be shown to be positive definite. In the multivariate case (a vector of estimators) we need to define a notion of when the variance of estimator is "lower" than another. Individual comparisons of variances can be helpful but incomplete. The proper notion of ordering is the following.

Definition 6.1.5. Let A and B be two positive definite matrices. Then we say $B > A$ if $B - A$ is positive definite.

This implies that $x^t(B-A)x > 0$ for all $x \in \mathbb{R}^n \backslash \{0_{n\times 1}\}$. Therefore $x^tBx > x^tAx$ and we can say that the quadratic form of A is always strictly lower than that of B. In the variance example, it says that the variance of a linear combinations of estimators (with covariance matrix A) is strictly lower than those of estimators with covariance matrix B.

6.2 Exercises

1. Let A be an $n \times n$ square matrix. Assume:

$$x^T A x = 0, \qquad \forall x \in \mathbb{R}^n. \tag{6.1}$$

   (a) Prove all diagonal components of A are $0 \in \mathbb{R}$.

   (b) Show by example that condition (6.1) does not imply $A = \mathbf{0}$.

Chapter 7

Determinants

Definition 7.0.1. Let A be an $n \times n$ matrix and let $n \geq 2$. The **minor** A_{ij} is obtained by deleting the i-th row and the j-th column of A.

Definition 7.0.2. Let A be a $n \times n$ matrix, its **determinant**, denoted by $det(A)$ is defined recursively in the following way.

1. If A is a 1×1 matrix, i.e. $\det(A) = a_{11}$.

2. The determinant for an $(n+1) \times (n+1)$ matrix A is defined as

$$det(A) = \sum_{j=1}^{n+1} (-1)^{1+j} a_{1j} det(A_{1j}).$$

For example, if A is a 2×2 matrix:

$$\det(A) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \Longrightarrow \quad det(A) := a_{11}a_{22} - a_{12}a_{21}$$

For example, let A be the following 3×3 matrix:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Then

$$\begin{aligned} \det(A) &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}). \end{aligned}$$

We present the following properties of determinants without proof.

Lemma 7.0.1 (Basic Properties of Determinants)**.** *Let $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_n]$ be an $n \times n$ matrix, where $\boldsymbol{a}_j$'s are the column vectors of A. Let $\boldsymbol{b}$ be an $n \times 1$ vector. Let c be a scalar.*

- $det([\boldsymbol{a}_1, ..., \boldsymbol{a}_j + \boldsymbol{b}, ..., \boldsymbol{a}_n]) = det([\boldsymbol{a}_1, ..., \boldsymbol{a}_j, ..., \boldsymbol{a}_n]) + det([\boldsymbol{a}_1, ..., \boldsymbol{b}, ..., \boldsymbol{a}_n])$
- $det([\boldsymbol{a}_1, ..., c\boldsymbol{a}_j, ..., \boldsymbol{a}_n]) = c\ det([\boldsymbol{a}_1, ..., \boldsymbol{a}_j, ..., \boldsymbol{a}_n])$
- $det([\boldsymbol{a}_1, ..., \boldsymbol{a}_i, ..., \boldsymbol{a}_j, ..., \boldsymbol{a}_n]) = -det([\boldsymbol{a}_1, ..., \boldsymbol{a}_j, ..., \boldsymbol{a}_i, ..., \boldsymbol{a}_n])$
- $det([\boldsymbol{a}_1, ..., \boldsymbol{a}_i, ..., \boldsymbol{a}_i, ..., \boldsymbol{a}_n]) = 0$

Lemma 7.0.2 (Properties of Determinants)**.** *Let A, B be an $n \times n$ matrix and $\alpha \in \mathbb{R}$. Then*

- $det(A) = det(A^T)$.
- $det(A) \neq 0$ *if and only if* A *is full rank.*
- $det(AB) = det(A)det(B)$.
- $det(\alpha A) = \alpha^n det(A)$.
- $det(I) = 1$.

7.1 Characteristic Polynomial

Lemma 7.1.1. *Let A be an $m \times m$ matrix and $\lambda \in \mathbb{C}$, then $\det(\lambda I - A)$ is a polynomial of degree n in λ and the coefficient of λ^n is 1.*

Proof. We will prove this by induction.

(i) For $n = 1$, $\det(\lambda I - A) = \lambda - a_{11}$. This a polynomial of degree 1 and the coefficient in front of λ is 1.

(ii) Assume for $n = k$ the determinant of a matrix in this form is a polynomial of λ with degree k and the coefficient of λ^k is 1. Consider $n = k + 1$. By expansion by the first row,

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} := |B|$$

$$= (\lambda - a_{11}) \det(B_{11}) + (-1)^{1+2}(-a_{12}) \det(B_{12}) + \cdots + (-1)^{1+n}(-a_{1n}) \det(B_{1n})$$

Note that each $\det(B_{1j})$ for $2 \leq j \leq n$ is a polynomial of degree $k - 1$. For $\det(B_{11})$, it is a polynomial of degree k and the coefficient of λ^k is 1 by our induction assumption. Hence this whole term is a polynomial of degree $k + 1$ and the coefficient of λ^{k+1} is again 1. Therefore, if A is an $n \times n$ matrix, $\det(\lambda I - A)$ is a polynomial of λ with degree n and the coefficient of λ^n is 1. We write it as

$$|\lambda I - A| = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_0.$$

We call this polynomial the **characteristic polynomial** of matrix A. We call λ is an eigenvalue of A if λ is a root of its characteristic polynomial.

□

Example 4. *If A is an $n \times n$ upper (lower) triangular matrix, then all its eigenvalues are its diagonal components:*

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ 0 & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda - a_{nn} \end{vmatrix} = \prod_{i=1}^{n} (\lambda - a_{ii})$$

Example 5. *Let*

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

then

$$det(\lambda I - A) = \lambda^2 + 1$$

Clearly, there is no real number λ satisfying $\lambda^2 + 1 = 0$. This example tells us there may be no real eigenvalues *even if the matrix A is* real. *But we know there are two roots of $\lambda^2+1 = 0$ in* complex numbers $\mathbb{C}$, *i and $-i$. That is if we consider A as a matrix over the field $\mathbb{C}$ (Recall we can always do this because $\mathbb{R}$ is a subfield of $\mathbb{C}$), it has two eigenvalues and they are in $\mathbb{C}$.* ♦

The nonexistence of real eigenvalues of a matrix A can be easily resolved if we always consider the matrix A is over the field $\mathbb{C}$ *even if it is real.* Recall:

> **Fundamental Theorem of Algebra:** Let $P(x) = x^n + b_{n-1}x^{n-1} + \cdots + b_0$ where $b_0, \cdots, b_{n-1} \in \mathbb{C}$, then there exists $x_1, \cdots, x_n \in \mathbb{C}$ such that $P(x) = \prod_{i=1}^{n}(x - x_i)$.

It tells us a polynomial with complex coefficients of degree n always has n complex roots if each root is counted up to its multiplicity. Clearly, every real coefficient polynomial can be considered as a complex coefficient polynomial. Hence $|\lambda A - I|$ always has n roots (possibly complex numbers) if each root is counted up to its multiplicity.

In this course, we will only consider matrices over $\mathbb{R}$. But from now on, when we talk about eigenvalues of a matrix A, even if A is real, we consider A as a complex matrix and allow its eigenvalues and eigenvectors to be complex. From above analysis, we can always write $det(\lambda I - A) = \prod_{i=1}^{n}(\lambda - \lambda_i)$. Thus λ is an eigenvalue of A if and only if $\lambda = \lambda_i$ for some i. In other words, $\lambda_1, \cdots, \lambda_n$ are all the eigenvalues of A and it is possible that $\lambda_i = \lambda_j$ for some $i \neq j$.

As an simple application of this result, we can have another characterization of when A is invertible.

Theorem 7.1.1. *Let A be an $n \times n$ matrix and $\lambda_1, \cdots, \lambda_n$ be the roots of its characteristic polynomial. Then*

$$\prod_{i=1}^{n} \lambda_i = \det(A)$$

Therefore A is invertible if and only if A does not have 0 eigenvalue.

Proof. Since $det(\lambda I - A) = \prod_{i=1}^{n}(\lambda - \lambda_i)$, evaluating at $\lambda = 0$ on both sides yields

$$|-A| = (-1)^n \prod_{i=1}^{n} \lambda_i$$

But the left hand side is $(-1)^n \det(A)$. Hence $\det(A) = \prod_{i=1}^{n} \lambda_i$. □

7.2 Vectorization and Continuity (Optional)

Let A, B be two $m \times n$ matrix. Then the vectorization of the matrix is

$$A = \begin{pmatrix} \uparrow & \cdots & \uparrow \\ a_1 & \cdots & a_n \\ \downarrow & \cdots & \downarrow \end{pmatrix}, \quad vec(A) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}_{(mn)\times 1}$$

where A is a matrix. We can define a distance metric for matrices based on the Euclidean norm.

$$d(A, B) = ||vec(A) - vec(B)||$$

Definition 7.2.1. Define $\Phi(L, n)$ be a collection of sets. Each set $\sigma \in \Phi(L, n)$ selects L indexes or less from $\{1, \ldots, n\}$ (possibly repeating indexes).

The size of the collection $\Phi(L, n)$ is finite because there is a finite number of permutations of the indexes.

Definition 7.2.2. . A function $f : \mathbb{R}^n \to \mathbb{R}$ is a finite multivariate polynomial if there exists a finite L such that f can be expressed as

$$f(x) = \sum_{\sigma \in \Phi(L,n)} \gamma_\sigma \prod_{k \in \sigma} x_k, \quad \gamma_\sigma \in \mathbb{R} \; forall \; \sigma \in \Phi(L, n)$$

we say that L is the order of the polynomial.

For example, $f(x) = x_1^4 + x_1 x_2 + x_1^3 x_2^2$ is a multivariate polynomial of order 5. Because polynomials are expressed as a finite addition and multiplication of elements in the vector $x \in \mathbb{R}^n$, finite polynomial functions are **continuous**.

Lemma 7.2.1. *Let A be a $n \times n$ matrix. Then* $\det(A) = f(vec(A))$*, where f is a polynomial of order n in the entries of the matrix.*

Proof. The proof procedes by induction. If $n = 2$ then, $det(A) = a_{11}a_{22} - a_{12}a_{21}$ which has a polynomial of order $L = 2$. Suppose that the determinant of every $n \times n$ matrix can be represented as a finite order polynomial of order n. We will show that this also holds for $n + 1$.

$$
\begin{aligned}
det(A) &= \sum_{j=1}^{n+1}(-1)^{1+j}a_{1j}det(A_{1j}) \\
&= \sum_{j=1}^{n+1}(-1)^{1+j}a_{1j}\sum_{\sigma\in\Phi(n,(n+1)^2)}\gamma_{j\sigma}\prod_{k\in\sigma}a_{i_k,j_k} && \text{Substituting Def. Polynomial} \\
&= \sum_{j=1}^{n+1}\sum_{\sigma\in\Phi(n,(n+1)^2)}(-1)^{1+j}\gamma_{j\sigma}\left(a_{1j}\prod_{k\in\sigma}a_{i_k,j_k}\right) && \text{Distributing Terms} \\
&= \sum_{\sigma\in\Phi(n,(n+1)^2)}\sum_{j=1}^{n+1}(-1)^{1+j}\gamma_{j\sigma}\left(a_{1j}\prod_{k\in\sigma}a_{i_k,j_k}\right) && \text{Exchanging Order of Sum} \\
&= \sum_{\sigma\in\Phi(n,(n+1)^2)}\gamma_\sigma a_{ij}\prod_{k\in\sigma}a_{i_kj_k} && \text{Substituting } \gamma_\sigma = \sum_{j=1}^{n+1}(-1)^{1+j}\gamma_{ij\sigma}. \\
&= \sum_{\sigma\in\Phi(n+1,(n+1)^2)}\tilde{\gamma}_\sigma\prod_{k\in\sigma}a_{i_kj_k} && \text{Adding One More Term to Product}
\end{aligned}
$$

In Line 2 we substitute the fact that $det(A_{ij})$ is a polynomial of order n that select elements from $vec(A)$, which has $(m + 1)^2$ elements. The third line groups together the coefficients and elements of the matrix that multiply each other.

□

We can always write a matrix in a vectorized form. That means that we can change the domain of the determinant to be $\mathbb{R}^{n^2}$.

Lemma 7.2.2. *The function* $det : \mathbb{R}^{n^2} \to \mathbb{R}$ *is a continuous function.*

Proof. Use Lemma 7.2.1 and the fact that multivariate polynomials are continuous. □

This leads to an interesting result. If a matrix is full rank, then "small" perturbations of the entries of a full rank (invertible) matrix preserve invertibility of the matrix. In other words, invertibility is not a "knife-edge" case. This is particularly important in settings where a matrix is estimated with uncertainty. It is also possible to show a similar characterization for positive definite matrices.

Corollary 7.2.1. *The set of full rank* $n \times n$ *matrices is an open set under the vectorized Euclidean norm.*

Proof. Since the determinant is a continuous function that implies that the pre-image of every open set is an open set. Let $\mathcal{A}$ denote the set of $n \times n$ full rank matrices, which can characterized by matrices with a non-zero determinant.

$$\mathcal{A} := \{vec(A) \in \mathbb{R}^{n^2} : \det(A) \neq 0\}$$

The set $U = \mathbb{R}\backslash\{0\}$ is an open set, therefore the pre-image $\det^{-1}(U)$ is also an open set, using the vectorized Euclidean norm. That means that the set of square full rank (invertible) matrices is an open set. □

7.3 Exercises

A matrix $B_{n\times n}$ is positive definite if $\forall x \in \mathbb{R}^n, x^T Bx > 0$. An equivalent definition of positive definiteness can be formulated using the determinant:

$$B = \begin{bmatrix} b_{11} & ... & b_{n1} \\ ... & ... & ... \\ b_{n1} & ... & b_{nn} \end{bmatrix}$$

Define the leading principal minor k of B, as the matrix formed by taking the upper left $(k \times k)$ submatrix. In other words:

$$B_1 = \begin{bmatrix} b_{11} \end{bmatrix}, B_2 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, ..., B_n = \begin{bmatrix} b_{11} & ... & b_{n1} \\ ... & ... & ... \\ b_{n1} & ... & b_{nn} \end{bmatrix}$$

A matrix is positive definite if and only if $\forall i \in \{1, ..., n\}, \det(B_i) > 0$. (Take this as a given, you do not need to prove it).

1. Define a function $F : \mathcal{M}_{n\times n} \to \mathbb{R}^n$. $F(B) = (\det(B_1), ..., \det(B_n))$. Reformulate the definition of positive definiteness in terms of $F(B)$.

2. Define a metric for the distance between two matrices, $d(A, B)$. Show that it is a metric: that it is non-negative, symmetric and satisfied the triangle inequality.

3. (Optional) Show that the function $F(B)$ is continuous.

4. (Optional) Show that the set of positive definite matrices of size (n) is an open set in $\mathcal{M}_{n\times n}$.

   Remark This shows that under small perturbations in the components of a positive definite matrix, the resulting matrix preserves the property of positive definiteness.

Chapter 8

Eigenvalues and Eigenvectors

The techniques we have studied so far allow us to solve static linear equation systems. In some applications it is useful to analyze the properties of **recursive** operations on matrices, which naturally arise in dynamic systems. The most useful concept in this area are eigenvalues and eigenvectors.

8.1 Review of Complex Numbers

A complex scalar $a \in \mathbb{C}$ has the form $a = a_R + a_{IM}i$ where a_R, a_{IM} are real numbers and $i = \sqrt{-1}$. Every complex scalar has a complex conjugate, which we define $\bar{a} = a_R - a_{IM}i$.

Properties 1. *The complex conjugate satisfies:*

(a) $\bar{a} = a$, *if and only if* $a_{IM} = 0$.

(b) $\overline{ab} = \bar{a}\bar{b}$.

(c) $\bar{a}a = a_R^2 + a_{IM}^2 \geq 0$.

(d) $\bar{a}a = 0$ *if and only if* $a = 0$.

Let x be a complex vector. Let A be an $m \times n$ matrix with complex entries. The conjugate of a matrix is the conjugate of the individual entries. Then we can define matrix multiplication analogously to the real numbers. We will also define the conjugate transpose, also known as the hermitian matrix.

Definition 8.1.1. Let A be an $n \times n$ matrix with $[a_{ij}] \in \mathbb{C}$. Then the matrix A^H is the **Hermitian matrix of** A if $[a_{ij}^H] = [\bar{a}_{ji}]$.

The hermitian matrix transposes the matrix and then finds the conjugate transpose of each entry.

We can define the image and the kernel in the complex numbers as well. Let A be an $m \times n$ matrix with real entries, then the image and kernel are,

$$Im(A) := \{z \in \mathbb{C}^m : z = Ax,\ x \in \mathbb{C}^n\}$$

$$Ker(A) := \{x \in \mathbb{C}^n : \mathbf{0}_{m\times 1} = Ax,\ x \in \mathbb{C}^n\}$$

Remark: The identity matrix spans the complex plane, that is $Im(I) = \mathbb{C}^n$. Therefore, using the same proof as Lemma 2.2.2 (which did not rely on the entries being real or complex) then full rank matrices have at most m columns. If a full rank matrix has $n = m$ complex-valued columns, then it is invertible.

8.2 Eigenvectors and Eigenvalues

Definition 8.2.1. Let A be a $n \times n$ real matrix. If there exists a non-zero vector $v \in \mathbb{C}^n$ and a complex scalar $\lambda \in \mathbb{C}$ such that $Av = \lambda v$, then we call λ is an eigenvalue of A, and v is the (right) eigenvector with corresponding eigenvalue λ.

Lemma 8.2.1. *Let A be an $n \times n$ matrix. Suppose that $v \in \mathbb{C}^n \backslash \{0_{n\times 1}\}$ are eigenvectors with corresponding eigenvalues $\lambda \in \mathbb{C}$, then:*

1. *If $\lambda = 0$, then A is not full rank.*
2. *$A^s v = \lambda^s v$, for all positive integers s.*
3. *The vector $\bar{v}$ is also an eigenvector of A with associated eigenvalue $\bar{\lambda}$.*

Proof. The proof has three parts:

1. Suppose that $\lambda = 0$ and that v is a non-zero vector. Then $v \in Ker(A)$. That means that $Ker(A) \neq \{\mathbf{0}_n\}$ and therefore A is not full rank.
2. The relationship holds for $t = 1$ by definition. Suppose that it holds for Step t. Then for Step $t + 1$, then $A^{t+1}x = AA^t x = A(\lambda^t v)$ which is equal to $\lambda^t Av = \lambda^{t+1}v$.
3. If $Av = \lambda v$, then we can apply the conjugate transpose. Since A has real entries, $\bar{A} = A$, which means that $\overline{Av} = A\bar{v} = \bar{\lambda}\bar{v}$. Therefore $\bar{v}$ is also an eigenvector of A with associated eigenvalue $\bar{\lambda}$.

□

8.2.1 Linear Independence and Diagonalizability

Definition 8.2.2. The vectors of $m \times n$ matrix A with complex entries are is said to be linearly independent if and only if $Ker(A) = \mathbf{0}_{n\times 1}$.

We can use this definition to show that the eigenvectors associated with distinct eigenvalues are linearly independent.

Theorem 8.2.1. *Let A be an $n \times n$ matrix. Assume $\lambda_1, \cdots, \lambda_r$ are distinct eigenvalues of A and $x_1, \cdots, x_r$ are associated eigenvectors. Then $x_1, \cdots, x_r$ are linearly independent.*

Proof. We show this by induction on r. When $r = 1$, this trivially holds because $x_1 \neq 0$. Assume $r = k < n$ and any set of k eigenvectors associated with distinct eigenvalues are linearly independent. Let $r = k + 1$. Assume $x_1, \cdots, x_{k+1}$ are eigenvectors corresponding to different eigenvalues $\lambda_1, \cdots, \lambda_{k+1}$. If

$$(*) \qquad c_1x_1 + \cdots + c_{k+1}x_{k+1} = \mathbf{0}_{m\times 1}$$

Then we can derive two new sets of restrictions:

$$\begin{aligned} \mathbf{0}_{m\times 1} &= c_1\lambda_{k+1}x_1 + \cdots + c_{k+1}\lambda_{k+1}x_{k+1} = 0 && \text{Multiplying } (*) \text{ by } \lambda_{k+1} \\ \mathbf{0}_{m\times 1} &= c_1Ax_1 + \cdots + c_{k+1}Ax_{k+1} && \text{Left-Multiplying } (*) \text{ by } A \end{aligned}$$

The last equation can be written as $c_1\lambda_1x_1 + c_2\lambda_2x_2 + \cdots + c_{k+1}\lambda_{k+1}x_{k+1} = 0$. Combining the two equations, we have that

$$c_1(\lambda_{k+1} - \lambda_1)x_1 + c_2(\lambda_{k+1} - \lambda_2)x_2 + \cdots + c_k(\lambda_{k+1} - \lambda_k)x_k = 0.$$

By induction assumption, $c_i(\lambda_{k+1} - \lambda_i) = 0$ for all $1 \leq i \leq k$. Since λ's are distinct, this is equivalent to $c_i = 0$ for all $1 \leq i \leq k$. Lastly, from $(*)$, we know $c_{k+1} = 0$ since $x_{k+1} \neq 0$. □

Definition 8.2.3. Let A be an $n \times n$ matrix and let Λ be a diagonal matrix of eigenvalues. Then the matrix A is said to be diagonalizable if there exists an $n \times n$ full rank matrix B such that $A = B\Lambda B^{-1}$.

Lemma 8.2.2. *Let A be an $n \times n$ matrix. If A is diagonalizable, then $A^t = B\Lambda^t B^{-1}$ for all positive integers t.*

Proof. The relationship holds for $t = 1$. Suppose that it also holds for Step t. Then for Step $(t + 1)$ □

Theorem 8.2.2. *Let A be an $n \times n$ matrix. A is diagonalizable if and only if A has n linearly independent eigenvectors.*

Proof. $\Longleftarrow$ Let B be a matrix whose columns are eigenvectors of A, which we denote b_j. Then $Ab_j = \lambda_j b_j$ for $j \in \{1, \ldots, n\}$. We can express this in matrix form as $AB = B\Lambda$, where Λ is a diagonal matrix with λ_j's on its diagonal. Since B is full rank, then $A = B\Lambda B^{-1}$.

$\Longrightarrow$ Suppose that A is diagonalizable and $A = B\Lambda B^{-1}$. Then we can reverse the steps to show that $AB = B\Lambda$ and $Ab_j = \lambda_j b_j$.

□

8.2.2 Symmetry

Matrices are guranteed to have real eigenvalues and eigenvectors when they are symmetric.

Theorem 8.2.3. *If A is an $n \times n$ symmetric real matrix. Then all its eigenvalues are real. Moreover, for each eigenvalue λ, there exist real eigenvectors.*

Proof. Let λ be an eigenvalue of A and $x \neq \mathbf{0}$ be a corresponding eigenvector. Then $Ax = \lambda x$. Hence the conjugate transpose is equal to:

$$\begin{aligned} \lambda \overline{x}^t x &= \overline{x}^t A x \\ &= (A^t \overline{x})^t x \\ &= \overline{Ax}^t x \\ &= \overline{\lambda} \overline{x}^t x \end{aligned}$$

Since $\overline{x}^t x \neq 0$, we have $\lambda = \overline{\lambda}$. That is λ is real. Moreover, suppose that $x \in \mathbb{C}$ where $x = a + bi$. Then by Lemma 8.2.1 the vector $\bar{x} = a - bi$ is also an eigenvector associated with $\bar{\lambda} = \lambda$. Because eigenvectors are non-zero, then either $a \neq 0$ or $b \neq 0$. Let us consider each case separately:

1. If $a \neq 0$. Then define $z = x + \bar{x} = 2a$. Then $Az = A(x + \bar{x}) = \lambda x + \lambda \bar{x}$, which is equal to $\lambda(x + \bar{x}) = \lambda z$. Therefore, z is a real non-zero vector such that $Az = \lambda z$.

2. If $b \neq 0$. Then define $z = i(x - \bar{x}) = -2b$. Then $Az = iA(x - \bar{x}) = i(\lambda x - \lambda \bar{x}) = \lambda z$. Therefore, z is a real non-zero vector such that $Az = \lambda z$.

□

The spectral theorem shows that every symmetric matrix is diagonalizable.

Definition 8.2.4. An matrix $m \times n$ matrix is orthogonal if $A^t A = I$.

Notice that orthogonal matrices satisfy the property that $A^t = A^{-1}$.

Theorem 8.2.4 (Spectral Theorem)**.** *Let A be an $n \times n$ symmetric matrix. There exists an orthogonal matrix Q such that $A = Q^t \Lambda Q$ where Λ is a diagonal matrix whose diagonal components are eigenvalues of A.*

Proof. We only need to show A has n orthogonal eigenvectors. Let $\lambda_1 = \max_{\|v\|=1} v^T A v$ and $v_1 \in \arg\max_{\|v\|=1} v^T A v$. Define $W_1 = \text{span}\{v_1\} = \{c_1 v_1 | c_1 \in \mathbf{R}^n\}$. We show λ_1 is an eigenvalue of A and v_1 is a corresponding eigenvector. Notice if we know $Av_1 = c_1 v_1$ for

some c_1 and $\lambda_1 = v_1^T A v_1 = c_1 v_1^T v_1 = c_1$ showing λ_1 is indeed an eigenvalue. Hence we only need to show $Av_1 \in W_1$. We show this by showing $Av_1 \perp W_1^\perp$. For any arbitrary $w \in W_1^\perp$, we know

$$\frac{v_1 + aw}{\|v_1 + aw\|} = \frac{v_1 + aw}{\sqrt{1 + a^2\|w\|^2}}$$

is a normal vector for any $a \in \mathbb{R}$. Hence by definition of v_1, we have

$$\begin{aligned} v_1^T A v_1 &\geq \frac{1}{1 + a^2\|w\|^2}(v_1 + aw)^T A (v_1 + aw) \\ &= \frac{v_1^T A v_1 + 2a v_1^T A w + a^2 w^T A w}{1 + a^2\|w\|^2} \end{aligned}$$

for all $a \in \mathbb{R}$, where the equality comes from the assumption that A is symmetric. But for this inequality to hold for all a, we must have $v_1^T A w = 0$ otherwise we can always find arbitrary small a such that this inequality is violated. Since $w \in W_1^\perp$ is arbitrary, we showed $Av_1 \in W_1$.

Suppose we have defined $v_1, \cdots, v_k$ and $\lambda_1, \cdots, \lambda_k$. Let

$$v_{k+1} \in \arg \max_{\substack{v \in \text{span}^\perp\{v_1, \cdots, v_k\} \\ \|v\|=1}} v^T A v,$$

$$\lambda_{k+1} = \max_{\substack{v \in \text{span}^\perp\{v_1, \cdots, v_k\} \\ \|v\|=1}} v^T A v$$

and $W_{k+1} = \text{span}\{v_{k+1}\}$. Then we can apply the same logic to show that $Av_{k+1} \in W_{k+1}$ and thus $Av_{k+1} = \lambda_{k+1} v_{k+1}$. Since this is finite dimensional, we can complete this process when $k = n$. This completes the proof. □

8.3 Exercises

Let P be an $n \times n$ matrix. Define a *stochastic matrix* P as an $n \times n$ matrix that has non-negative entries where the entries of each column sum to one. Let π be a non-negative vector whose entries sum to one. Show that π does not belong to the kernel. Further show that $P\pi$ is a vector whose entries sum to one

1. This questions studies the convergence properties of stochastic matrices:

   (a) (3 points) Now suppose that $\lim_{m\to\infty} P^m \to P^*$. Show that P^* is also a markov matrix and show that π does not belong to its kernel. (Hint: Show that every P^n is markov).

   (b) (5 points) Suppose that P^* is such that for every π, $P^*\pi = \pi^*$, for a fixed π^*. Write down what the matrix P^* has to be for $\pi^* = (0.2, 0.3, 0.4, 0.1)$ if P^* is 4×4.

   (c) (2 points) Construct an example of a 2×2 symmetric matrix P that doesn't converge. (Hint use zeros and ones only). Compute its eigenvalues. Use the spectral decomposition to give a reason why it doesn't converge.

   (d) (3 points) Show that the following asymmetric P converges to a P^* such that $P^*\pi = \pi^*$. Compute P^* and π^*.

$$P = \begin{bmatrix} 0.5 & 0 \\ 0.5 & 1 \end{bmatrix}$$

2. This questions asks you to analyze the eigenvalues of stochastic matrices:

   (a) (3 points) Let $v \in \mathbb{R}^n$. Show that the entries of the vector Pv add up to $\sum_{j=1}^n v_j$.

   (b) (9 points) Let $v^* \in \mathbb{R}^n, v \neq 0$ be an eigenvector of P, with corresponding eigenvalue λ. Prove the following statements:

      i. (1 point) $P^s v^* = \lambda^s v^*, s \in \mathbb{N}$.

      ii. (4 points) Show that if $\sum_{j=1}^n v_j^* \neq 0$, then $\lambda = 1$. [Hint: show that P^s is also markov].

      iii. (4 points) Show that if $\sum_{j=1}^n v_j^* = 0, v^* \neq 0$, then $|\lambda| \leq 1$. [Hint: show that for any fixed $v \neq 0$ (not necessarily an eigenvector), $\sup_P ||Pv|| \leq M < \infty$, P markov].

Part II

Differentiation

Chapter 9

Introduction to Differentiation

9.1 Review of Convergence

Recall the definition of convergence from the first part of the course:

Definition 9.1.1. Let $f : [a, b] \to \mathbb{R}$, be a real-valued function, $p \in [a, b]$. We write $f(x) \to q$ as $x \to p$ or $\lim_{x\to p} f(x) = q$ if there exists $q \in \mathbb{R}$ with the following property: $\forall \epsilon > 0$ there exists $\delta > 0$ such that:

$$|f(x) - q| < \epsilon, \quad \forall x \in (p - \delta, p + \delta) \cap [a, b], \quad x \neq p$$

Definition 9.1.2. $\lim_{x\downarrow p} f(x) = q$ if $\exists q \in \mathbb{R}$ such that $\forall \epsilon > 0, \exists \delta > 0$ such:

$$|f(x) - q| < \epsilon, \forall x \in (p, p + \delta) \cap [a, b] \backslash \{p\}$$

The following theorem gives four different characterizations to the definitions above:

Theorem 9.1.1 (Equivalent Limit Definitions)**.** *Let $f : [a, b] \to \mathbb{R}$, be a real-valued function, $p \in [a, b]$. The following are equivalent:*

1. $\lim_{x\to p} f(x) = q$
2. $\lim_{n\to\infty} f(x_n) = q$, *for every sequence* $\{x_n\}_{n=0}^{\infty}, \quad x_n \neq p$ *such that* $\lim_{n\to\infty} x_n = p$
3. $\lim_{x\downarrow p} f(x) = \lim_{x\uparrow p} f(x) = q$
4. $\lim_{n\to\infty} f(x_n^+) = \lim_{n\to\infty} f(x_n^-) = q$, $\forall \{x_n^+\}_{n=0}^{\infty}$, $\forall \{x_n^-\}_{n=0}^{\infty}$ *such that* $x_n^+ > p > x_n^-$ *and* x_n^+, x_n^- *converge to* p.

5. $\lim_{x\to p} |f(x) - q| = 0$.

These equivalences are proven in the Appendix. They will be very useful when we want to show whether a function is differentiable or not. This will become clear soon.

Definition 9.1.3. The function f is continuous at some point c of its domain $[a, b]$ if

$$\lim_{x\to c} f(x) = f(c).$$

Theorem 9.1.2 (Change of Variable)**.** *If there exists $q \in \mathbb{R}$ such that $\lim_{y\to y_0} h(y) = q$ and $f(x_0) = y_0$ and f is a continuous function, then $\lim_{x\to x_0} h(f(x)) = \lim_{y\to y_0} h(y)$.*

Proof. Let $x_n \in \mathbb{R}$ be a convergent sequence such that $x_n \to x_0$. Then $y_n = f(x_n)$ is a convergent sequence because f is continuous (using the second definition), which converges to $y_0 = f(x_0)$. If $\lim_{y\to y_0} h(y_n) = q$ that means that for every sequence y_n such that $y_n \to y_0$, $\lim_{y_n\to y_0} h(y_n) = q$. □

9.2 Definition of Differentiability

Definition 9.2.1. Let $f : [a, b] \to \mathbb{R}$, be a real-valued function. We say that f is differentiable at $x_0 \in [a, b]$ if the limit:

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists and is finite. We denote it by:

$$f'(x_0) := \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

and say that $f'(x_0)$ is the derivative of f at x_0. If f is differentiable at every $x \in [a, b]$, we say that f is differentiable.

Given Theorem 9.1.1 we can characterize the derivative of a function as:

Theorem 9.2.1. *Let $f : [a, b] \to \mathbb{R}$, be a real-valued function, $x_0 \in [a, b]$. The following are equivalent:*

1. *f is differentiable at x_0, with derivative $f'(x_0)$*
2. *$\lim_{n\to\infty} \frac{f(x_n)-f(x_0)}{x_n-x_0} = f'(x_0)$, for every sequence $\{x_n\}_{n=0}^{\infty}, x_n \neq x_0$, such that $\lim_{n\to\infty} x_n = x_0$*

3. $\lim_{x\downarrow x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\uparrow x_0} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$

9.2.1 Examples

The following example computes the derivative of some common functions and shows some functions that are not differentiable.

Example 6. *1. Let $c \in \mathbb{R}$ and $f(x) = c$. Then $f'(x) = 0, \forall x \in \mathbb{R}$.*

2. Let $n \geq 1$ and $f(x) = x^n$. Then $f'(x) = nx^{n-1}$.

3. Let $f(x) = e^x$. Then $f'(x) = e^x$.

4. $f(x) = |x|$ is not differentiable at $x = 0$.

5. $f(x) = \begin{cases} x \cdot \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ *is not differentiable at $x = 0$.*

Proof.

1. $f(x) = c$:
$$f'(x_0) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x\to x_0} \frac{c - c}{x - x_0} = 0$$

2. $f(x) = x^n$:

$$\begin{aligned} f(x) - f(x_0) &= x^n - x_0^n \\ &= (x - x_0) \cdot (x^{n-1} + x^{n-2}x_0 + x^{n-3}x_0^2 + \ldots + x \cdot x_0^{n-2} + x_0^{n-1}) \\ &= (x - x_0) \sum_{k=0}^{n-1} x^k x_0^{n-k-1} \end{aligned}$$

Then,

$$\begin{aligned} f'(x_0) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} &= \lim_{x\to x_0} \sum_{k=0}^{n-1} x^k x_0^{n-k-1} \\ &= \sum_{k=0}^{n-1} x_0^k x_0^{n-k-1} \\ &= nx_0^{n-1} \end{aligned}$$

3. $f(x) = e^x$:

$$\begin{aligned}\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} &= \lim_{x\to x_0} \frac{e^x - e^{x_0}}{x - x_0} \\ &= e^{x_0}\left(\lim_{x\to x_0} \frac{e^{x-x_0} - 1}{x - x_0}\right) \\ &= e^{x_0}\left(\lim_{x\to 0} \frac{e^x - 1}{x}\right)\end{aligned}$$

Recall the definition of e,

$$e := \lim_{y\to 0}(1+y)^{1/y}$$

Using the continuity of log at 1,

$$\lim_{y\to 0} \quad log(1+y)^{1/y} = \lim_{y\to 0} \quad \frac{1}{y} log(1+y) = 1$$

Using the continuity of $1/x$ at 1,

$$\lim_{y\to 0} \quad \frac{y}{log(1+y)} = 1$$

Defining $x = log(1+y)$,

$$\lim_{x\to 0} \frac{e^x - 1}{x} = 1$$

Thus,

$$f'(x_0) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = e^{x_0}\left(\lim_{x\to 0} \frac{e^x - 1}{x}\right) = e^{x_0}$$

4. $f(x) = |x|$:

We prove it using the characterizations given by Theorem 9.2.1:

$$\lim_{x\downarrow x_0} \frac{f(x) - f(0)}{x} = 1 \neq \lim_{x\uparrow x_0} \frac{f(x) - f(0)}{x} = -1$$

Given that the limits from above and from below are not equal to each other, the function is not differentiable at 0.

5. $f(x) = \begin{cases} x \cdot \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$:

$$\frac{f(x)-f(0)}{x} = \frac{x\sin(\frac{1}{x})}{x} = \sin\left(\frac{1}{x}\right)$$

Consider the sequences $x_n = \frac{1}{2n\pi}$ and $y_n = \frac{1}{2n\pi+\pi/2}, n \in \mathbb{N}$. They both converge to 0.

$$\sin(1/x_n) = \sin(2n\pi) = 0, \quad n \in \mathbb{N}$$

$$\sin(1/y_n) = \sin(2n\pi + \pi/2) = 1, \quad n \in \mathbb{N}$$

By Theorem 9.2.1, f is not differentiable.

□

Below lists several useful differentiation rules without proof.

Theorem 9.2.2. *Suppose $f, g : [a, b] \to \mathbb{R}$, be real-valued functions, differentiable at $x_0 \in [a, b]$ and $k \in \mathbb{R}$. Then:*

1. $(kf)'(x_0) = kf'(x_0)$
2. $(f + g)'(x_0) = f'(x_0) + g'(x_0)$
3. $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$
4. *if $g(x_0) \neq 0$, then* $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0)-g'(x_0)f(x_0)}{g(x_0)^2}$

Proof. See Rudin et al. (1964) □

Example 7. *Let $n \geq 1$ and $f(x) = x^n$. Then $f'(x) = nx^{n-1}$. Here we provide an easier proof using the product rule.*

Proof. By induction. For the base case, $n = 1$:

$$f(x) = x, \quad f'(x_0) = \lim_{x\to x_0} \frac{x - x_0}{x - x_0} = 1 = nx_0^{n-1}$$

Assume it holds for $n = k$:

$$f(x) = x^k, \quad f'(x_0) = kx_0^{k-1}$$

For $n = k + 1$:

$$f(x) = x^{k+1} = x \cdot x^k$$

Using the product rule in Theorem 9.2.2,

$$f'(x_0) = 1 \cdot x_0^k + x \cdot kx_0^{k-1} = x_0^k + kx_0^k = (k+1)x_0^k$$

□

9.3 Differentiability Implies Continuity

The following theorem states the relationship between continuity and differentiation of a function.

Theorem 9.3.1. *Let $f : [a, b] \to \mathbb{R}$, be a real-valued function, differentiable at $x_0 \in [a, b]$. Then f is continuous at x_0.*

Proof.

$$\begin{aligned} f(x) &= f(x) + f(x_0) - f(x_0) && (9.1) \\ &= f(x_0) + \left(\frac{f(x) - f(x_0)}{x - x_0}\right) \cdot (x - x_0) && (9.2) \end{aligned}$$

Taking limits when $x \to x_0$,

$$\lim_{x \to x_0} f(x) = f(x_0) + \lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0}\right) \cdot (x - x_0) \quad (9.3)$$

Given that $\lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0}\right)$ exists and is finite and $\lim_{x \to x_0} x - x_0 = 0$,

$$\Rightarrow \quad \lim_{x \to x_0} f(x) = f(x_0)$$

□

Note that the converse of this theorem is not true. That is, if a function $f : [a, b] \to \mathbb{R}$ is continuous at some point $x_0 \in [a, b]$, it need not be differentiable at that point. For example, the function

$$f(x) = \begin{cases} x \cdot \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is continuous at 0, but is not differentiable at that point. Here is another simpler example. $f(x) = |x|$ is continuous at $x = 0$ but not differentiable at $x = 0$.

9.4 First Order Conditions

Definition 9.4.1. Let $f : A \to \mathbb{R}$, where $A \subseteq \mathbb{R}$.

1. $x_0 \in A$ is a global maximum of f if $f(x_0) \geq f(x), \quad \forall x \in A$
2. $x_0 \in A$ is a global minimum of f if $f(x_0) \leq f(x), \quad \forall x \in A$
3. $x_0 \in A$ is a local maximum of f if $\exists \delta > 0$ such that $f(x_0) \geq f(x), \quad \forall x \in A \cap (x_0 - \delta, x_0 + \delta)$
4. $x_0 \in A$ is a local minimum of f if $\exists \delta > 0$ such that $f(x_0) \leq f(x), \quad \forall x \in A \cap (x_0 - \delta, x_0 + \delta)$

Theorem 9.4.1. *Let $f : [a, b] \to \mathbb{R}$. If f has a local maximum (minimum) at $x_0 \in (a, b)$ and f is differentiable at x_0, then $f'(x_0) = 0$.*

Proof. Let f have a local maximum at $x_0 \in (a, b)$. Then, there exists a $\delta > 0$ such that:

- For all $x \in (x_0, x_0 + \delta)$:

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0 \quad \Rightarrow \lim_{x \downarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

- For all $x \in (x_0 - \delta, x_0)$:

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0 \quad \Rightarrow \lim_{x \uparrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

The limits exist, given that f is differentiable at x_0. By Theorem 9.2.1,

$$0 \leq f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

□

Comment 1. The reserve is not true. A counterexample is $f(x) = x^3$. Note that $f'(0) = 0$ but 0 is not a local maximum or minimum.

Comment 2. The assumption of differentiability is important. If we drop the condition on differentiability, the result is no longer true. A counterexample is $f(x) = |x|$. Note that 0 is indeed a local minimum (actually a global minimum) of $f(x)$, but the derivative at $x = 0$ is not zero, simply because it is not even differentiable at $x = 0$.

9.5 Intermediate Value Theorem

Theorem 9.5.1 (Intermediate Value Theorem for Derivatives)**.** *Let $f : [a, b] \to \mathbb{R}$ continuous and differentiable on $[a, b]$. If $f'(a) < \lambda < f'(b)$, there exists $x \in (a, b)$ such that $f'(x) = \lambda$.*

Proof. Let λ such that $f'(a) < \lambda < f'(b)$. Define $g(t) := f(t) - \lambda t$. Then:

$$g'(t) = f'(t) - \lambda, \quad g'(a) < 0, \quad g'(b) > 0$$

This means that g is decreasing on a and increasing on b, so we can find $x_1, x_2 \in (a, b)$ such that $g(x_1) < g(a)$ and $g(x_2) < g(b)$. Thus, g attains a minimum at some x in the interior of $[a, b]$. By Theorem 9.4.1, $g'(x) = f'(x) - \lambda = 0$. Then:

$$f'(x) = \lambda$$

□

Comment. It is incorrect to invoke the intermediate value theorem for continuous functions. This is because derivatives are not necessarily continuous. For example, the function:

$$f(x) = \begin{cases} x^2 \cdot \sin(\frac{1}{x}) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is differentiable at every point (notice the difference with a similar example discussed previously). However, the derivative is not continuous at 0. Although we cannot claim that the derivative of a function is continuous, derivatives and continuous functions have something in common: they take on all the intermediate values.

9.6 Chain Rule

Theorem 9.6.1 (Chain rule)**.** *Let I and J be two intervals in $\mathbb{R}$. Let $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$, $f(I) \subseteq J$. If f is differentiable at $x_0 \in I$ and g is differentiable at $f(x_0) \in J$, then:*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$$

Proof. Define the residuals for $x \neq x_0$, $y \neq y_0$:

$$r(x; x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0)$$

$$u(y; y_0) = g(y) - g(y_0) - g'(y_0)(y - y_0)$$

By the definition of differentiability of f and g:

$$\lim_{x \to x_0} \frac{r(x, x_0)}{x - x_0} = \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right] = 0$$

$$\lim_{y \to y_0} \frac{u(y, y_0)}{y - y_0} = \lim_{y \to y_0} \left[\frac{g(y) - g(y_0)}{y - y_0} - f'(y_0) \right] = 0$$

Before we take limits, we rewrite the expression for the slope of $(g \circ f)(x_0)$. Let $y = f(x)$:

$$\begin{aligned}
\frac{g(f(x)) - g(f(x_0))}{(x - x_0)} &= \frac{g(y) - g(y_0)}{(x - x_0)} \\
&= \frac{g'(y_0)(y - y_0) + u(y; y_0)}{(x - x_0)} && \text{Rewriting } u(y; y_0) \\
&= \left[g'(y_0) + \frac{u(y; y_0)}{(y - y_0)} \right] \frac{(y - y_0)}{(x - x_0)} && \text{Factorizing } (y - y_0). \\
&= \left[g'(y_0) + \frac{u(y; y_0)}{(y - y_0)} \right] \frac{(f(x) - f(x_0))}{(x - x_0)} && \text{Substituting } y = f(x) \\
&= \left[g'(y_0) + \frac{u(y; y_0)}{(y - y_0)} \right] \left[f'(x_0) + \frac{r(x; x_0)}{x - x_0} \right] && \text{Rewriting } r(x; x_0)
\end{aligned}$$

This means that overall slope of $(g \circ f)$ between x_0 and x is equal to the multiplication of the slope of g between y_0 and y times the slope of f between x_0 and x. We can take limits

on both sides:

$$\lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{(x-x_0)}=\lim_{x\to x_0}\left[g'(y_0)+\frac{u(y;y_0)}{(y-y_0)}\right]\lim_{x\to x_0}\left[f'(x_0)+\frac{r(x;x_0)}{x-x_0}\right]$$

In order to distribute the limits we need to show that the individual limits exists and are finite, using the property that we used for $r(x;x_0)$ and $u(y;y_0)$.

$$\lim_{x\to x_0}\left[f'(x_0)+\frac{r(x;x_0)}{x-x_0}\right]=f'(x_0)+\lim_{x\to x_0}\left[\frac{r(x;x_0)}{x-x_0}\right]=f'(x_0)$$

$$\lim_{x\to x_0}\left[g'(y_0)+\frac{u(y;y_0)}{y-y_0}\right]=g'(y_0)+\lim_{x\to x_0}\left[\frac{u(y;y_0)}{y-y_0}\right]=g'(f(x_0))$$

The last result follows by using Theorem 9.1.2, setting $h(y)=\frac{u(y;y_0)}{y-y_0}$. Therefore $\lim_{x\to x_0}\left[\frac{u(f(x);y_0)}{f(x)-f(x_0)}\right]=\lim_{y\to y_0}\left[\frac{u(y;y_0)}{y-y_0}\right]$. This completes the proof, because it shows that:

$$\lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{(x-x_0)}=g'(y_0)f'(x_0)=g'(f(x_0))f'(x_0)$$

□

9.7 Properties Appendix: Equivalent Notions of Continuity

Definition 9.7.1. Let $f : [a, b] \to \mathbb{R}$, be a real-valued function, $p \in [a, b]$. We write $f(x) \to q$ as $x \to p$ or $\lim_{x\to p} f(x) = q$ if there exists $q \in \mathbb{R}$ with the following property: $\forall \epsilon > 0$ there exists $\delta > 0$ such that:

$$|f(x) - q| < \epsilon, \quad \forall x \in (p - \delta, p + \delta) \cap [a, b], \quad x \neq p$$

Proof. 1. We will prove (1) $\iff$ (5).

Suppose that for every $\epsilon > 0$ there exists $\delta > 0$ such that :

$$|f(x) - q| < \epsilon, \ \forall x \in (p - \delta, p + \delta) \cap [a, b], x \neq p$$

Since $|f(x) - q| = |\ |f(x) - q| - 0|$, the two notions of convergence are equivalent.

2. We will prove that (1) $\iff$ (2).

$\implies$ Suppose that $\lim_{x\to p} f(x) = q$. We will show that every sequence needs to converge.

Suppose that not every sequence $f(x_n)$ converges to $q : \exists \{x_n\}_{n=0}^{\infty}, x_n \neq p, x_n \to p$ such that $\exists \epsilon > 0$ s.t. $\forall N \in \mathbb{N}, \exists n \geq N$ such that $d(f(x_n), q) > \epsilon$. However, we also know that $\forall \epsilon > 0, \exists \delta > 0$ such that $d(f(x), q) < \epsilon$ for all $x \in (p-\delta, p+\delta) \cap [a, b] \backslash \{p\}$. Because x_n convergent, we know that for very large n, x_n must be contained in $(p - \delta, p + \delta) \cap [a, b] \backslash \{p\}$. This means that $d(f(x_n), q) < \epsilon$ for large enough n. This contradicts the assumption that not every sequence $f(x_n)$ converges to q. Therefore, we have shown that every sequence **must** converge to q.

$\impliedby$ Suppose that it is true that if $x_n \to p$ then $\lim_{n\to\infty} f(x_n) \to q$. Now we will show the $\epsilon - \delta$ definition holds.

Suppose that the either $\lim_{x\to p} f(x)$ does not exist, or it does not converge to q. Then $\exists \epsilon > 0$ such that $\forall \delta > 0, \exists x \in (p-\delta, p+\delta) \cap [a, b] \backslash \{p\}, d(f(x), q) > \epsilon$. This means that if we set $\delta = \delta_n = 1/n$ we can choose an choose x_n such that $d(f(x_n), q) > \epsilon$. Because $\delta_n \to 0, x_n \to p, x_n \neq p$. However, because every $f(x_n)$ converges to q, for large n then $d(f(x_n), q) < \epsilon$. This generates a contradiction with how x_n was defined. Therefore, $\lim_{x\to p} f(x)$ must exist and be equal to q.

3. Now we will prove (1) $\iff$ (3).

$\implies$ Suppose that if $\lim_{x\to p} f(x) = q$. Since $(p, p+\delta) \subset (p-\delta, p+\delta)$ it always follows that $\exists \delta > 0$ such that $d(f(x), q) < \epsilon$, for all $x \in (p, p+\delta) \cap [a, b] \backslash \{p\}$.

$\impliedby$ If $\lim_{x\uparrow p} f(x) = \lim_{x\downarrow p} f(x) = q$, then $\forall \epsilon > 0, \exists \delta_1, \delta_2 > 0$ such that:

$$d(f(x), q) < \epsilon, \forall x \in V_1 = (p, p+\delta_1) \cap [a, b] \backslash \{p\}$$

$$d(f(x), q) < \epsilon, \forall x \in V_2 = (p-\delta, p+\delta_2) \cap [a, b] \backslash \{p\}$$

Then set $\delta^* = \min\{\delta_1, \delta_2\}$. Then $V^* = (p-\delta^*, p+\delta^*) \cap [a, b] \backslash \{p\} \subset V_1 \cap V_2$. Therefore, $d(f(x), q) < \epsilon, \forall x \in V^*$. This completes the proof.

4. Now we will prove (2) $\iff$ (4).

   $\implies$ Suppose that it is true that if $x_n \to p$ then $\lim_{n\to\infty} f(x_n) \to q$. Then since x_n^+ ad x_n^- are special cases of this type of sequence, then $\lim_{n\to\infty} f(x_n^+) = \lim_{n\to\infty} f(x_n^-) = q$.
   $\impliedby$ Now suppose that $\lim_{n\to\infty} f(x_n^+) = \lim_{n\to\infty} f(x_n^-) = q$ for all x_n^+ and x_n^-. Let x_n be an arbitrary convergent sequence (with values potentially above and below p). Now construct two sub-sequences $f(x_n^-)$ and $f(x_n^+)$, that separate the terms in x_n that are below and above p, respectively. You start off with the first element in x_n, if it is below p assign it to x_n^-, otherwise to x_n^+. Repeat this process for all x_n in order to construct a sequence. That means that $\forall \epsilon > 0$, $\exists N_1$ such that $\forall n \geq N_1, d(f(x_n^-, q)) < \epsilon$ and $\exists N_2$ such that $\forall n \geq N_2, d(f(x_n^+), q) < \epsilon$. Let $N^* > N_1 + N_2$. Then $\forall n \geq N^*, d(f(x_n), q) < \epsilon$. Since we found an N^* for every $\epsilon > 0$, then $f(x_n)$ must converge to q.

□

9.8 Exercises

1. Let $f(x) = \begin{cases} x^\alpha \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$. For what values of α is $f(x)$ differentiable at $x = 0$?

2. Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions. Let $y_0 = g(x_0)$ for some $x_0 \in \mathbb{R}$. Find examples for the following cases when:

   (a) g is differentiable at x_0 and f is not differentiable at y_0;

   (b) g is not differentiable at x_0 and f is differentiable at y_0;

   (c) g is not differentiable at x_0 and f is not differentiable at y_0,

   but $f \circ g(x)$ is differentiable.

3. (Exercise 11 on page 186, Pugh) Assume that $f : (-1, 1) \to \mathbb{R}$ and $f'(0)$ exists. If $\alpha_n, \beta_n \to 0$ as $n \to \infty$, define the different quotient

$$D_n = \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n}.$$

   (a) Prove that $\lim_{n\to\infty} D_n = f'(0)$ under each of the following conditions (Hint: First rewrite this expression in terms of $\frac{f(\beta_n)-f(0)}{\beta_n}$ and $\frac{f(\alpha_n)-f(0)}{\alpha_n}$ and use the sequential definition of the limit.

      i. $\alpha_n < 0 < \beta_n$.

      ii. $0 < \alpha_n < \beta_n$ and $\frac{\beta_n}{\beta_n - \alpha_n} \leq M$.

      iii. $f'(x)$ exists and is continuous for all $x \in (-1, 1)$.

   (b) Set $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$. Observe that f is differentiable everywhere in $(-1, 1)$ and $f'(0) = 0$. Find α_n and β_n that tend to 0 in such a way that D_n converges to a limit unequal to $f'(0)$.

Chapter 10

Mean Value Theorems

10.1 Mean Value Theorems

Theorem 10.1.1 (Rolle's Theorem)**.** *$f : [a, b] \to \mathbb{R}$ continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists $x \in (a, b)$ such that $f'(x) = 0$*

Proof. Define

$$x_1 = \arg\min_{x\in[a,b]} f(x), \quad m = \min_{x\in[a,b]} f(x)$$

$$x_2 = \arg\max_{x\in[a,b]} f(x), \quad M = \max_{x\in[a,b]} f(x)$$

- If $m = M$, f is constant and $f'(x) = 0, \forall x \in [a, b]$
- If $m < M$, at least one of x_1 or x_2 is different from both a and b, given that $f(x_1) < f(x_2)$ and $f(a) = f(b)$. Without loss of generality, assume $x_1 \in (a, b)$. By Theorem 9.4.1, $f'(x_1) = 0$.

□

Theorem 10.1.2 (Cauchy's Mean Value Theorem)**.** *Suppose $f, g : [a, b] \to \mathbb{R}$ are continuous and differentiable on (a, b). There exists $x_0 \in (a, b)$ such that*

$$f'(x_0)(g(b) - g(a)) = g'(x_0)(f(b) - f(a))$$

Proof. Define $h(t) := f(t)(g(b) - g(a)) - g(t)(f(b) - f(a))$. h is continuous on $[a, b]$, differentiable on (a, b) and $h(a) = h(b)$. By Rolle's Theorem (Theorem 10.1.1), there exists an $x_0 \in (a, b)$ such that $h'(x_0) = 0$. This happens if, and only if,

$$f'(x_0)(g(b) - g(a)) = g'(x_0)(f(b) - f(a))$$

□

Theorem 10.1.3 (Mean Value Theorem)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is continuous and differentiable on (a, b). There exists $x_0 \in (a, b)$ such that*

$$f(b) - f(a) = f'(x_0)(b - a)$$

Proof. Set $g(x) = x$ in Cauchy's Mean Value Theorem (Theorem 10.1.2). □

Corollary 10.1.1. *Let $f : [a, b] \to \mathbb{R}$ is continuous, differentiable on (a, b) and*

$$\sup_{x \in (a,b)} |f'(x)| \leq M$$

Then,

$$|f(x) - f(x')| \leq M|x - x'|, \quad x, x' \in [a, b]$$

Proof. Let $x, x' \in [a, b], x < x'$. By Mean Value Theorem (Theorem 10.1.3) there exists $\zeta \in (x, x')$ such that $f(x) - f(x') = f'(\zeta)(x - x')$, and hence

$$|f(x) - f(x')| = |f'(\zeta)(x - x')| = |f'(\zeta)| \cdot |x - x'| \leq M|x - x'|.$$

□

10.2 L'Hospital's Rule

Theorem 10.2.1 (L'Hospital's Rule)**.** *Suppose f and g are differentiable on (a, b), $g'(x) \neq 0, \forall x \in (a, b)$, where $-\infty \leq a \leq b \leq \infty$. Suppose:*

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = A, \quad -\infty \leq A \leq \infty$$

If either:

1. $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$, *or*
2. $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = \infty$

Then, $\lim_{x \to a} \frac{f(x)}{g(x)} = A$

10.3 Derivatives of Monotone Functions

Definition 10.3.1. Let $f : I \to \mathbb{R}$. If for all $x_1, x_2 \in I$ s.t. $x_1 < x_2$,

1. $f(x_1) \leq (\geq) f(x_2)$, we say that f is monotonically increasing (decreasing).
2. $f(x_1) < (>) f(x_2)$, we say that f is strictly monotonically increasing (decreasing).

The next theorem characterizes monotonic functions in terms of their derivatives:

Theorem 10.3.1. *Let $f : [a, b] \to \mathbb{R}$ continuous and differentiable on (a, b).*

1. *f is increasing on (a, b) $\iff$ $f'(x) \geq 0, \forall x \in (a, b)$*
2. *f is decreasing on (a, b) $\iff$ $f'(x) \leq 0, \forall x \in (a, b)$*
3. *f is strictly increasing on (a, b) if $f'(x) > 0, \forall x \in (a, b)$*
4. *f is strictly decreasing on (a, b) if $f'(x) < 0, \forall x \in (a, b)$*

Proof. 1. $\Rightarrow$: f is increasing $\Rightarrow$ for all $x < x'$, $\frac{f(x')-f(x)}{x'-x} \geq 0$. Taking limits:

$$f'(x) = \lim_{x' \downarrow x} \frac{f(x') - f(x)}{x' - x} \geq 0$$

$\Leftarrow$: $f'(x) \geq 0$ for all $x \in (a, b)$. Let $x_1 < x_2$. By the Mean Value Theorem, there exists $\zeta \in (x_1, x_2)$ such that:

$$f(x_2) - f(x_1) = f'(\zeta)(x_2 - x_1) \geq 0$$

Then, $f(x_2) \geq f(x_1)$.

2. Analogous to 1.
3. $f'(x) > 0$ for all $x \in (a, b)$. Let $x_1 < x_2$. By the Mean Value Theorem, there exists $\zeta \in (x_1, x_2)$ such that:

$$f(x_2) - f(x_1) = f'(\zeta)(x_2 - x_1) > 0$$

Then, $f(x_2) > f(x_1)$.

4. Analogous to 3.

□

Note that 3. and 4. go only in one direction: if the derivative is strictly positive (negative), the function is strictly increasing (decreasing). However, a function that is strictly increasing (decreasing) does not necessarily have strictly positive (negative) derivative at every point in the domain. An example of such a function is $f(x) = x^3$. In this case, f is strictly increasing, although $f'(0) = 0$.

10.4 Inverse Function Theorem

Theorem 10.4.1 (Inverse Function Theorem)**.** *Let $f : (a, b) \to (c, d)$ be surjective, continuous and differentiable on (a, b), and $f'(x) \neq 0, \forall x \in (a, b)$. Then f is a homeomorphism and its inverse f^{-1} is differentiable, with:*

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

Proof. If $f'(x) \neq 0, \forall x \in (a, b)$, by the Intermediate Value Theorem for Derivatives, $f'(x)$ is either positive for all $x \in (a, b)$, or negative. Assume, without loss of generality, that $f'(x) > 0, \forall x \in (a, b)$.

Let $a < x_1 < x_2 < b$. By the Mean Value Theorem, there exists $\zeta \in (x_1, x_2)$ such that:

$$f(x_2) - f(x_1) = f'(\zeta)(x_2 - x_1) > 0$$

Then, f is strictly monotonically increasing, so it is injective. Since, by assumption, it is also surjective, its inverse f^{-1} exists and is well defined. Moreover, since f is differentiable, it is continuous on (a, b).

Now, lets prove that a strictly monotonic and continuous function is a homeomorphism. Let $y_0 \in (c, d)$ and $\epsilon > 0$. Denote $x_0 = f^{-1}(y_0)$ and define $y^- = f(x_0 - \epsilon)$ and $y^+ = f(x_0 + \epsilon)$. Let $\delta = \min\{|y^+ - y_0|, |y^- - y_0|\}$.

Since f is monotonic, f^{-1} is also monotonic, so $f^{-1}(y_0 + \delta) \leq x_0 + \epsilon$, $f^{-1}(y_0 - \delta) \geq x_0 - \epsilon$ and $f^{-1}(y_0 - \delta, y_0 + \delta)$ is an interval. Moreover, f is continuous, so $f^{-1}(y_0 - \delta, y_0 + \delta)$ is an open set, which means that $f^{-1}(y_0 - \delta, y_0 + \delta) \subseteq (x_0 - \epsilon, x_0 + \epsilon)$, so f^{-1} is continuous and f is a homeomorphism.

Now, lets show that:

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

Let $x_0 = f^{-1}(y_0), \quad x = f^{-1}(y)$.

$$\begin{aligned}
(f^{-1})'(y_0) &= \lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} && (10.1)\\
&= \lim_{x \to x_0} \frac{x - x_0}{f(x) - f(x_0)} && (10.2)\\
&= \frac{1}{f'(x_0)} && (10.3)\\
&= \frac{1}{f'(f^{-1}(y_0))} && (10.4)
\end{aligned}$$

The second equality is true because f^{-1} is continuous, which implies that $y \to y_0$ if and only if $x \to x_0$. □

Example 8. *Let* $y = sin(x), \quad x \in (-\pi/2, \pi/2)$. *Find* $(f^{-1})'(y)$.

Proof. $f^{-1}(y) = arcsin(y)$. Then, by the Inverse Function Theorem:

$$(f^{-1})'(y) = \frac{1}{cos(arcsin(y))} = \frac{1}{\sqrt{1 - sin^2(arcsin(y))}} = \frac{1}{\sqrt{1 - y^2}}$$

□

10.5 Application: Auctions

In this section we present an example that applies the Chain Rule and the Inverse Function Theorem. We present the economic context first to establish where the problem arises. We then present a purely mathematical formulation which simplifies some features of the actual problem so that the application of the theorems is more transparent.

10.5.1 Economic Context

A single object is traded at an auction with two potential bidders, $i \in \{1, 2\}$. Each individual has a valuation for the object, $v_i \in [0, 1]$, which represents the maximum amount they are willing to pay for the object. Valuations are **private**, which means that bidder i does not know the valuation of bidder j. They are also **independent**. which means that bidder i cannot infer any additional information about v_j, based on his own realization. Both players know the probability distribution of (v_i, v_j). At the time of the auction, each individual must decide an amount b_i to bid. The rules for allocating the object follow a **first price auction**. Whoever bids the highest amount pays b_i and receives the object. Losers do not have to pay and do not receive the object. The optimal bid depends on $i's$ beliefs about the other player's strategy.

We will state this problem in a mathematical form so that we can apply some of the differentiation techniques in this chapter. The function we want to optimize is:

$$U(b, v) = (v - b)\mathbb{P}(b > \sigma(v_j))$$

A bidder's expected utility is the net amount that a bidder receives if she wins the auction $(v - b)$ times the probability of winning. The utility function captures the main trade-off in auctions: If you bid higher then you have a higher probability of winning the object but you also receive a lower net payoff in the event that you win.

The probability depends on the strategy of the other player, $\sigma(v_j)$, which is a function of her valuation. Let F be the cumulative density function of v_j, which is assumed to be strictly increasing (the variable is continuous).

10.5.2 Mathematical Formulation

After some manipulations the problem can be rewritten as:

$$U(b) = (v - b)F(\sigma^{-1}(b))$$

where v is a constant. Define the supremum of the function as

$$U^* = \sup_{b \in \mathbb{R}_+} U(b)$$

We formulate assumptions on the functions, sometimes referred to as regularity conditions, which ensures that some of the objects that we are analyzing have desirable properties.

Assumption 10.5.1. *(Regularity Conditions)*

(a) *The function $F : \mathbb{R}_+ \to \mathbb{R}_+$ is twice differentiable and $F' > 0$.*

(b) *The function $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ is surjective, continuous on $\mathbb{R}_+$, differentiable on $\mathbb{R}_{++}$ and its derivative is strictly positive.*

(c) $F(\sigma^{-1}(0)) = 0$.

Let us analyze what these regularity conditions. First, the positive derivative ensures that F, σ are strictly increasing and therefore $F^{-1}(\sigma^{-1}(b))$ is strictly increasing. This ensures that there is a tradeoff: Bidding more increases the value of $F(\sigma^{-1}(b))$ and decreases $(v - b)$. Such an economic trade-off guarantees that the problem has an interior solution. The second derivatives are required so that we can apply our theorems. The last condition states the probability of winning the auction if you bid zero, is also zero.

10.5.3 Existence Interior Solution

First we show that the problem has an interior solution when $v > 0$.

Lemma 10.5.1. *(Existence Solution) Suppose that 10.5.1 holds. For each valuation $v \in [0,1]$, there exists an optimal bid $0 \leq b^*(v) \leq v$ such that $U(b^*) = U^*$. Furthermore, if the valuation is zero then the optimal bid is zero, $b^*(0) = 0$. If the valuation is positive, $v > 0$, then the bid is also positive $b^*(v) > 0$.*

Proof. We break down the proof into multiple parts. Define $h(b) := F(\sigma^{-1}(b))$.

(a) **The function $h(b)$ is strictly increasing**.

By Lemma 10.3.1, $F' > 0$ and $\sigma' > 0$ implies that F, σ are increasing. By the Inverse Function Theorem, the inverse of σ exists and

$$(\sigma^{-1})'(b) = \frac{1}{\sigma'(\sigma^{-1}(b)}$$

Since $\sigma' > 0$ for every point on its domain and it is surjective, then $(\sigma^{-1})'(b) > 0$ for all $b \in \mathbb{R}_+$. Since the composite of two strictly increasing functions is also increasing, $F(\sigma^{-1}(b))$ is strictly increasing in b.

(b) **Choices $b > v$ are suboptimal**.

Suppose that we set $b = v$, then $U(b) = 0$. For $b > v \geq 0$, $F(\sigma^{-1}(b)) > F(\sigma^{-1}(0)) = 0$ because the function is strictly increasing. Furthermore, for $b > v$, the term $(b - v)$ is strictly negative by construction. Therefore $U(b) < 0$ for all $b > v$. That means that it cannot be optimal to choose any $b > v$.

(c) **Existence optimal $0 \leq b^*(v) \leq v$ for all $v \geq 0$**

That means that without loss of generality we can restrict attention to the interval $[0, v]$, which is a compact set. We can also show that $U(b)$ is continuous because it is the composite of continuous functions. By the extreme value theorem, there exists a $b^* \in [0, v]$ such that $U(b^*) = U^*$.

(d) **Special Cases** If $v = 0$, then $b^*(0) \in [0, 0]$. Therefore $b^*(0) = 0$. Now we consider the case where $v > 0$. If $b = 0$ or $b = v$, then $U(b) = 0$. If $\tilde{b} = v/2$ then $h(\tilde{b}) > 0$ and $b - v > 0$. This implies that $U(\tilde{b}) > 0$. This does not imply that $\tilde{b}$ is optimal, but rather that the corner solutions are suboptimal. That means that $b^*(v) \in (0, v)$.

□

10.5.4 First Order Conditions

Lemma 10.5.2. *Suppose that 10.5.1 holds, then $b^*(v)$ solves the following first order conditions.*

$$(v-b)F'(\sigma^{-1}(b))\frac{\partial}{\partial b}\sigma^{-1}(b) + (-b)F(\sigma^{-1}(b)) = 0$$

Proof. In the previous lemma we showed that $b^*(v)$ is an interior maximum. Therefore, we can use the first order conditions to identify it. Let $U(b) = (v-b)h(b)$ where $h(b) := F(\sigma^{-1}(b))$. Then using the product rule:

$$U'(b) = (v-b)h'(b) - h(b) = 0$$

Using the chain rule:

$$(v-b)F'(\sigma^{-1}(b))(\sigma^{-1})'(b) - h(b) = 0$$

Using the inverse function theorem and plugging in the definition of $h(b)$:

$$(v-b)F'(\sigma^{-1}(b))\frac{1}{\sigma'(\sigma^{-1}(b))} - F(\sigma^{-1}(b)) = 0 \tag{10.5}$$

□

10.6 Exercises

1. In the auctions example.

   (a) Assume in addition that $\sigma(v)$ is a function such that $\forall v \in [0,1], b^*(v) = \sigma(v)$ (there is a symmetric equilibrium). Use Equation 10.5 to show that:

   $$\sigma(v) = v - \sigma'(v)\frac{F(v)}{F'(v)}$$

   The right hand side is called the virtual value.

   (b) Using the above equation and the signs of the derivatives, show that if $\forall v \in [0,1], b^*(v) = \sigma(v)$ then $\forall v \in [0,1], \sigma(v) \leq v$ (this show that in a symmetric equilibrium everyone bids weakly below their valuation).

2. Assume f function is continuous on $[0,\infty)$ and differentiable on $(0,\infty)$. Suppose $f(0) = 0$ and f' is increasing on $(0,\infty)$. Prove

   $$g(x) = \frac{f(x)}{x}$$

   is increasing on $(0,\infty)$.

Chapter 11

Taylor Expansion

In this section we will explore the properties of higher derivatives. We will prove Taylor's theorem which concerns the fit of polynomial approximations of a function around a certain point. The theorem has wide applicability in economics, with two main uses. First, it is used for its desirable approximation properties. In **econometrics** it used as a tool to deal with non-linear criterion functions and derive the asymptotic distributions of estimators. This and other types of approximations are used in **macroeconomics** to compute numerical solutions to macro-models. Second, it is used as a tool to analyze the signs of the derivatives. At the end of the chapter we give an example where Taylor's theorem can be used to characterize risk averse consumers. Taylor's theorem is a powerful tool that requires differentiability of the function up to a certain order.

This chapter is organized as follows. We start of with a definition of higher-order derivatives. In the remainder of the chapter we prove Taylor's theorem by breaking down each of its components. First, we analyze the properties of polynomial approximations, which are proven primarily with algebraic manipulations. Second, we prove a recursive version of the mean value theorem. Finally, we prove the main statement of Taylor's theorem incorporating the previous steps.

Definition 11.0.1. Let $f : (a, b) \to \mathbb{R}$ be a real-valued function. Let $x \in (a, b)$ and define $f^{(0)}(x) = f(x)$. Suppose that $f^{(m)} : (a, b) \to \mathbb{R}$ exists. We say that $f^{(m)}$ is differentiable at $x_0 \in (a, b)$ if there exists a finite $L \in \mathbb{R}$ such that:

$$\lim_{x \to x_0} \left| \frac{f^{(m)}(x) - f^{(m)}(x_0)}{x - x_0} - L \right| = 0$$

We define $f^{(m+1)}(x_0) := L$ as the $(m + 1)^{th}$ order derivative of f evaluated at x_0. If $f^{(m)}$ is differentiable at every $x \in (a, b)$, we say that f is $(m + 1)-$order differentiable.

11.1 Polynomial Approximation

Definition 11.1.1. Suppose that $f : (a, b) \to \mathbb{R}$ is $M-$order differentiable. Define the $M-$order Taylor approximation at $x \in (a, b)$:

$$P(h) := \sum_{m=0}^{M} \frac{f^{(m)}(x)h^m}{m!} = f(x) + f^{(1)}(x)h + \ldots + \frac{f^{(M)}(x)h^M}{M!}$$

We present the following illustration of the Taylor approximation and its derivatives with respect to h for a 2-order differentiable function.

$$\begin{aligned} P(h) &= f(x) + f^{(1)}(x)h + \frac{1}{2}f^{(2)}(x)h^2 \\ P^{(1)}(h) &= f^{(1)}(x) + f^{(2)}(x)h \\ P^{(2)}(h) &= f^{(2)}(x) \end{aligned}$$

This leads to a set of interesting properties. For example, $P^{(s)}(0) = f^{(s)}(x)$ for $s \in \{0, \ldots, M\}$. Notice that the terms $f^{(s)}(x)$ are fixed coefficients, the only thing that varies is h. Some of the terms vanish with higher s because the derivative of a constant term is zero. We formalize these results for arbitrary M.

Lemma 11.1.1. *Suppose that $f : (a, b) \to \mathbb{R}$ is $M-$order differentiable and suppose that $P(h)$ is the Taylor approximation at $x \in (a, b)$. Define $R(h) := f(x + h) - P(h)$. Then $P(h), R(h)$ are M-order differentiable functions and for $0 \leq s \leq M$.*

1. $P^{(s)}(h) = \sum_{m=s}^{M} \frac{f^{(m)}(x)h^{m-s}}{(m-s)!}$.
2. $R^{(s)}(h) = f^{(s)}(x + h) - P^{(s)}(h)$.

Proof. We will prove each part of the theorem separately.

1. We show the first part by induction. If $s = 0$ then the result follows by the definition of $P(h)$. Now suppose that it holds for some $0 \leq s \leq M - 1$. That means that $P^{(s)}(h) = \sum_{m=s}^{M} \frac{f^{(m)}(x)h^{m-s}}{(m-s)!}$. This is a polynomial in h, which is differentiable. The terms $f^{(m)}(x)$ are fixed coefficients that do not depend on h. We will show that it holds for $s + 1$.

   Since $s \leq M - 1$, we can decompose it as

$$P^{(s)}(h) = f^{(s)}(x) + \sum_{m=(s+1)}^{M} \frac{f^{(m)}(x)h^{m-s}}{(m-s)!}$$

The derivative of the first term is zero because it does not depend on h. For $(m-s) \geq 1$ we can differentiate each power term $(m-s)h^{m-s}$ as the derivative $(m-s)h^{m-s-1}$. Then the derivative exists and is equal to

$$P^{(s+1)}(h) = \sum_{m=(s+1)}^{M} \frac{f^{(m)}(x)h^{m-(s+1)}}{(m-(s+1))!}$$

2. We need to differentiate the term $f(x+h)$ with respect to h. Notice that x is fixed in this formulation. The chain rule is a good tool to address this issue.

   Define $g(h) := x+h$ and $w(h) := f(x+h) = f(g(h))$. For $s=0$, $w^{(0)} = w(h) = f^0(x+h)$. Suppose that it holds for $0 \leq s \leq M-1$. Then $w^{(s)} = f^{(s)}(x+h)$. The derivative is $g^{(1)}(h) = 1$. Since $s \leq M-1$, the derivative of $f^{(s)}(y)$ exists by assumption of the theorem. Given that both derivatives exist, we can use the chain rule in Theorem 9.6.1 to show that $w^{s+1}(h) = f^{(s+1)}(g(h))g^{(1)}(h) = f^{(s+1)}(x+h)$.

   We can then combine the two results to show that

$$R^{(s)}(h) = w^{(s)}(h) - P(h) = f^{(s)}(x+h) - P^{(s)}(h)$$

□

Corollary 11.1.1 (Properties Taylor Residual)**.** *Suppose that the assumptions of Lemma 11.1.1 hold, then*

(a) *For all* $1 \leq s \leq M$, $\quad R^{(s)}(0) = 0$.

(b) $R^{(M-1)}(h) = f^{(M-1)}(x+h) - f^{(M-1)}(x) - f^M(x)h$.

Proof. (a) By Lemma 11.1.1, $P^{(s)}(h) = \sum_{m=s}^{M} \frac{f^{(m)}(x)h^{m-s}}{(m-s)!}$. If $m=s$, then $h^0 = 1$. Otherwise, if $m > s$, then $h^{m-s} = 0$ evaluated at $h=0$. Therefore, since $0! = 1$,

$$P^{(s)}(0) = \frac{f^s(x)}{0!} = f^s(x)$$

We can use the second part of the lemma to show that $R^{(s)}(0) = f^{(s)}(x+0) - P^{(s)}(0) = f^{(s)}(x) - f^{(s)}(x) = 0$.

(b) $P^{(M-1)}(h) = \sum_{m=M-1}^{M} \frac{f^{(m)}(x)h^{m-M}}{(m-M)!}$, which is equal to $f^{M-1}(x)h^0/0! + f^M(x)h^1/1! = f^{M-1}(x) + f^M(x)h$. Therefore the M^{th} derivative of the residual is $R^{(M-1)}(h) = f^{(M-1)}(x+h) - f^{M-1}(x) - f^M(x)h$ using the second part of the Lemma. □

11.2 Recursive Mean Value Theorem

Lemma 11.2.1 (Recursive Mean-Value-Theorem)**.** *Suppose that $f : (a, b) \to \mathbb{R}$ is $M-$order differentiable and suppose that $P(h)$ is the Taylor approximation at $x \in (a, b)$. Define $R(h) := f(x + h) - P(h)$ and $\theta_0 := h$. Suppose that $M \geq 1$, then*

$$R(h) = R^{(s)}(\theta_s) \prod_{m=0}^{s-1} \theta_m \tag{11.1}$$

where $\theta_m \in (0, \theta_{m-1})$ for all $1 \leq m \leq s$ and $s \in \{1, \ldots, M\}$.

Proof. Assume WLOG that $h > 0$. By Lemma 11.1.1, $R(h)$ is $M-$order differentiable. We will prove this by induction.

(i) For $s = 1$. by Corollary Properties Taylor Residual (Corollary 11.1.1), $R(0) = 0$. Therefore, by the Mean Value Theorem, there exists $\theta_1 \in (0, h)$ such that,

$$R(h) = R(h) - R(0) = R^{(1)}(\theta_1)h$$

Setting $\theta_0 = h$ completes the definition. If $M = 1$, then we are done, otherwise, we can continue.

(ii) Suppose that Equation 11.1 holds for some $s \in \{1, \ldots, M - 1\}$. We will show that it holds for $s + 1$. By Properties Taylor Residual (Corollary 11.1.1), $R^{(s)}(0) = 0$. Furthermore, $R^{(s)}(h)$ is differentiable because $s \leq M - 1$. Therefore, by the Mean Value Theorem, there exists $\theta_{s+1} \in (0, \theta_s)$ such that

$$R^s(\theta_s) = R^{(s)}(\theta_s) - R^{(s)}(0) = R^{(s+1)}(\theta_{s+1})\theta_s$$

Substituting this into Equation 11.1, we get the equation:

$$\begin{aligned}
R(h) &= R^{(s)}(\theta_s) \prod_{m=0}^{s-1} \theta_m && \text{Assumption Induction Step} \\
&= R^{(s+1)}(\theta_{s+1})\theta_s \prod_{m=0}^{s-1} \theta_m && \text{By the MVT} \\
&= R^{(s+1)}(\theta_{s+1}) \prod_{m=0}^{s} \theta_m && \text{Grouping Terms}
\end{aligned}$$

□

11.3 Taylor Theorem

Theorem 11.3.1 (Taylor's Theorem)**.** *Let $f : (a, b) \to \mathbb{R}$ be M-order differentiable and let $P(h)$ be the associated Taylor polynomial evaluated at $x \in (a, b)$. Assume $M \geq 1$. Define $R(h) = f(x + h) - P(h)$. Then:*

(i) $\lim_{h\to 0} \frac{R(h)}{h^M} = 0$

(ii) *$P(h)$ is the only polynomial of degree lower than or equal to M with Property (i).*

(iii) *If, in addition, f is $(M + 1)$-th order differentiable, there exists $\zeta \in (x, x + h)$ such that:*

$$f(x + h) = P(h) + \frac{f^{(M+1)}(\zeta)h^{M+1}}{(M + 1)!}$$

We will break down the proofs into three parts.

11.3.1 Rate of Convergence

Proof of Taylor Theorem (i). WLOG assume that $h > 0$. We consider two exhaustive cases:

(a) If $M = 1$, then $R(h) = f(x+h) - f(x) - f^{(1)}(x)h$. Then by definition of differentiability:

$$\lim_{h\to 0} \frac{|R(h)|}{h} = \lim_{h\to 0} \left| \frac{f(x+h) - f(x)}{h} - f^{(1)}(x) \right| = 0$$

(b) If $M \geq 2$, then proceed with the following proof. In Lemma 11.2.1, set $s = M-1$. Then $R(h) = R^{(M-1)}(\theta_{M-1}) \prod_{m=0}^{M-2} \theta_m$ with the property that $0 < \theta_m < \theta_{m-1}$ and $\theta_0 = h$. We can show that $\theta_{M-1} < \theta_{M-2} < \ldots < \theta_0 = h$. Taking the absolute value on both sides of the equation,

$$\begin{aligned}
|R(h)| &= \left| R^{(M-1)}(\theta_{M-1}) \prod_{m=0}^{M-2} \theta_m \right| && \text{Absolute Value of Equation 11.1.} \\
&= |R^{(M-1)}(\theta_{M-1})| \prod_{m=0}^{M-2} |\theta_m| && \text{Distributing Absolute Value} \\
&\leq |R^{(M-1)}(\theta_{M-1})| h^{M-1} && \text{Because } \theta_m < h, \text{ for all } m \in \{1, \ldots, M-1\}
\end{aligned}$$

On the other hand, by Properties Taylor Residual (b) $R^{M-1}(\theta_{M-1}) = f^{M-1}(x+\theta_{M-1}) - f^{M-1}(x) - f^M(x)\theta_{M-1}$. We can reformulate the inequality as follows,

$$\begin{aligned}
\frac{|R(h)|}{h^M} &\leq \frac{|R^{(M-1)}(\theta_{M-1})| h^{M-1}}{h^M} && \text{Dividing Inequality by } h^M \\
&= \frac{|R^{(M-1)}(\theta_{M-1})|}{h} && \text{Cancelling out terms.} \\
&\leq \frac{|R^{(M-1)}(\theta_{M-1})|}{\theta_{M-1}} && \text{Since } \theta_{M-1} < h. \\
&= \frac{|f^{M-1}(x+\theta_{M-1}) - f^{M-1}(x) - f^M(x)\theta_{M-1}|}{\theta_{M-1}} && \text{By Properties Taylor Residual.}
\end{aligned}$$

To complete this part of the proof we apply the definition of M-order differentiability. Define

$$u(\theta_{M-1}) := \frac{|f^{M-1}(x+\theta_{M-1}) - f^{M-1}(x) - f^M(x)\theta_{M-1}|}{\theta_{M-1}}$$

Since f is M-order differentiable, then by Definition 11.0.1,

$$\lim_{\theta_{M-1}\to 0} u(\theta_{M-1}) = 0$$

The proof is almost complete, but we need to take the limits with respect to h not θ_{M-1} so we have to do some technical manipulations so that we can exchange the limit.

Since $\lim_{\theta_{M-1}\to 0} u(\theta_{M-1})$ exists, then by the sequential definition of a limit, which we stated in 9.1.1, $\lim_{n\to\infty} u(\theta_{M-1,n}) = 0$ for every sequence s.t. $\theta_{M-1,n} \to 0$.

Let $\{h_n\}$ be an arbitrary sequence such that $h_n \to 0$. For every h_n choose the corresponding value $\theta_{M-1,n}$ found in the Recursive Mean-Value-Theorem, which satisfies the property that $0 < \theta_{M-1,n} < h_n$. Therefore $\theta_{M-1,n}$ converges and therefore

$$\lim_{n\to\infty} u(\theta_{M-1,n}) = 0$$

Since the sequence $\{h_n\}$ was arbitrary, that means that $\lim_{h\to 0} u(\theta_{M-1}) = 0$ and therefore

$$\lim_{h\to 0} \frac{|R(h)|}{h^M} \leq \lim_{h\to 0} u(\theta_{M-1}) = 0$$

□

11.3.2 Uniqueness of the Approximation

Taylor's Theorem Part (ii). Let:

$$P(h) = a_0 + a_1 h + \ldots + a_M h^M$$

$$Q(h) = b_0 + b_1 h + \ldots + b_M h^M$$

where the coefficients in $Q(h)$ are allowed to be zero at this stage. Suppose $P \neq Q$ are two polynomials such that:

$$\lim_{h\to 0} \frac{f(x+h) - P(h)}{h^M} = 0$$

$$\lim_{h\to 0} \frac{f(x+h) - Q(h)}{h^M} = 0$$

Then:

$$\frac{f(x+h) - Q(h)}{h^M} = \frac{f(x+h) - P(h)}{h^M} + \frac{P(h) - Q(h)}{h^M}$$

which means that $\lim_{h\to 0} \frac{P(h)-Q(h)}{h^M} = 0$. If this is the case then the polynomial also converges at slower rates,

$$\lim_{h\to 0} \frac{P(h) - Q(h)}{h^s} = \lim_{h\to 0} \frac{P(h) - Q(h)}{h^M} \lim_{h\to 0} h^{M-s} = 0, \quad s \in \{1, \ldots, M\}$$

Suppose that there exists $0 \leq k \leq M$ such that $a_k \neq b_k$. Let k_0 be the smallest such k. Then for $h \neq 0$,

$$\frac{P(h) - Q(h)}{h^{k_0}} = \frac{\sum_{k=k_0}^{M} (a_k - b_k) h^k}{h^{k_0}} = \sum_{k=k_0}^{M} (a_k - b_k) h^{k-k_0}$$

Therefore,

$$\lim_{h\to 0} \frac{P(h) - Q(h)}{h^{k_0}} = a_k - b_k$$

Since $a_k - b_k \neq 0$, this is a contradiction. Therefore, $a_k = b_k$ for all $k \in \{1, \ldots, r\}$.

□

11.3.3 Form of the Residual

From Properties Taylor Residual (Corollary 11.1.1)

$$R^{M-1}(h) = f^{M-1}(x+h) - f^{M-1}(x) - f^{M}(x)h$$

If $f^{M+1}(y)$ exists, we can differentiate twice with respect to h, we can show that

$$R^{M+1}(h) = f^{M+1}(x+h)$$

From the definition, $R(h) = f(x+h) - P(h)$. Define $g(h) := h^{M+1}$. Then we can show that

$$g^{(m)}(h) = \frac{(M+1)!}{((M+1)-m)!}h^{M+1-m}, \quad m \in \{1, \dots, M+1\}$$

It follows that $g^{(m)}(0) = 0$ for all $m \in \{1, \dots, M\}$. Then we can use the Cauchy's Mean Value Theorem recursively,

$$\begin{aligned}
\frac{R(h)}{g(h)} &= \frac{R(h) - R(0)}{g(h) - g(0)} && \text{Because } R(0) = g(0) = 0. \\
&= \frac{R^{(1)}(\theta_1)}{g^{(1)}(\theta_1)} && \text{(I) By Cauchy's Mean Value Theorem} \\
&= \frac{R^{(1)}(\theta_1) - R^{(1)}(0)}{g^{(1)}(\theta_1)} && \text{(II) } R^{(1)}(0) = 0 \text{ by Properties Taylor Residual} \\
&= \frac{R^{(1)}(\theta_1) - R^{(1)}(0)}{g^{(1)}(\theta_1) - g^{(1)}(0)} && \text{(III) Because } g^{(1)}(0) = 0. \\
&= \dots \\
&= \frac{R^{(M+1)}(\theta_{M+1}) - R^{(M+1)}(0)}{g^{(M+1)}(\theta_{M+1}) - g^{(M+1)}(0)} \\
&= \frac{f^{(M+1)}(x+\theta_{M+1})}{(M+1)!} && \text{Plug-in } R^{M+1}(\theta_{M+1}) = f^{(M+1)}(x+\theta_{M+1}).
\end{aligned}$$

We repeat steps $(I) - (III)$ recursively until we obtain the final expression. To complete the proof, multiply both sides by $g(h) = h^{M+1}$.

$$R(h) = \frac{f^{(M+1)}(x+\theta_{M+1})}{(M+1)!}h^{M+1}$$

11.4 Continuous Differentiability

Definition 11.4.1 (Continuous differentiability)**.** Let $f : (a, b) \to \mathbb{R}$ be a $m-$order differentiable function. If $f^{(m)}$ is continuous we say that f is $m-$order continuously differentiable and denote it by $f \in C^m$.

Because of Theorem 9.3.1, if f^m exists then all its lower-order derivatives are continuous. However, not all differentiable functions are continuously differentiable.

Remark It is important to note that this property was not required to prove Taylor's theorem. We only relied on the definition of differentiability.

11.5 Application: Risk Aversion

Taylor's theorem can be very useful analyze problems that have sign restrictions. For example, in decision theory, risk aversion can be characterized using second derivatives. Suppose that a consumer is offered a choice between two assets. One asset pays x with complete certainty. The other pays $x + \epsilon$ with half probability, and $x - \epsilon$ with half probability, where $\epsilon > 0$. In expectation, it pays x. Suppose that $U(x)$ represents a consumer's utility function. Then a consumer is said to be **risk averse** if for all $x, \epsilon \in \mathbb{R}$,

$$U(x) \geq \frac{1}{2}U(x + \epsilon) + \frac{1}{2}U(x - \epsilon) \quad \text{(Risk Aversion)}$$

This captures the idea that a consumer prefers an asset with a certainty rather over a risky asset, even if both give the same return in expectation. Suppose that we assume that the utility function U is twice continuously differentiable. What can we say about the sign of the derivatives?

The following lemma, which is derived using Taylor's theorem, turns out to be very useful.

Lemma 11.5.1. *Let $U : \mathbb{R} \to \mathbb{R}$ be twice differentiable. Then for $x \in (a, b)$,*

$$\lim_{\epsilon \to 0} \frac{U(x + \epsilon) + U(x - \epsilon) - 2U(x)}{\epsilon^2} = U^2(x)$$

Proof. By using the first part of Taylor's Theorem

$$U(x + \epsilon) = U(x) + U^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(x)\epsilon^2 + R_1(\epsilon)$$

$$U(x - \epsilon) = U(x) - U^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(x)\epsilon^2 + R_2(\epsilon)$$

Adding these two expressions together and diving by 2, we get

$$U(x + \epsilon) + U(x - \epsilon) = 2U(x) + U^{(2)}(x)\epsilon^2 + R_1(\epsilon) + R_2(\epsilon)$$

Notice that the terms involving $f^1(x)$ cancel out by construction. If we rearrange the terms we get

$$\frac{U(x + \epsilon) + U(x - \epsilon) - 2U(x)}{\epsilon^2} = U^{(2)}(x) + \frac{R_1(\epsilon)}{\epsilon^2} + \frac{R_2(\epsilon)}{\epsilon^2}$$

By the first part of Taylor's theorem $\lim_{\epsilon\to 0}\frac{R_1(\epsilon)}{\epsilon^2} = \lim_{\epsilon\to 0}\frac{R_2(\epsilon)}{\epsilon^2} = 0$. Therefore,

$$\lim_{\epsilon\to 0}\frac{U(x+\epsilon)+U(x-\epsilon)-2U(x)}{\epsilon^2} = U^2(x)$$

□

This allows us to formulate the following equivalence theorem.

Lemma 11.5.2. *Let U be a twice differentiable utility function. Then a consumer is risk averse if and only if $U^2(x) \leq 0$ for all $x \in (a,b)$.*

Proof. $\implies$ Suppose that a consumer is risk averse, then for all $x, \epsilon \in \mathbb{R}$,

$$U(x) \geq \frac{1}{2}U(x+\epsilon) + \frac{1}{2}U(x-\epsilon) \quad \text{(Risk Aversion)}$$

By rearranging the equation and dividing by $\epsilon^2 > 0$,

$$\frac{U(x+\epsilon)+U(x-\epsilon)-2U(x)}{\epsilon^2} \leq 0$$

By Lemma 11.5.1 $U^{(2)}(x) = \lim_{\epsilon\to 0}\frac{U(x+\epsilon)+U(x-\epsilon)-2U(x)}{\epsilon^2}$. Such a limit exists because U is twice differentiable. We can show that this is equal to zero by taking the limit on both sides of the inequality.

$\impliedby$ Suppose that $U^(2)(x) \leq 0$ for all $x \in \mathbb{R}$. Then we can use the third part of Taylor's theorem:

$$U(x+\epsilon) = U(x) + U^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(\zeta_1)\epsilon^2, \quad \zeta_1 \in (x, x+\epsilon)$$

$$U(x-\epsilon) = U(x) - U^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(\zeta_1)\epsilon^2 \quad \zeta_2 \in (x-\epsilon, x)$$

By averaging the two equations and rearranging the terms we can show that

$$\frac{1}{2}U(x+\epsilon) + \frac{1}{2}U(x-\epsilon) - U(x) = \frac{1}{2}U^{(2)}(\zeta_1)\epsilon^2 + \frac{1}{2}U^{(2)}(\zeta_2)\epsilon^2$$

The right-hand side is less than or equal to zero because the second derivative is non-positive regardless of the choice of ζ_1, ζ_2. Therefore, the consumer is risk averse.

□

The proof is interesting because we use both parts of Taylor's Theorem.

11.6 Properties Appendix: Common Strategies

Polynomial expansions can be manipulated in different ways to highlight different derivatives. Suppose that we have a twice differentiable function evaluated at two points and expanded around x,

$$f(x+\epsilon) = f(x) + f^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(\zeta_1)\epsilon^2, \quad \zeta_1 \in (x, x+\epsilon)$$
$$f(x-\epsilon) = f(x) - f^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(\zeta_2)\epsilon^2 \quad zeta_2 \in (x-\epsilon, x).$$

Additive Strategy: Cancels out first derivative. Useful if we know sign of second derivative.

$$f(x+\epsilon) + f(x-\epsilon) = 2f(x) + \frac{1}{2}(U^{(2)}(\zeta_1) + U^{(2)}(\zeta_2))\epsilon^2$$

Let f be 3-order differentiable function.

$$f(x+\epsilon) = f(x) + f^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(x)\epsilon^2 + \frac{1}{6}U^{(3)}(\zeta_1)\epsilon^3, \quad \zeta_1 \in (x, x+\epsilon)$$
$$f(x-\epsilon) = f(x) - f^{(1)}(x)\epsilon + \frac{1}{2}U^{(2)}(x)\epsilon^2 - \frac{1}{6}U^{(3)}(\zeta_2)\epsilon^3 \quad \zeta_2 \in (x-\epsilon, x).$$

Subtraction Strategy: Cancels out second derivative. Useful if we know properties of first and third derivative. Sometimes the first derivative is zero at a local maximum, which simplifies the equation further.

$$f(x+\epsilon) - f(x-\epsilon) = 2f^{(1)}(x) + \frac{1}{3}(U^{(3)}(\zeta_1) + U^{(3)}(\zeta_2))\epsilon^3.$$

Transforming Problem Let $x_1, x_2 \in \mathbb{R}$. Then we can always express $f(x_1), f(x_2)$ in the above form by choosing $x = \frac{1}{2}(x_1 + x_2)$ and choosing $\epsilon = x_2 - x$.

11.7 Exercises

1. Suppose $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable. Assume $f(0) > 0$, $f'(0) < 0$ and $f''(x) < 0$ for all $x \in \mathbb{R}$. Prove there exists $\xi \in \left(0, -\frac{f(0)}{f'(0)}\right)$ such that $f(\xi) = 0$.

2. Assume $f : [a, b] \to \mathbb{R}$ is twice differentiable and $f'(a) = f'(b) = 0$. Prove there exists $\xi \in (a, b)$ such that
$$\big|f''(\xi)\big| \geq \frac{4}{(b-a)^2}\big|f(b) - f(a)\big|.$$
(Hint: expand $f\left(\frac{a+b}{2}\right)$ at a and b respectively)

3. Let $f : [a, b] \to \mathbb{R}$ be twice differentiable. Assume $\sup_{x\in[a,b]} |f''(x)| \leq M$ for some constant M. Assume also f achieves its global maximum at some point x^* in (a, b). Prove
$$\big|f'(a)\big| + \big|f'(b)\big| \leq M(b - a).$$

Chapter 12

First-Order Differentiation in $\mathbb{R}^n$

12.1 Definition Differentiation

Recall that $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $x \in \mathbb{R}$ if the following limit exists and is finite:

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

We say that the derivative of f at x is:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

This is equivalent to saying that f is differentiable at x, with derivative $f'(x)$, if there exists a function $r : \mathbb{R} \to \mathbb{R}$ such that:

$$f(x+h) - f(x) = f'(x) \cdot h + r(h)$$

And the remainder r is "sublinear":

$$\lim_{h \to 0} \frac{r(h)}{h} = 0$$

Note that, for a given x, the term $f'(x)h$ is linear in h, so we can interpret the derivative $f'(x)$ not as a number, but as a linear operator in $\mathbb{R}$, that maps h to $f'(x)h$. This is a natural way to extend the concept of derivative to $\mathbb{R}^n$:

Definition 12.1.1. Let $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$. The function f is differentiable at $p \in U$, if there exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$f(p+v) - f(p) = T(v) + R(v)$$

and the remainder function R is sublinear:

$$\lim_{v \to 0} \frac{\|R(v)\|}{\|v\|} = 0$$

We say that the derivative (also called total derivative or Fréchet derivative) is $(Df)_p = T$.

This is equivalent to saying that $f : U \to \mathbb{R}^m$ is differentiable at $p \in U$ if there exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{v \to 0} \frac{\|f(p+v) - f(p) - T(v)\|}{\|v\|} = 0$$

Theorem 12.1.1. *If f is differentiable at $p \in U$, then the derivative is uniquely determined by:*

$$(Df)_p(u) = \lim_{t \to 0} \frac{f(p+tu) - f(p)}{t}$$

Proof. Let T be a linear map satisfying $f(p+v) - f(p) = T(v) + R(v)$ and $\lim_{v \to 0} \frac{\|R(v)\|}{\|v\|} = 0$.

$$\begin{aligned} \lim_{t \to 0} \frac{f(p+tu) - f(p)}{t} &= \lim_{t \to 0} \frac{T(tu)}{t} + \frac{R(tu)}{t} && (12.1) \\ &= \lim_{t \to 0} \frac{tT(u)}{t} + \frac{R(tu)}{t} && (12.2) \\ &= T(u) + \lim_{t \to 0} \frac{R(tu)}{t\,\|u\|} \cdot \|u\| && (12.3) \\ & && (12.4) \end{aligned}$$

Given that $\|u\|$ is finite and R is sublinear, the second term vanishes, so:

$$\lim_{t \to 0} \frac{f(p+tu) - f(p)}{t} = T(u)$$

Since limits are unique, if there are two such transformations T and T', they must be equal to each other: $T = T'$. □

12.2 Continuity

Now, we state some of the theorems we saw in the univariate case, extended for the multivariate case.

Theorem 12.2.1. *Let $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$. Suppose f is differentiable at p. Then f is continuous at p.*

Proof. $(Df)_p : \mathbb{R}^n \to \mathbb{R}^m$ is a finite linear map, from $\mathbb{R}^n$ to a normed vector space $\mathbb{R}^m$.

$$\begin{aligned} \lim_{v\to 0} \|f(p+v) - f(p)\| &= \lim_{v\to 0} \|(Df)_p(v) + R(v)\| && (12.5)\\ &\leq \lim_{v\to 0} \|(Df)_p\| \cdot \|v\| + \|R(v)\| && (12.6)\\ &= 0 && (12.7) \end{aligned}$$

given that $\|(Df)_p\| < \infty$, $\lim_{v\to 0} \|v\| = 0$ and $\lim_{v\to 0} \|R(v)\| = 0$. □

Theorem 12.2.2. *Let $f, g : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ be differentiable at $p \in U$, $\alpha \in \mathbb{R}$. Then:*

1. $(D(f + \alpha g))_p = (Df)_p + \alpha(Dg)_p$
2. *If $f(p) = c$, for all $p \in U$, then $(Df)_p = 0$*
3. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping, $f(v) = Av, A \in \mathbb{R}^m \times \mathbb{R}^n$, then A is the Jacobian matrix for all $p \in U$.*
4. *If $h : \mathbb{R}^{2n} \to \mathbb{R}$ is a bilinear form, $h(p) = p_1^t A p_2$, $p = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \in \mathbb{R}^{2n}$, $A \in \mathbb{R}^n \times \mathbb{R}^n$, then $(J)_p = [p_2^t A^t, p_1^t A]$.*

12.2.1 Differentiation of Vector Valued Functions

Theorem 12.2.3. *Let $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$. Then, f is differentiable at $p \in U$ if and only if each of its components f_i is differentiable at p. Furthermore, the derivative of the i-th component is the i-th component of the derivative.*

Proof. $\Rightarrow$: Let f be differentiable and define the projection on the i-th dimension as:

$$\pi_i : \mathbb{R}^n \to \mathbb{R}, \quad \pi_i(w_1, \ldots, w_i, \ldots, w_n) = w_i$$

Clearly, π_i is linear, so it is differentiable. Then, $f_i = \pi_i \circ f$ is differentiable and:

$$(Df_i)_p = (D\pi_i)_{f(p)}(Df)_p$$

Moreover, the projection π_i can be represented by the $1 \times n$ vector that has 1 in the i-th component and 0 elsewhere:

$$A = (0, \ldots, 1, \ldots, 0)$$

Thus we know that $(D\pi_i)_{f(p)}$ is represented by the matrix A. So:

$$(Df_i)_p = A(Df)_p = \pi_i \circ (Df)_p$$

$\Leftarrow$: Suppose each f_i is differentiable, with derivative $(Df_i)_p$. Construct:

$$T = \begin{bmatrix} (Df_1)_p \\ \vdots \\ (Df_m)_p \end{bmatrix}$$

$$\Rightarrow \quad f(p+h) - f(p) - T \cdot h = \begin{bmatrix} f_1(p+h) - f_1(p) - (Df_1)_p \cdot h \\ \vdots \\ f_m(p+h) - f_m(p) - (Df_m)_p \cdot h \end{bmatrix}$$

Taking limits, this converges if and only if each component converges. Therefore, T is indeed the derivative of f. $\square$

This theorem is important, because it shows that what makes calculus in $\mathbb{R}^n$ different from calculus in $\mathbb{R}$ is the multidimensionality of the domain, and not of the range.

12.3 Special Theorems

12.3.1 Chain Rule

Theorem 12.3.1 (Chain Rule)**.** *Let $U \subseteq \mathbb{R}^n$ and $W \subseteq \mathbb{R}^m$ be open sets. Let $f : U \to \mathbb{R}^m$ be differentiable at $p \in U$ and $f(U) \subseteq W$. Let $g : W \to \mathbb{R}^l$ be differentiable at $f(p) \in W$. Define $h = g \circ f$. Then h is differentiable at $p \in U$ and $(Dh)_p = (Dg)_{f(p)} \cdot (Df)_p$*

Proof.

$$\begin{aligned} f(p+v) - f(p) &= (Df)_p(v) + R(v) \\ g(f(p)+u) - g(f(p)) &= (Dg)_{f(p)}(u) + S(u) \\ g(f(p+v)) &= g(f(p) + (Df)_p(v) + R(v)) \\ &= g(f(p)) + (Dg)_{f(p)}((Df)_p(v) + R(v)) + S((Df)_p(v) + R(v)) \end{aligned}$$

Therefore,

$$\begin{aligned} g(f(p+v)) - g(f(p)) &= (Dg)_{f(p)}((Df)_p(v) + R(v)) + S((Df)_p(v) + R(v)) \\ &= (Dg)_{f(p)}(Df)_p(v) + (Dg)_{f(p)}R(v) + S((Df)_p(v) + R(v)) \end{aligned}$$

It now suffices to show that the last two terms are sublinear:

1. $(Dg)_{f(p)}R(v)$:

$$\lim_{v\to 0} \frac{\left\|(Dg)_{f(p)}R(v)\right\|}{\|v\|} \leq \lim_{v\to 0} \left\|(Dg)_{f(p)}\right\| \cdot \frac{\|R(v)\|}{\|v\|} = 0$$

   as the first term is finite anr R is sublinear.

2. $S((Df)_p(v) + R(v))$:

$$\lim_{v\to 0} \frac{\|S((Df)_p(v) + R(v))\|}{\|v\|} = \lim_{v\to 0} \frac{\|S((Df)_p(v) + R(v))\|}{\|(Df)_p(v) + R(v)\|} \cdot \frac{\|(Df)_p(v) + R(v)\|}{\|v\|}$$

   The limit when $v \to 0$ of the last term is finite:

$$\frac{\|(Df)_p(v) + R(v)\|}{\|v\|} \leq \frac{\|(Df)_p(v)\|}{\|v\|} + \frac{\|R(v)\|}{\|v\|} \leq \frac{\|(Df)_p\|\,\|v\|}{\|v\|} + \frac{\|R(v)\|}{\|v\|} = \|(Df)_p\| + \frac{\|R(v)\|}{\|v\|}$$

□

12.3.2 Mean-Value Theorem

Theorem 12.3.2 (Mean Value Theorem)**.** *Let $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$. Assume f is differentiable on U and the segment $[p, q]$ is contained in U. Then:*

$$|f(q) - f(p)| \leq M\,|q - p|\,, \quad M = \sup_{x \in U}\{\|(Df)_x\|\}$$

Proof. Assume the segment $[p, q]$ is contained in U. The segment can be parameterized as:

$$p + t(q - p), \quad t \in [0.1]$$

Define:

$$g : [0, 1] \to \mathbb{R}, \quad g(t) := (f(p) - f(q))^T \cdot f(p + t(q - p))$$

$$\Rightarrow \quad g'(t) = (f(p) - f(q))^T (Df)_{p+t(q-p)}(q - p)$$

By the Mean Value Theorem in $\mathbb{R}$, there exists $\zeta \in (0, 1)$ such that:

$$g(1) - g(0) = g'(\zeta) = (f(p) - f(q))^T (Df)_{p+\zeta(q-p)}(q - p)$$

$$g(1) - g(0) = (f(p) - f(q))^T \cdot (f(q) - f(p)) = -\|f(p) - f(q)\|^2$$

$$\Rightarrow \quad \|f(p) - f(q)\|^2 = (f(p) - f(q))^T (Df)_{p+\zeta(q-p)}(p - q)$$

By the Cauchy-Schwarz Inequality:

$$\|f(p) - f(q)\| \leq \left\|(Df)_{p+\zeta(q-p)}\right\| \cdot \|p - q\| \leq M\,\|p - q\|$$

□

Corollary 12.3.1. *Assume U is connected. Let $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$ be differentiable and $(Df)_x = 0$. Then f is constant.*

Proof. Let $x \in U$. Define $P(x) := \{y \in U | f(x) = f(y)\}$. Lets show that $P(x)$ is open: Let $y \in P(x)$. Since U is open, there exists an ϵ-neighborhood of y, $O_y \subseteq U$, which is open. Let $z \in O_y$. The segment $[y, z] \subseteq O_y$. Then, $|f(y) - f(z)| \leq M\,|y - z| = 0$. This implies that $f(x) = f(y) = f(z)$ for every $z \in O_y$. Then $z \in P(x)$, which implies $O_y \subseteq P(x)$, so $P(x)$ is open.

Now we show $P(x) = U, \forall x \in U$. Assume $P(x) \neq U$. That is, assume there exists $x \in U, P(x) \neq U$. $P(x)$ and $\cup_{y \notin P(x)} P(y)$ are both open, disjoint and $U = P(x) \cup \left(\cup_{y \notin P(x)} P(y)\right)$. This implies that U is disconnected, which is a contradiction. Therefore, $P(x) = U$. □

12.4 Partial Derivatives

Definition 12.4.1. Let $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$. Define the ij-th partial derivative of f at p as:

$$\frac{\partial f_i(p)}{\partial x_j} = \lim_{t\to 0} \frac{f_i(p + te_j) - f_i(p)}{t}$$

Theorem 12.4.1. *Let $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$ be differentiable. Then, the partial derivatives exist and are the entries of the matrix that represents the total derivative.*

Proof. Recall that the total derivative $(Df)_p$ is a linear map. This means that there exists a matrix of size $m \times n$ that represents $(Df)_p$. Let A be the matrix that represents the derivative $(Df)_p$. Then:

$$(Df)_p(e_j) = Ae_j = \lim_{t\to 0} \frac{f(p + te_j) - f(p)}{t} = \begin{bmatrix} \frac{\partial f_1(p)}{\partial x_j} \\ \vdots \\ \frac{\partial f_m(p)}{\partial x_j} \end{bmatrix}$$

Then:

$$A = \begin{bmatrix} \frac{\partial f_1(p)}{\partial x_1} & \cdots & \frac{\partial f_1(p)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m(p)}{\partial x_1} & \cdots & \frac{\partial f_m(p)}{\partial x_n} \end{bmatrix}$$

□

Note that Theorem 12.4.1 states that if the derivative exists, then the partials also exist. A natural question is whether the converse is true. If the partial derivatives exist, is the function f differentiable? The following example shows that this is not the case.

Example 9. *Let:*

$$f(x) = \begin{cases} 0 & \text{if } x, y = 0 \\ \frac{xy}{x^2+y^2} & \text{otherwise} \end{cases}$$

f is not continuous at $(x, y) = (0, 0)$. To see this, take:

$$(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{n}\right) \xrightarrow{n\to\infty} (0, 0)$$

$$f(x_n, y_n) = \frac{1}{2}, \quad \forall n \geq 1$$

But $f(0, 0) = 0$, so f is not continuous. However, the partials exists. Note, however, that the partials are not continuous.

In the above example, we saw that the existence of the partials is not sufficient for the function to be differentiable. In particular, the partial derivatives of the function in the example existed, but were not continuous. The following theorem states a sufficient condition for f to be differentiable.

Theorem 12.4.2. *Let $f : U \to \mathbb{R}^m, U \subseteq \mathbb{R}^n$. If the partial derivatives of f exist and are continuous then f is differentiable.*

Proof. Assume the partials exist and are continuous. Without loss of generality, assume that $m = 1$ (Theorem 12.2.3). Let $h \in \mathbb{R}^n$.

$$\begin{aligned} f(x+h) - f(x) &= f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n) \\ &= f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, x_2 + h_2, \ldots, x_n + h_n) \\ &+ f(x_1, x_2 + h_2, \ldots, x_n + h_n) - f(x_1, x_2, x_3 + h_3 \ldots, x_n + h_n) \\ &+ f(x_1, x_2, x_3 + h_3, \ldots, x_n + h_n) - f(x_1, x_2, x_3, x_4 + h_4, \ldots, x_n + h_n) \\ &\quad \ldots \\ &+ f(x_1, x_2, \ldots, x_{n-1}, x_n + h_n) - f(x_1, x_2, \ldots, x_n) \end{aligned}$$

We are "moving" component by component on each line. Using the Mean Value Theorem:

$$\begin{aligned} &= \frac{\partial f}{\partial x_1}(\theta_1, x_2 + h_2, \ldots, x_n + h_n)h_1 \\ &+ \frac{\partial f}{\partial x_2}(x_1, \theta_2, x_3 + h_3, \ldots, x_n + h_n)h_2 \\ &+ \ldots \\ &+ \frac{\partial f}{\partial x_n}(x_1, \ldots, x_{n-1}, \theta_n)h_n \end{aligned}$$

where $\theta_1 \in (x_1, x_1 + h_1), \ldots, \theta_n \in (x_n, x_n + h_n)$. Then:

$$\begin{aligned} f(x+h) &- f(x) - A \cdot h \\ &= \left(\frac{\partial f}{\partial x_1}(\theta_1, x_2 + h_2, \ldots, x_n + h_n) - \frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x_1, \ldots, x_{n-1}, \theta_n) - \frac{\partial f}{\partial x_n}(x) \right) \cdot h \\ &= z(h) \cdot h \end{aligned}$$

By Cauchy-Schwarz Inequality:

$$\frac{\|f(x+h) - f(x) - A \cdot h\|}{\|h\|} \leq \frac{\|z(h)\|}{\|h\|} \|h\| = \|z(h)\| \xrightarrow{h \to 0} 0$$

where the last inequality follows because the partials are continuous. Therefore, f is differentiable. $\square$

12.5 Exercises

1. (Euler's Equations) Assume $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable. Fix $(x, y) \in \mathbb{R}^2$. Define $g(t) = f(tx, ty)$ for all $t > 0$. Show g is differentiable and

$$g'(t) = x\frac{\partial f}{\partial x}(tx, ty) + y\frac{\partial f}{\partial y}(tx, ty).$$

Assume in addition, there exists $\alpha > 0$ such that

$$f(tx, ty) = t^\alpha f(x, y) \quad \forall t > 0 \quad \text{and} \quad \forall (x, y) \in \mathrm{R}^2. \tag{12.8}$$

Show for all $(x, y) \in \mathbb{R}^2$,

$$x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) = \alpha f(x, y). \tag{12.9}$$

A function with the property (17.1) is said to be homogeneous of degree α. The equation (17.2) is called Euler's formula.

2. (Exercise 16 on page 347, Pugh) Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f = (x, y, z)$ and $g = w$ where

$$\begin{aligned} w &= w(x, y, z) = xy + yz + zx \\ x &= x(s, t) = st \quad y = y(s, t) = s\cos t \quad z = z(s, t) = s\sin t \end{aligned}$$

   (a) Find the matrices that represent the linear transformations $(Df)_p$ and $(Dg)_q$ where $p = (s_0, t_0) = (0, 1)$ and $q = f(p)$.

   (b) Use the Chain rule to calculate the 1×2 matrix $[\partial w/\partial s, \partial w/\partial t]$ that represents $(D(g \circ f))_p$.

   (c) Plug the functions $x = x(s, t)$, $y = y(s, t)$ and $z = z(s, t)$ directly into $w = w(x, y, z)$ and recalculate $[\partial w/\partial s, \partial w/\partial t]$, verifying the answer given in (b).

Chapter 13

Second-Order Differentiation in $\mathbb{R}^n$

13.1 Bilinear Maps

Definition 13.1.1. Let $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ is a bilinear map if for all $\alpha_i, \beta_j \in \mathbb{R}$ and $u_i, v_j \in \mathbb{R}^n$ such that $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, k'\}$.

$$g\left(\sum_{j=1}^{k'} \beta_j v_j, \sum_{i=1}^{k} \alpha_i u_i\right) = \sum_{i=1}^{k}\sum_{j=1}^{k'} \alpha_i \beta_j g(v_j, u_i)$$

for any positive integer k and k'.

Lemma 13.1.1 (Bilinear Matrix Representation)**.** *Let $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ be a bilinear map. If $m = 1$, then the function g is uniquely represented by an $n \times n$ matrix H, such that $g(x, y) = x^t H y$ where $x, y \in \mathbb{R}^n$.*

Proof. First we show that if $g(x, y) = x^t H y$ then g is a bilinear map. Suppose that $x = \sum_{i=1}^{k} \alpha_i u_i \in \mathbb{R}^n$ and $y = \sum_{j=1}^{k'} \beta_j v_j \in \mathbb{R}^n$. Then,

$$\begin{aligned} x^t H y &= \left[\sum_{i=1}^{k} \alpha_i u_i\right]^t H \left[\sum_{j=1}^{k'} \beta_j v_j\right] && \text{Plugging-in Linear Combinations} \\ &= \left[\sum_{i=1}^{k} \alpha_i u_i^t\right] H \left[\sum_{j=1}^{k'} \beta_j v_j\right] && \text{Distributing Transpose} \end{aligned}$$

We can distribute the sum on either side of h and rearrange the equation to prove the desired

result.

$$
\begin{aligned}
&= \sum_{i=1}^{k} \alpha_i u_i^t H \left[\sum_{j=1}^{k'} \beta_j v_j \right] \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k'} \alpha_i \beta_j u_i^t H v_j \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k'} \alpha_i \beta_j g(u_i, v_j)
\end{aligned}
$$

Second we show that every bilinear map can be represented with an $n \times n$ matrix H. Define the entries $h_{ij} = g(e_i, e_j)$, where e_i, e_j are **elementary basis vectors** (have 1 in a single coordinate and zero otherwise). This is a similar argument to when we proved the unique representation of a linear map in Lemma 1.3.2. Then we can write all vectors in the Euclidean space as linear combinations of the elementary basis vectors. Let $u, v \in \mathbb{R}^n$, then $u = \sum_{i=1}^{n} u_i e_i$ and $v = \sum_{j=1}^{n} u_j e_j$.

□

13.2 Function Spaces

Definition 13.2.1. Suppose that $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$. We will denote this as $g(v)(u)$ where $u, v \in \mathbb{R}^n$. Define $g(v)(\cdot)$ as a function that is constructed by fixing one of the inputs of g.

Lemma 13.2.1. *Suppose that $g : \mathbb{R}^n \times \mathbb{R}^n$ is a bilinear map. Then $g(v)(\cdot)$ is a linear map.*

Proof. Let $\alpha_1, \beta_1 \in \mathbb{R}$ and $u_1, u_2 \in \mathbb{R}$. Since g is a bilinear map we can distribute linear combinations of its second argument.

$$g(v)(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 g(v)(u_1) + \alpha_2 g(v)(u_2)$$

□

We define a metric between two linear maps f, g as

$$d(f, g) = ||f - g||$$

where $||T|| := \sup_{x \in \mathbb{R}^n : ||x|| = 1} ||T(x)||$. Using the operator norm inequality, this implies that $||f(v) - g(v)|| \leq ||f - g||\ ||v||$. We can show that $||T||$ is a well-defined metric over the space of linear maps.

1. $||f - g|| = ||g - f||$ (Symmetry).
2. $||f - g|| \leq ||f - h|| + ||h - g||$ (Triangle Inequality).
3. $||f - g|| \geq 0$ and $||f - g|| = 0$ if and only if $f = g$.

Proof. We prove each iterm

1. Let $f(x) = Ax$, $g(x) = Bx$ and $h(x) = Cx$. Then (i) $||f - g|| = ||g - f||$ because $||Ax - Bx|| = ||Bx - Ax||$ for all $x \in \mathbb{R}^n$.
2. (ii) Follows from the fact that $||Ax - Bx|| \leq ||Ax - Cx|| + ||Cx - Bx||$ (using the triangle inequality for the Euclidean norm) for all $||x|| = 1$. We can take the supremum on both sides to show that $||f - g|| \leq ||f - h|| + ||h - g||$.
3. The norm is always non-negative by construction. Furthermore, if $||f - g|| = 0$ that means that $||(A - B)x|| = 0$ for all $x \in \mathbb{R}^n$ such that $||x|| = 1$. It can be shown that this also holds for all non-zero $x \in \mathbb{R}^n$ by scaling the vector. Then that means that $Ker(A - B) = \mathbb{R}^n$. Therefore $A - B = 0_{m \times n}$ and $A = B$. On the other hand if $f - g = 0_{m \times n}$ then $||f - g|| = 0$.

□

13.3 Second-Order Derivatives

Definition 13.3.1. Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable with total derivative (Df). We say that f is twice differentiable if there exists a bilinear map $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ such that:

$$(Df)_{x+v} - (Df)_x = T(v)(\cdot) + R(v)(\cdot), \quad \lim_{v \to 0} \frac{\|R(v)(\cdot)\|}{\|v\|} = 0$$

We denote the derivative as $(D^2 f)_x := T$ and call it the second derivative of f. The norm used in the numerator is the operator norm. The definition implies that $R(v)(\cdot)$ is linear on its second argument (but not necessarily the first).

The definition of second-order differentiability implies that:

$$(Df)_{x+v}(u) - (Df)_x(u) = T(v)(u) + R(v)(u)$$

Each element in this equation is a vector in $\mathbb{R}^m$. The equation is more easily interpreted if we simplify some features of the problem. Let $m = 1$. Suppose that f is a function that measures the profits of company and that $u \in \mathbb{R}^2$ is a proposed price change in two of its products. The value $(Df)_x$ represents the effects of an average price change at current prices x (call this the "bad times" prices).

However, the marginal effects of the price could differ depending on the current state of prices. Suppose that we evaluate the marginal change at a different level $x + v$ ("good times"), which we denote $(Df)_{x+v}(u)$. The function $T(v, u)$ is a bilinear approximation to these simultaneous changes in prices (change in overall level and marginal changes). Consequently the functions $(Df)_{x+v}$ and $(Df)_x$ represent the effect of all possible price changes at each level.

Remark 1: In the example we described there appear to be an artificial distinction between changes in overall price levels and marginal price changes. If the function is twice differentiable there need not be. In the next section we show that the function $T(v)(u)$ is symmetric.

Remark 2: When f is a real-valued function, that is, when $m = 1$ the representation is much simpler. By Lemma 13.1.1 there exists an $n \times n$ matrix that represents it. We call this the **Hessian** matrix.

13.4 Symmetry

Theorem 13.4.1. *If $(D^2 f)_p$ exists, it is symmetric:*

$$(D^2 f)_p(v)(w) = (D^2 f)_p(w)(v)$$

Proof. Without loss of generality, assume $m = 1$ (as symmetry concerns only the arguments of f, not its values). Let $f : \mathbb{R}^n \to \mathbb{R}$. Fix $v, w \in \mathbb{R}^n$. Let $t \in [0, 1]$ and define $g : [0, 1] \to \mathbb{R}$, where:

$$g(s) = f(p + tv + stw) - f(p + stw)$$

Using the Chain Rule,

$$g'(s) = Df_{p+tv+stw}(tw) - Df_{p+stw}(tw)$$

By the Mean Value Theorem, $g(1) - g(0) = g'(\theta), \theta \in (0, 1)$, therefore

$$g(1) - g(0) = Df_{p+tv+\theta tw}(tw) - Df_{p+\theta tw}(tw) \tag{13.1}$$

On the other hand, by definition of the second order derivative:

$$(Df)_{p+tv+\theta tw} - (Df)_p = (D^2 f)_p(tv + \theta tw)(\cdot) + R(tv + \theta tw)(\cdot) \tag{13.2}$$

$$(Df)_{p+\theta tw} - (Df)_p = (D^2 f)_p(\theta tw)(\cdot) + S(\theta tw)(\cdot) \tag{13.3}$$

Since $(D^2 f)_p$ is bilinear then it is linear in its first argument,

$$(D^2 f)_p(tv + \theta tw)(\cdot) - (D^2 f)_p(\theta tw)(\cdot) = (D^2 f)_p(tv)(\cdot)$$

We can subtract Equation 13.3 from 13.2 to obtain a new equation

$$(Df)_{p+tv+\theta tw} - (Df)_{p+\theta tw} = (D^2 f)_p(tv)(\cdot) + R(tv + \theta tw)(\cdot) - S(\theta tw)(\cdot)$$

We can plug the right hand side into Equation 13.1, evaluated at the vector (tw),

$$g(1) - g(0) = (D^2 f)_p(tv)(tw) + R(tv + \theta tw)(tw) - S(\theta tw)(tw) \tag{13.4}$$

We can divide both sides by t^2,

$$\begin{aligned}\frac{g(1)-g(0)}{t^2} &= \frac{(D^2f)_p(tv)(tw)}{t^2} + \frac{R(tv+\theta tw)(tw)}{t^2} - \frac{S(\theta tw)(tw)}{t^2} \\ &= (D^2f)_p(v)(w) + \frac{R(tv+\theta tw)(tw)}{t^2} - \frac{S(\theta tw)(tw)}{t^2} && \text{Because } (D^2f)_p \text{ bilinear.} \\ &= (D^2f)_p(v)(w) + \frac{R(tv+\theta tw)(w)}{t} - \frac{S(\theta tw)(w)}{t} && \text{By Diff, } R,S \text{ linear in second arg.}\end{aligned}$$

By definition of differentiability, R, S are sublinear in the first argument. Therefore we can take limits on both sides to show that,

$$\frac{g(1)-g(0)}{t^2} = (D^2f)_p(v)(w)$$

To complete the proof we show that $g(1)-g(0)$ is symmetric in the vectors v, w.

$$g(0) = f(p+tv) - f(p)$$

$$g(1) = f(p+tv+tw) - f(p+tw)$$

Combining the two equations

$$g(1)-g(0) = f(p+tv+tw) - f(p+tv) - f(p+tw) + f(p)$$

which is symmetric in v, w. Therefore $(D^2f)_p(v,w) = \lim_{t\to 0}(g(1)-g(0))/t^2$ is also symmetric in the choice of v, w and therefore

$$(D^2f)_p(v)(w) = (D^2f)_p(w)(v)$$

□

Corollary 13.4.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$. Suppose that f is twice differentiable. Then, there exists a symmetric matrix representation (Hessian) for $(D^2f)_x$.*

Proof. If $m = 1$, a matrix representation exists by Lemma Bilinear Matrix Representation. By Theorem 13.4.1 the linear map is symmetric. Since the entries of the Hessian matrix are $h_{ij} = D^2f(e_i)(e_j)$ and $D^2f(v)(w) = D^2f(w)(v)$, then $h_{ji} = D^2f(e_j)(e_i) = D^2f(e_i)(e_j) = h_{ij}$.

□

13.5 Taylor's Expansion Theorem

In this section we present a special case of Taylor's theorem for twice differentiable real-valued functions.

Theorem 13.5.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice differentiable in an open set containing the vectors $x, y \in \mathbb{R}^n$, then*

$$f(y) = f(x) + J_x(y-x) + \frac{1}{2}(y-x)^t H_x(y-x) + R(y,x)$$

where J is the Jacobian matrix associated with Df_x and H is the hessian associated with $(D^2 f)_x$, where $R(\cdot)$ satisfies $\lim_{y\to x} \frac{||R(y,x)||}{||y-x||^2} = 0$. Alternatively this can also be expressed as

$$f(y) = f(x) + J_x(y-x) + \frac{1}{2}(y-x)^t H_{x+\theta(y-x)}(y-x)$$

where $\theta \in (0,1)$.

Proof. Let $g(t) := f(x + t(y-x))$. Then by Taylor's expansion theorem (Theorem 11.3.1),

$$g(t) = g(0) + g^{(1)}(0)t + \frac{1}{2}g^{(2)}(0)t^2 + R^*(x, t(y-x))$$

where the vectors y, x are fixed. Using the chain rule, we can show that $g^{(1)}(t) = Df_{x+t(y-x)}(y-x)$ and $g^{(1)}(0) = Df_x(y-x)$. The second term can be represented in matrix form as $J_x(y-x)$, where J is the Jacobian of the function. The residual is an unknown function of x and the vector $t(y-x)$ with the property that $\lim_{t\to 0} R^*(x, t(y-x))/t^2 = 0$. On the other hand, by the definition of a second order derivative,

$$Df_{x+t(y-x)} - Df_x = D^2 f_x(t(y-x))(\cdot) + S(t(y-x))(\cdot)$$

where R is sublinear in its first argument. We can evaluate the linear maps on either side in the direction $(y-x)$,

$$Df_{x+t(y-x)}(y-x) - Df_x(y-x) = D^2 f_x(t(y-x))(y-x) + S(t(y-x))(y-x)$$

We can substitute the left-hand side with $g^{(1)}(t) - g^{(1)}(0)$ (using our previous result).

$$g^{(1)}(t) - g^{(1)}(0) = D^2 f_x(t(y-x))(y-x) + S(t(y-x))(y-x)$$

We can divide both sides by t. The second derivative is bilinear, so $D^2 f_x(t(y-x))(y-x)/t$

is equal to $D^2 f_x(y-x)(y-x)$ (which does not depend on t).

$$\frac{g^{(1)}(t) - g^{(1)}(0)}{t} = D^2 f_x(y-x)(y-x) + \frac{S(t(y-x))(y-x)}{t}$$

Taking limits on both sides we can show that

$$g^{(2)}(0) = \lim_{t \to 0} \frac{g^{(1)}(t) - g^{(1)}(0)}{t} = D^2 f_x(y-x)(y-x)$$

The term involving the residual converges to zero because it is sublinear. The term $D^2 f_x(y-x)(y-x)$ can be represented in terms of the Hessian as $(y-x)^t H(y-x)$. Finally notice that $g(0) = f(x)$ and $g(1) = f(y)$. We can combine our results to show that

$$f(y) = f(x) + J_x(y-x) + \frac{1}{2}(y-x)^t H_x(y-x) + R(y,x)$$

where the residual is $R(x,y) = R^*(x, y-x)$. We can multiply and divide by t^2 to do the following change of variable.

$$\begin{aligned}
\lim_{y \to x} \frac{||R^*(x, t(y-x))||}{||y-x||^2} &= \lim_{y \to x} \frac{||R^*(x, t(y-x))||}{||t(y-x)||^2} t^2 \\
&= \lim_{v \to 0} \frac{||R^*(x,v)||}{||v||^2} t^2 = 0
\end{aligned}$$

We can use similar techniques to show that $g^{(2)}(\theta) = (y-x)^t H_{x+\theta(y-x)}(y-x)$, applying the third part of the univeriate Taylor theorem. Therefore, we can alternatively state the theorem as

$$f(y) = f(x) + J_x(y-x) + \frac{1}{2}(y-x)^t H_{x+\theta(y-x)}(y-x)$$

where $\theta \in (0,1)$.

□

13.6 Exercises

1. We showed that a matrix representation exists for a linear map. Why does it have to be unique?

2. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^3 + x_2^3 \end{pmatrix}.$$

   Prove for any $p \in \mathbb{R}^2$, the matrix that represents $(D^2 f)_p$ is

$$\begin{pmatrix} 6p_1 & 0 \\ 0 & 6p_2 \end{pmatrix}.$$

3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as

$$f(x) = x^T A^T A x$$

   where A is an $n \times n$ matrix. Calculate the matrices that represent $(Df)_x$.

4. Assume that X is an $n \times k$ full rank matrix and that $Y \in \mathbb{R}^n$. Show that $\hat{\beta} = (X^t X)^{-1} X^t Y$ is the solution to the least squares criterion function by computing the first order conditions of

$$(Y - X\beta)^t (Y - X\beta)$$

Chapter 14

Comparative Statics

The main focus of this chapter is to present the implicit function theorem (IFT), which is frequently used in economic theory for comparative statics. In a canonical comparative statics setting there is a set of **endogenous variables** and a set of **exogenous variables** or **parameters**. An individual agent makes an optimal choice, which is encoded in a set of equations. We are interested in understanding how those choices depend on the underlying parameters because it allows to answer questions about policy changes and how heterogeneity of the parameters impacts the model.

In order to prove the main theorem we take an intermediate step to prove the contraction mapping theorem (CMT), which can be used to characterize existences and uniqueness of solutions in certain cases. The main parts of our proof of the (IFT) transform the problem so that we can apply the contraction mapping theorem (CMT). The (CMT) is of independent interest the foundation for finding solutions to life-cycle models in macroeconomics and structural microeconomics. To make the theorem useful on its own we need additional structure, e.g. optimization model + blackwell sufficiency conditions, which we do not cover here. However, the proof is interesting and is an opportunity to practice some of the concepts seen in the math camp so far.

In the chapter we cover two problems that arise in consumer theory and highlight how the IFT is useful to derive answers to economic questions.

14.1 Contraction Mapping Theorem

14.1.1 Preliminaries

Definition 14.1.1. Let M be a metric space. A sequence $\{x_n\}$ is **Cauchy** if for all $\epsilon > 0$ there exists an integer N such that $k, n \geq N$,

$$d(x_k, x_n) \leq \epsilon$$

Definition 14.1.2. A metric space M is complete if each Cauchy sequence in M converges to a limit in M.

By Theorem 24 in Pugh and Pugh (2002), the Euclidean space $\mathbb{R}^M$ is a complete metric space.

14.1.2 Unique Fixed Points

Definition 14.1.3. Let M be a metric space. A contraction of M is a mapping $f : M \to M$ such that for some constant $\rho < 1$ and all $x, y \in M$ we have

$$d(f(x), f(y)) \leq \rho \, d(x, y)$$

Theorem 14.1.1 (Contraction Mapping Theorem)**.** *Suppose that $f : M \to M$ is a contraction and that the space is complete. Then f has a unique fixed-point p and for any $x \in M$, the iterate $f^n := f \circ f \circ \cdots f(x)$ converges to p as $n \to \infty$.*[1]

Proof. Choose any $x_0 \in M$ and define $x_n = f^n(x_0)$. We will break down the proof into three parts.

(a) We show that for all $n \in \mathbb{N}$,

$$d(x_n, x_{n+1}) \leq \rho^n d(x_0, x_1) \tag{14.1}$$

We can show this by induction. For $n = 1$, $d(x_1, x_2) = d(f(x_0), f(x_1)) \leq \rho d(x_0, x_1)$ because f is a contraction.

Suppose that the relationship holds for some n. Then $d(x_{n+1}, x_{n+2}) = d(f(x_n), f(x_{n+1})) \leq \rho d(x_n, x_{n+1})$. By assumption of the induction step, $d(x_{n+1}, x_{n+2}) \leq \rho^{n+1} d(x_0, x_1)$.

(b) We show that the sequence $\{x_n\}$ is Cauchy. If $N \leq m \leq n$.

$$\begin{aligned}
d(x_m, x_n) &\leq d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \ldots d(x_{n-1}, x_n) && \text{Triangle Inequality (Recursive)} \\
&\leq \rho^m d(x_0, x_1) + \rho^{m+1} d(x_0, x_1) + \ldots + \rho^{n-1} d(x_0, x_1) && \text{By Equation 14.1} \\
&\leq \rho^m \left(1 + \rho + \rho^2 + \ldots \rho^{n-m-1}\right) \; d(x_0, x_1) && \text{Factorizing } d(x_0, x_1). \\
&\leq \rho^m \sum_{l=0}^{\infty} \rho^l d(x_0, x_1) && \text{Finite series less than infinite sum} \\
&\leq \frac{\rho^m}{1-\rho} d(x_0, x_1) && \text{Series converges because } \rho < 1. \\
&\leq \frac{\rho^N}{1-\rho} d(x_0, x_1) && \text{Since } \rho < 1 \text{ and } N \leq m \text{ by def.}
\end{aligned}$$

Since $\rho < 1$ we can choose N large enough so that $\frac{\rho^N}{1-\rho} d(x_0, x_1) < \epsilon$ for an arbitrary $\epsilon > 0$. Therefore $\{x_n\}$ is Cauchy.

[1] This is not a derivative, it is function iterated multiple times.

(c) Show that the sequence converges. Since M is a complete metric space and the sequence is Cauchy, $x_n \to x^*$.

(d) Show that if a function satisfies the contraction property, it is continuous at x^*. Fix $\epsilon > 0$, and choose $\delta = \epsilon$, then for all x, y, $d(f(x^*), f(y)) \leq \rho d(x^*, y) < \epsilon$. Therefore, the function is continuous.

(e) The vector x^* is a fixed point because:

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} f(x_{n-1}) = f(\lim_{n\to\infty} x_{n-1}) = f(x^*)$$

(f) Show that the fixed point is unique. Suppose that $x^* = f(x^*)$ and $y^* = f(y^*)$. Suppose that $x^* \neq y^*$, then $d(x^*, y^*) > 0$. However, $d(x^*, y^*) = d(f(x^*), f(y^*)) < \rho d(x^*, y^*)$ which is a contradiction. Therefore, $x^* = y^*$.

□

14.2 Implicit Function Theorem

Assumption 14.2.1 (Continuous Differentiability)**.** *Let $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^{n+m}$, be a continuously differentiable mapping. Define*

$$B(\theta, y) := \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{bmatrix} \quad A(\theta, y) := \begin{bmatrix} \frac{\partial f_1}{\partial \theta_1} & \cdots & \frac{\partial f_1}{\partial \theta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial \theta_1} & \cdots & \frac{\partial f_m}{\partial \theta_n} \end{bmatrix}$$

Assumption 14.2.2 (Equilibrium Condition)**.** *Let $z_0 \in \mathbb{R}^m$. There exists a $(\theta_0, y_0) \in \mathbb{R}^{n+m}$ such that $f(\theta_0, y_0) = z_0$.*

Assumption 14.2.3 (Full rank)**.** *Assume that $B := B(\theta_0, y_0)$ is full rank.*

Theorem 14.2.1. *Suppose that Assumptions Continuous Differentiability, Full rank and Equilibrium Condition hold. Then there exist open sets $V \subseteq \mathbb{R}^{n+m}$ and $\Theta \subseteq \mathbb{R}^n$ with the property that $(\theta_0, y_0) \in V$. Furthermore, for all $\theta \in \Theta$,*

(a) There exists a unique y such that $(\theta, y) \in V$ and $f(\theta, y) = z_0$.

(b) Let $y = g(\theta)$ be an implicit function of θ. Then $g : \Theta \to \mathbb{R}^m$ is continuously differentiable. Furthermore,

(i) $g(\theta_0) = y_0$.

(ii) $f(\theta, g(\theta)) = z_0$ for all $\theta \in \Theta$.

(iii) $(Dg)_{\theta_0} = -B^{-1}A$.

14.3 Proof of Implicit Value Theorem

Proof. Without loss of generality assume that $(\theta_0, y_0) = (0, 0)$ and that $z_0 = 0$. The Taylor expression for f is

$$f(\theta, y) = A\theta + By + R(\theta, y)$$

where R is sublinear and continuously differentiable with respect to (θ, y) since the other terms are continuously differentiable. Solving $f(\theta, y) = 0$ for some θ, is equivalent to solving,

$$y = -B^{-1}(A\theta + R(\theta, y))$$

where we use the fact that B^{-1} is full rank. If $R(\theta, y)$ does not depend on y then the proof is complete. Otherwise, the equation has y on both the left and right hand sides. We will show that there exists a unique y by defining a contraction mapping. Fix an arbitrary θ and treat A, B as constants, define the following function of y:

$$K_\theta(y) := -B^{-1}(A\theta + R(\theta, y))$$

Then we can find the difference between two values of y as:

$$\begin{aligned}
||K_\theta(y_1) - K_\theta(y_2)|| &= ||B^{-1}(R(\theta, y_1) - R(\theta, y_2))|| && \text{Substituting Definition} \\
&\leq ||B^{-1}||\ ||R(\theta, y_1) - R(\theta, y_2)|| && \text{Operator Norm Inequality} \\
&= ||B^{-1}||\ \Big|\Big|\frac{\partial R(x, \tilde{y})}{\partial y}\Big|\Big|\ ||y_1 - y_2|| && \text{Multivariate Mean Value Theorem}
\end{aligned}$$

The scalar $||B^{-1}||$ is a finite constant because finite linear maps have finite operator norm. The derivative of R is continuously differentiable around $(0, 0)$ (recall that we centered (θ_0, y_0) to be zero) and $(DR)_{(0,0)} = 0$. Therefore we can bound the partial derivative term by bounding the domain of (θ, y). Choose $||\theta||, ||y|| \leq r$ such that

$$\Big|\Big|\frac{\partial R(\theta, \tilde{y})}{\partial y}\Big|\Big| < \frac{1}{2||B^{-1}||}$$

That shows that $K_\theta(y)$ has the contraction property for $||x||, ||y|| \leq r$, because

$$||K_\theta(y_1) - K_\theta(y_2)|| \leq \frac{1}{2}||y_1 - y_2||$$

This is the key part of the proof because it allows us to apply the contraction mapping theorem. We also need additional steps to cover other implications of the theorem. For

example, we need to ensure that (θ_0, y_0) is contained in the set V. To ensure this notice that $K_\theta(0)$ is the value that maps our original value of y (which we centered around zero) to some other point. We need to make sure this is less than or equal to r (so that the contraction maps it to the same set).

$$||K_\theta(0)|| \leq ||B^{-1}||\ ||\theta|| + ||B^{-1}||\ ||R(\theta, 0)||$$

Combining the triangle inequality and the operator norm. We can make the residual arbitrarily small, e.g. $||x|| \leq \tau$, such that $||K_\theta(0)|| \leq r/2$. Then define the set Θ as the ball with radius τ around zero, and let M be a closed ball of radius r around zero in $\mathbb{R}^m$.

By applying the contraction mapping theorem there exists a unique y for value of $\theta \in \Theta$.

□

14.4 Application: Savings under Uncertainty

In this example we will consider a two period model where an investor is deciding how much to save for the future. Her utility function is given by the function g and depends on the amount she receives in each period. She has an initial wealth w and is planning to invest an amount x in the asset. In the first periods she consumes an amount $w - x$. The amount she consumes in the future depends on the state of the world, which is uncertain. There are $s \in \{1, \dots S\}$ possible states, each given a net income of θb_s in addition to the amount she saved in the first period.

Assumption 14.4.1 (Preferences)**.** *Let $g : \mathbb{R} \to \mathbb{R}$. g is C^3, the function is strictly increasing $g'(x) > 0, \forall x \in \mathbb{R}$ and has strictly negative second derivative $g''(x) < 0, \forall x \in \mathbb{R}$.*

A strictly positive derivative ensures that the utility function is strictly increasing (more is better). A negative second derivative captures risk aversion. The third derivative captures absolute risk aversion, which we will explore in this chapter. The function g is her per-period utility. Her expected utility is

$$f(x, w, \theta) = g(w - x) + \sum_{s=1}^{S} \pi_s g(x + \theta b_s)$$

Interpretation The sum represents the expected value over the different states, with probabilities satisfying $\pi_s \in [0, 1], \sum_{s=1}^{S} \pi_s = 1$. We also assume that $\theta > 0$. The payoffs are $b_s \in \mathbb{R}$. We could also add a discount factor to the analysis with minimal changes.

Assumption 14.4.2 (Zero-Expected Income)**.** $\sum_{s=1}^{S} \pi_s b_s = 0$.

This assumption ensures that the consumer saves some of her wealth for the next period, to compensate for the fact that she does not receive any income in the next period (in expectation). The following questions were from a previous quiz, and highlight interesting techniques used to solve the exercises.

Problem We wish to analyze the properties of the solution x^* using our differentiation theorems. Our objective is to figure out the relationship between savings and initial wealth and the variance of the asset (capture by θ).

14.4.1 Convex Combinations

1. Show that $\sum_s \pi_s g'(y_s) > 0, \forall \{y_s\} \in \mathbb{R}$ and that $\sum_s \pi_s g''(y_s) < 0, \forall \{y_s\} \in \mathbb{R}$.

   Solution. For each y_s,

   $$\begin{aligned} g'(y_s) > 0, \forall s \in \{1, ..., S\} && \text{By assumption} \\ \pi_s g'(y_s) \geq (>) 0 && \text{The inequality is strict for at least one } s \\ \sum_s \pi_s g'(y_s) > 0 \end{aligned}$$

   The inequality is weak because $\pi_s \geq 0$. It has to be strict for at least one s because $\sum_s \pi_s = 1$. A similar logic follows for $g''(y_s)$:

   $$\begin{aligned} g''(y_s) < 0, \forall s \in \{1, ..., S\} && \text{By assumption} \\ \pi_s g''(y_s) \leq (<) 0 && \text{The inequality is strict for at least one } s \\ \sum_s \pi_s g''(y_s) < 0 \end{aligned}$$

   □

14.4.2 FOC + Implicit Function Theorem

1. Define $h(x, w, \theta) := \frac{\partial f}{\partial x}$. Compute $h(x)$.

   Solution. Notice that is a composite function of $g(y)$ and $y = w - x$. Using the chain rule we find that $\frac{\partial g(w-x)}{\partial x} = \frac{\partial g(y)}{\partial y}|_{(w-x)} \frac{\partial (w-x)}{\partial x} = g'(w-x)(-1)$. Using the chain rule we can also show that $\frac{\partial g(x+\theta b_s)}{\partial x} = \frac{\partial g(y)}{\partial y}|_{(x+\theta b_s)} \frac{\partial (x+\theta b_s)}{\partial x} = g'(x + \theta b_s)$.

   Finally we use the fact that a derivative of a linear combination of functions is just a linear combination of the derivatives:

$$\begin{aligned} h(x, w, \theta) = \frac{\partial f}{\partial x} &= \frac{\partial g(w-x)}{\partial x} + \sum_s \pi_s \frac{\partial g(x + \theta b_s)}{\partial x} \\ &= -g'(w-x) + \sum_s \pi_s g'(x + \theta b_s) \end{aligned}$$

   □

2. Compute $\frac{\partial h}{\partial x}, \frac{\partial h}{\partial w}, \frac{\partial h}{\partial \theta}$.

   Proof. Using the chain rule in a similar way as before, we can show that:

$$\begin{aligned} \frac{\partial g'(w-x)}{\partial x} &= -g''(w-x) \\ \frac{\partial g'(w-x)}{\partial w} &= g''(w-x) \\ \frac{\partial g'(w-x)}{\partial \theta} &= 0 \\ \frac{\partial g'(x+\theta b_s)}{\partial x} &= g''(x+\theta b_s) \\ \frac{\partial g'(x+\theta b_s)}{\partial w} &= 0 \\ \frac{\partial g'(x+\theta b_s)}{\partial \theta} &= g''(x+\theta b_s) b_s \end{aligned}$$

Therefore we can compute each of the partial derivatives of h:

$$\frac{\partial h}{\partial x} = g''(w-x) + \sum_s \pi_s g''(x+\theta b_s) \qquad \text{(Endogenous Variable)}$$

$$\frac{\partial h}{\partial w} = -g''(w-x) \qquad \text{(Parameter)}$$
$$\frac{\partial h}{\partial \theta} = \sum_s \pi_s g''(x+\theta b_s) b_s \qquad \text{(Parameter)}$$

□

Define the optimal x^* as the one that satisfies the first order condition $h(x^*, w, \theta) = 0$. Assume that $\frac{\partial h(x^*,w,\theta)}{\partial x} \neq 0$.

3. Use the implicit function theorem to show that $x^*(w,\theta)$ is increasing with respect to w.

Proof. The implicit function theorem says that: $\frac{\partial x^*(w,\theta)}{\partial w} = -\frac{\partial h}{\partial x}|^{-1}_{(x^*,w,\theta)} \frac{\partial h}{\partial w}|_{(x^*,w,\theta)}$
Because of the result in 1.(a), $\frac{\partial h}{\partial x}|_{(x^*,w,\theta)} < 0$ and $\frac{\partial h}{\partial w}|_{(x^*,w,\theta)} > 0$. Therefore $\frac{\partial x^*(w,\theta)}{\partial w} > 0$. This means that $x^*(w,\theta)$ is increasing in w.

Note: The implicit function theorem can be applied to each parameter separately. In matrix form the implicit function theorem says the jacobian of $x^*(w,\theta)$ with respect to (w,θ) is equal to $-J^{-1}_{(x^*,w,\theta),x} J_{(x^*,w,\theta),(x,w)}$, where $J_{(x^*,w,\theta),x}$ is the jacobian w.r.t to x and $J_{(x^*,w,\theta),(w,\theta)}$ with respect to (w,θ), which is equal to $[J_{(x^*,w,\theta),(w)}, J_{(x^*,w,\theta),(\theta)}]$. Therefore, there is no loss of generality in considering each parameter separately.

□

4. Compute an expression for $\frac{\partial x^*(w,\theta)}{\partial \theta}$.

Proof. By the implicit function theorem:
$\frac{\partial x^*(w,\theta)}{\partial \theta} = -\frac{\partial h}{\partial x}|^{-1}_{(x^*,w,\theta)} \frac{\partial h}{\partial \theta}|_{(x^*,w,\theta)}$
$\frac{\partial x^*(w,\theta)}{\partial \theta} = -\frac{\sum_s \pi_s g''(x+\theta b_s) b_s}{g''(w-x)+\sum_s \pi_s g''(x+\theta b_s)}$

□

14.4.3 Absolute Risk Aversion

The coefficient of absolute risk aversion is defined as

$$A(x) = -\frac{g''(x)}{g'(x)}$$

We will assume that $A(x)$ is decreasing $\forall x \in \mathbb{R}$. We will show that this is related to properties of the third-order derivative. The importance of the coefficient of absolute risk aversion emerges in comparative statics exercises with savings, because when we derive the first order conditions, we sometimes find terms involving the third derivative.

5. Rewrite the equation as: $g''(x) = -A(x)g'(x)$. Show that $g'''(x) >?0$.

   Proof. We can show that

$$\begin{aligned} & g''(x) = -A(x)g'(x) \forall x \in \mathbb{R} \\ \implies & g'''(x) = -A(x)g''(x) - A'(x)g'(x) \end{aligned}$$

   Since $g'(x) > 0$ and $A'(x) < 0$ (since A is decreasing), then $-A'(x)g'(x) > 0$. The value $A(x) = -\frac{g''(x)}{g'(x)}$ is positive because $g''(x)$ is negative, $g'(x)$ is positive. Therefore $-A(x)g''(x)$ is positive. Putting the two things together:

$$g'''(x) > 0$$

   □

6. Show that $\sum_{s=1}^{S} \pi_s g''(x + \theta b_s) b_s > 0$ and use it to find the sign of $\frac{\partial x^*(w,\theta)}{\partial \theta}$. [Hint: show that $\sum_{s=1}^{S} \pi_s g''(x) b_s = 0$ and construct a Taylor expansion between x and $x + \theta b_s$ for each s].

   Proof. Fixed typo: Originally the equation was stated in terms of $v''(y)$. Changed it to $g''(y)$. This question also requires you to assume that $\theta > 0$.

$$\sum_{s=1}^{S} \pi_s b_s = 0 \implies g''(x) \sum_{s=1}^{S} \pi_s = \sum_{s=1}^{S} \pi_s b_s g''(x) = 0$$

Therefore we can rewrite the following equation as:

$$\begin{aligned}\sum_{s=1}^{S} \pi_s g''(x+\theta b_s) b_s &= \sum_{s=1}^{S} \pi_s g''(x+\theta b_s) b_s - \sum_{s=1}^{S} \pi_s g''(x) \\ &= \sum_{s=1}^{S} \pi_s b_s [g''(x+\theta b_s) - g''(x)]\end{aligned}$$

We can do a first order taylor expansion between x and $x+\theta b_s$. Let $\xi_s \in [x, x+\theta b_s]$:

$$\begin{aligned}\sum_{s=1}^{S} \pi_s b_s [g''(x+\theta b_s) - g''(x)] &= \sum_{s=1}^{S} \pi_s b_s g'''(\xi_s) \theta b_s \\ &= \sum_{s=1}^{S} \pi_s g'''(\xi_s) \theta (b_s)^2 > 0\end{aligned}$$

The last result follows by assuming that $\theta > 0$ and the fact that $g'''(x) > 0 \forall x \in \mathbb{R}$.

Therefore $\frac{\partial h(x^*, w, \theta)}{\partial \theta} > 0$. Since $\frac{\partial h(x^*, w, \theta)}{\partial x} < 0$, then by the implicit function theorem. $\frac{\partial x^*(w, \theta)}{\partial \theta} > 0$.

□

14.5 Exercises

1. Consider the Auctions Example in previous chapters. Show that $b^*(v)$ is increasing in v. [Hint: Use the implicit function theorem].

2. Consider the following Keynesian IS-LM model. Suppose

$$\begin{aligned} Y &= C(Y-T) + I(r) + G \\ M &= L(Y, r) \end{aligned}$$

where Y is GDP, T is taxes, r is interest rate, G is government spending and M is money supply. The functions $C(\cdot)$, $I(\cdot)$ and $L(\cdot, \cdot)$ are consumption function, investment function and money supply function respectively. Assume they are continuously differentiable and

$$0 < C'(x) < 1, \quad I'(r) < 0, \quad \frac{\partial L}{\partial Y} > 0, \quad \text{and} \quad \frac{\partial L}{\partial r} < 0.$$

Suppose G, M and T are independent variables which can be controlled, Y and r are dependent variables determined by G, M and T. Analyze the relationships between $\{Y, r\}$ and $\{G, M, T\}$.

Chapter 15

Concavity (Convexity)

This section draws most of its results from Sundaram et al. (1996) Chapter 7. In the majority of the cases the proofs are taken directly.

15.1 Set Definition

Definition 15.1.1. Let $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ convex. The subgraph of f and the epigraph of f are defined as:

$$\text{sub } f = \{(x, y) \in \mathcal{D} \times \mathbb{R} \mid f(x) \geq y\} \tag{15.1}$$

$$\text{epi } f = \{(x, y) \in \mathcal{D} \times \mathbb{R} \mid f(x) \leq y\} \tag{15.2}$$

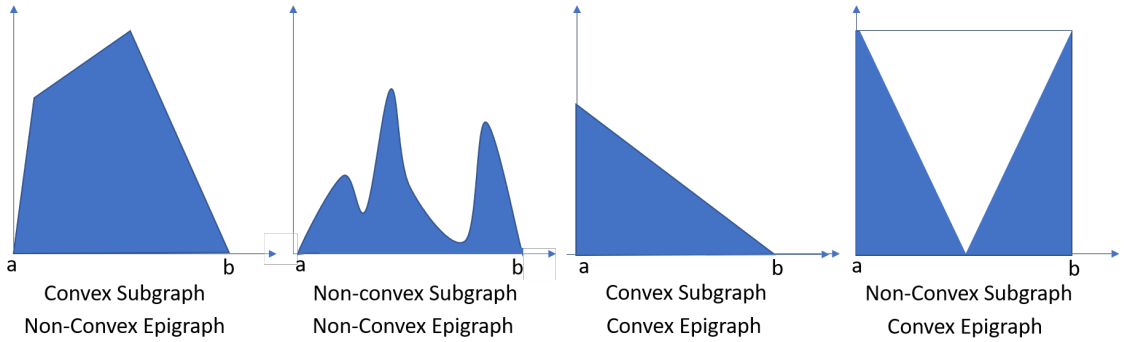

Figure 15.1: The figure depict the subgraph of a function defined over the interval $[a, b]$. The shaded region is the subgraph and the epigraph is the blank region above the subgraph. Notice that the subgraph and epigraph are defined with weak inequalities so they intersect at $y = f(x)$.

Definition 15.1.2. A real-valued function is said to be **concave** if sub f is a convex set. It is said to be **convex** if epi f is a convex set.

Theorem 15.1.1. *A function $f : \mathcal{D} \to \mathbb{R}^n$ defined on a convex set $\mathcal{D} \subseteq \mathbb{R}^n$, is a concave function if and only if for all $x, y \in \mathcal{D},\ \lambda \in (0, 1)$,*

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$$

Similarly, the function is convex if for all $x, y \in \mathcal{D},\ \lambda \in (0, 1)$

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Proof. We will only prove the relationship for concave functions. The proof for convex functions is analogous.

$\implies$ Suppose that the subgraph of f is a convex set. Let x_1, x_2 be arbitrary points in $\mathcal{D}$. Then $(x_1, f(x_1))$ and $(x_2, f(x_2))$ are contained in sub f. Since the set is convex, if $\lambda \in (0, 1)$ then $x_\lambda := \lambda(x_1, f(x_1)) + (1 - \lambda)(x_2, f(x_2))$ is contained in sub f. This can be rewritten as $(\lambda x_1 + (1 - \lambda)x_2, \lambda f(x_1) + (1 - \lambda)f(x_2))$. By definition, a point (w, z) is contained in the subgraph if $f(w) \geq z$, therefore,

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2))$$

$\impliedby$ Now suppose that we choose arbitrary points $(x_1, y_1), (x_2, y_2) \in$ sub f, i.e. $x_1, x_2 \in \mathcal{D}$ and $f(x_1) \geq y_1$ and $f(x_2) \geq y_2$. We want to show that for $\lambda \in (0, 1)$, $(x_\lambda, y_\lambda) := \lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \in$ sub f. Notice

$$f(x_\lambda) = f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \geq \lambda y_1 + (1 - \lambda)y_2 = y_\lambda$$

Because $f(x_\lambda) \geq y_\lambda$, we have shown that the subgraph is convex.

□

15.1.1 Strict Concavity (Convexity)

Definition 15.1.3. A real-valued function f defined over a convex set $\mathcal{D} \subseteq \mathbb{R}^n$ is said to be **strictly concave** if for all $x, y \in \mathcal{D}$ such that $x \neq y$ and for all $\lambda \in (0,1)$,

$$f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda) f(y)$$

Strictly convex functions are defined analogously by exchanging the inequality.

Lemma 15.1.1. *A function $f : \mathcal{D} \to \mathbb{R}$ is concave on $\mathcal{D}$ if and only if the function $-f$ is convex on $\mathcal{D}$. It is strictly concave if and only if $-f$ is strictly convex.*

Proof. Let $x_1, x_2 \in \mathcal{D}$ and $\lambda \in (0,1)$ then

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2) \quad \iff \quad -f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda(-1) f(x_1) (1-\lambda)(-1) f(x_2)$$

□

This lemma helps us establish that we can prove the majority of theorems for concave function, without loss of generality.

15.1.2 Conic Combinations of Concave Functions

Lemma 15.1.2. *Let $\mathcal{F}$ be a collection of real-valued concave functions defined on a convex set $\mathcal{D} \subseteq \mathbb{R}^n$. Then for all positive integers K, vectors of weights $\theta \in \mathbb{R}_+^K$ and functions $f_k \in \mathcal{F}$, then $f := \sum_{k=1}^K \theta_k f_k$ is also concave.*

Proof. Let $x_1, x_2 \in \mathcal{D}$ and let $\lambda \in (0,1)$. Since each function is concave

$$f_k(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f_k(x_1) + (1-\lambda) f_k(x_2) \quad \forall k \in \{1, \ldots, K\}$$

Multiplying each term by a non-negative quantity preserves the inequality. We can then sum the terms over $k \in \{1, \ldots, K\}$.

$$\sum_{k=1}^K \theta_k f_k(\lambda x_1 + (1-\lambda)x_2) \geq \lambda \sum_{k=1}^K \theta_k f_k(x_1) + (1-\lambda) \sum_{k=1}^K \theta_k f_k(x_2) \quad \forall k \in \{1, \ldots, K\}$$

which implies that $f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2)$.

□

15.2 Derivative Characterization

15.2.1 First Derivative

Theorem 15.2.1. *Let $\mathcal{D}$ be an open and convex set in $\mathbb{R}^n$, and let $f : \mathcal{D} \to \mathbb{R}$ be differentiable on $\mathcal{D}$. Then f is concave on $\mathcal{D}$ if and only if*

$$Df(x)(y-x) \geq f(y) - f(x), \quad \forall x, y \in \mathcal{D}$$

Proof. $\implies$ Suppose that f is concave. Let $x, y \in \mathcal{D}$, then for all $t \in (0,1)$, $f(ty+(1-t)x) \geq tf(y) + (1-t)f(x)$. We can subtract $f(x)$ on either side and divide by t,

$$\frac{f(ty+(1-t)x) - f(x)}{t} \geq \frac{tf(y) + (1-t)f(x) - f(x)}{t} = f(y) - f(x), \quad t \in (0,1)$$

The second equality follows because some of the terms cancel out. Furthermore, we can rewrite $ty + (1-t)x$ as $x + t(y-x)$. Then we can take the limit from above.

$$\lim_{t \downarrow 0} \frac{f(x+t(y-x)) - f(x)}{t} \geq f(y) - f(x)$$

Define $g(t) := f(x + t(y-x))$ and let $g'(t) = \lim_{t \to 0} \frac{g(t)-g(0)}{t}$. Using the chain rule we can show that $g'(t) = Df_x(y-x)$. Since the limit exists then it must be equal to $\lim_{t \downarrow 0} \frac{g(t)-g(0)}{t}$ because of Theorem Equivalent Limit Definitions. therefore,

$$Df_x(y-x) \geq f(y) - f(x)$$

$\impliedby$ Now suppose that for all $x_1, x_2 \in \mathcal{D}$ we have

$$Df_{x_1}(x_2 - x_1) \geq f(x_2) - f(x_1)$$

Pick any $x, y \in \mathcal{D}$ and $\lambda \in (0,1)$. We will show that we must have $f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y)$, which will establish that f is concave on $\mathcal{D}$. For expositional convenience, define the convex combination

$$z := \lambda x + (1-\lambda)y$$

By assumption we also have,

$$Df(z)(x-z) \geq f(x) - f(z) \tag{15.3}$$

$$Df(z)(y-z) \geq f(y) - f(z) \tag{15.4}$$

Notice that $\lambda(x-z)+(1-\lambda)(y-z)=0$. Multiplying Equation 15.3 by $\lambda/(1-\lambda)$ and adding the two equations, we obtain

$$\frac{\lambda}{1-\lambda}f(x)+f(y)-\frac{1}{1-\lambda}f(z)\leq 0$$

Then multiplying the equation by $(1-\lambda)$ and rearranging,

$$\lambda f(x)+(1-\lambda)f(y)\leq f(z)=f(\lambda x+(1-\lambda)y)$$

□

15.2.2 Second Derivative

Definition 15.2.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable, real-valued function. The second derivative $D^2 f_x$ evaluated at a point $x \in \mathbb{R}^n$ is negative definite if for all $v \in \mathbb{R}^n$ such that $v \neq 0$, $D^2 f_x(v)(v) < 0$. If the inequality is weak for at least one v, then $D^2 f_x$ is negative semi-definite.

For real-valued functions, it can be shown that negative (semi) definiteness of the second derivative is equivalent to saying that its associated hessian matrix is negative (semi) definite. Notice that symmetry is already guaranteed by Theorem 13.4.1 so we do not need to verify it as part of the definition of negative (semi) definiteness.

Theorem 15.2.2. *Let $f : \mathcal{D} \to \mathbb{R}$ be a twice differentiable function, where $\mathcal{D} \subseteq \mathbb{R}^n$ is open and convex. Then*

1. *f concave if and only if $D^2 f_x$ is a negative semi-definite matrix for all $x \in \mathcal{D}$.*
2. *$D^2 f_x$ is negative definite for all $x \in \mathcal{D}$, then f is strictly concave.*

An analogous result holds for (strictly) convex functions and the Hessian being (positive definite) positive semi-definite.

Proof. We break down the proof into two parts.

$\Longleftarrow$ Let $x_1, x_2 \in \mathbb{R}^R$ and $\lambda \in (0, 1)$. For notational simplicity define $h := x_1 - x_2$ and $x_\lambda := \lambda x_1 + (1 - \lambda)x_2$. We will do two separate taylor expansions of f around x_λ.

$$f(x_1) = f(x_\lambda) + (Df)_{x_\lambda}((1-\lambda)h) + \frac{1}{2}(D^2 f)_\theta((1-\lambda)h)((1-\lambda)h)$$
$$f(x_2) = f(x_\lambda) - (Df)_{x_\lambda}(\lambda h) + \frac{1}{2}(D^2 f)_\psi(\lambda h)(\lambda h)$$

where θ is contained in the line segment between x_1 and x_λ and ψ is contained in the line segment between x_λ and x_2. Multiply the first equation by λ, the second equation by $(1-\lambda)$. We can use linearity of $(Df)x_\lambda$ to simplify each equation.

$$\lambda f(x_1) = \lambda f(x_\lambda) + (Df)_{x_\lambda}(\lambda(1-\lambda)h) + \frac{1}{2}(D^2 f)_\theta((1-\lambda)h)((1-\lambda)h)$$
$$(1-\lambda)f(x_2) = (1-\lambda)f(x_\lambda) - (Df)_{x_\lambda}((1-\lambda)\lambda h) + \frac{1}{2}(D^2 f)_\psi(\lambda h)(\lambda h)$$

We can add the two equations together

$$\lambda f(x_1) + (1-\lambda)f(x_2) = f(x_\lambda) + \frac{1}{2}\lambda(D^2 f)_\theta((1-\lambda)h)((1-\lambda)h) + \frac{1}{2}(1-\lambda)(D^2 f)_\psi(\lambda h)(\lambda h)$$

If $D^2 f$ is negative semi-definite then the two terms in the equation are weakly negative then we can prove concavity

$$\lambda f(x_1) + (1-\lambda) f(x_2) \leq f(x_\lambda)$$

Similarly if $D^2 f$ is negative definite, then we get strict concavity,

$$\lambda f(x_1) + (1-\lambda) f(x_2) < f(x_\lambda)$$

$\implies$ Let $x \in \mathcal{D}$. Choose an arbitrary non-zero $\epsilon \in \mathbb{R}^n$. Define $x_2 = x + t\epsilon, x_1 = x - t\epsilon$ for a scalar $t > 0$. Since the set $\mathcal{D}$ is open, there exists a small enough t such that $x_1, x_2 \in \mathbb{R}^n$. Carry out the following Taylor expansions using the form that includes the residual as a separate term.

$$f(x + t\epsilon) = f(x) + (Df)_{x_\lambda}(t\epsilon) + \frac{1}{2}(D^2 f)_x(t\epsilon)(t\epsilon) + \frac{R(x, x+t\epsilon)}{t^2}$$
$$f(x - t\epsilon) = f(x) - (Df)_{x_\lambda}(t\epsilon) + \frac{1}{2}(D^2 f)_x(t\epsilon)(t\epsilon) + \frac{S(x, x-t\epsilon)}{t^2}$$

We can add the terms together and divide by 1/2,

$$\frac{1}{2}f(x + t\epsilon) + \frac{1}{2}f(x - t\epsilon) = f(x) + (D^2 f)_x(t\epsilon)(t\epsilon) + R(t\epsilon)(t\epsilon) + S(-t\epsilon)(t\epsilon)$$

Rearranging the equation and dividing by t. We simplify the equation by using bilinearity of $(D^2 f)_x$,

$$\frac{\frac{1}{2}f(x + t\epsilon) + \frac{1}{2}f(x - t\epsilon) - f(x)}{t^2} = (D^2 f)_x(\epsilon)(\epsilon) + \frac{R(x, x+t\epsilon)}{t^2} + \frac{S(x, x-t\epsilon)}{t^2}$$

We can take the limit of the residuals, consider the first residual,

$$\lim_{t\to 0} \frac{R(x, x+t\epsilon)}{t^2} \frac{||\epsilon||}{||\epsilon||} = \lim_{t\to 0} \frac{R(x, x+t\epsilon)}{||t\epsilon||^2} ||\epsilon||^2 = \lim_{t\epsilon \to 0} \frac{R(x, x+t\epsilon)}{||t\epsilon||^2} = 0$$

We can apply a similar strategy with $S(x, x - t\epsilon)$. The left-hand side is non-positive by the definition of concavity since $x = \frac{1}{2}(x + t\epsilon) + \frac{1}{2}(x - t\epsilon)$. We can take limits on both sides to show that

$$\lim_{t\to 0} \frac{\frac{1}{2}f(x + t\epsilon) + \frac{1}{2}f(x - t\epsilon) - f(x)}{t^2} = D^2 f_x(\epsilon)(\epsilon)$$

Therefore, we it follows that $D^2 f_x(\epsilon)(\epsilon) \leq 0$, for any arbitrary non-zero $\epsilon \in \mathbb{R}^n$. Therefore, D_x^2 is negative-semi definite. $\square$

15.3 Special Topological Properties

Theorem 15.3.1. *Let $f : \mathcal{D} \to \mathbb{R}$ be a concave function define on a convex set $\mathcal{D} \subseteq \mathbb{R}^n$. Then f is continuous in the interior of D.*

Proof. Let $x \in int(\mathcal{D})$ then there exists an open ball $B(x, \epsilon) \subseteq \mathcal{D}$ for some $\epsilon > 0$ such that $f(B(x, \epsilon))$ (prove as an exercise). Choose an $\epsilon^* \in (0, \epsilon)$ and define A as the set of vectors $z \in \mathbb{R}$ such that $||z - x|| = \epsilon^*$. By construction, $A \subseteq B(x, \epsilon) \subseteq \mathcal{D}$.

WLOG choose an arbitrary sequence $x_k \in B(x, \epsilon^*)$ such that $x_k \to x$. For all k there exists $z_k \in A$ such that $x_k = \theta_k x + (1 - \theta)z_k$ for some $\theta \in (0, 1)$. The vector $z_k - x$ is in the same direction as $x_k - x$ but is constrained to have a particular length that does not depend on k. This guarantees that as $k \to \infty$, $\theta_k \to 1$. Therefore, by concavity of f,

$$f(x_k) = f(\theta_k x + (1 - \theta_k)z_k) \geq \theta_k f(x) + (1 - \theta_k)f(z_k)$$

Taking limits on both sides and since $\theta_k \to 1$ (because x_k converges to x and $z_k - x$ has fixed length).

$$\liminf_{k\to\infty} f(x_k) \geq f(x) + \liminf_{k\to\infty} (1 - \theta_k)z_k = f(x)$$

Similarly, we can find a vector $w_k \in A$ and $\lambda_k \in (0, 1)$ such that $x = \lambda_k x_k + (1 - \lambda_k)w_k$. We can once again exploit concavity of f, to show that:

$$f(x) = f(\lambda_k x_k + (1 - \lambda_k)w_k) \geq \lambda_k f(x_k) + (1 - \lambda_k)f(w_k)$$

Since $\lambda_k \to 1$ as $k \to \infty$, by taking limits we obtain,

$$f(x) \geq \limsup_{k\to\infty} f(x_k)$$

Since $f(x) \leq \liminf_{k\to\infty} f(x_k) \leq \limsup_{k\to\infty} f(x_k) \leq f(x)$, then $f(x) = \lim_{k\to\infty} f(x_k)$ for any arbitrary sequence $\{x_k\}$. Therefore, f must be continuous.

□

Chapter 16

Quasiconcavity

16.1 Set Definition

The proofs in this section come from Sundaram et al. (1996), Chapter 8.

Definition 16.1.1. Let $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$ convex. The upper contour set of f and the lower contour set of f at $a \in \mathbb{R}$, are

$$U_f(a) = \{x \in \mathcal{D} \mid f(x) \geq a\} \tag{16.1}$$

$$L_f(a) = \{x \in \mathcal{D} \mid f(x) \leq a\} \tag{16.2}$$

Definition 16.1.2. A real-valued function is said to be **quasiconcave** if for all $a \in \mathbb{R}$, $U_f(a)$ is a convex set. It is said to be **quasiconvex** if for all $a \in \mathbb{R}$, $L_f(a)$ is a convex set.

Theorem 16.1.1. *A function $f : \mathcal{D} \to \mathbb{R}$ defined on a convex set $\mathcal{D} \subseteq \mathbb{R}^n$, is a quasiconcave function if and only if for all $x, y \in \mathcal{D},\ \lambda \in (0, 1)$,*

$$f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}$$

Similarly, the function is quasiconvex if for all $x, y \in \mathcal{D},\ \lambda \in (0, 1)$

$$f(\lambda x + (1 - \lambda)y) \leq \max f(x), f(y)$$

Proof. We will only show the proof for quasiconcavity, the proof for quasiconvexity is analogous.

$\implies$ Suppose that f is quasiconcave, i.e. that $U_f(a)$ is a convex set for each $a \in \mathbb{R}$. Let $x, y \in \mathcal{D}$ and $\lambda \in (0, 1)$. Assume without loss of generality, that $f(x) \geq f(y)$. Letting

$a = f(y)$, we have $x, y \in U_f(a)$. By the convexity of $U_f(a)$, we have $\lambda x + (1-\lambda)y \in U_f(a)$, which means

$$f(\lambda x + (1-\lambda)y) \geq a = f(y) = \min\{f(x), f(y)\}$$

$\Longleftarrow$ Now suppose we have $f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}$ for all $x, y \in \mathcal{D}$ and for all $\lambda \in (0,1)$. Let $a \in \mathbb{R}$. If $U_f(a)$ is empty or contains only one point point, it is convex, so suppose that it contains at least two points x and y. Then $f(x) \geq a$ and $f(y) \geq a$, so $\min\{f(x), f(y)\} \geq a$. Now, for any $\lambda \in (0,1)$, we have $f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}$ by hypothesis and so $\lambda x + (1-\lambda)y \in U_f(a)$. Since $a \in \mathbb{R}$ was arbitrary, the proof is complete.

□

16.1.1 Strict Quasi Concavity (Quasi Convexity)

Definition 16.1.3. A real-valued function f defined over a convex set $\mathcal{D} \subseteq \mathbb{R}^n$ is said to be **strictly quasiconcave** if for all $x, y \in \mathcal{D}$ such that $x \neq y$ and for all $\lambda \in (0,1)$,

$$f(\lambda x + (1-\lambda)y) > \min\{f(x), f(y)\}$$

Strictly quasiconvex functions satisfy,

$$f(\lambda x + (1-\lambda)y) < \max\{f(x), f(y)\}$$

Lemma 16.1.1. *The function $f : \mathcal{D} \to \mathbb{R}$ is quasiconcave on $\mathcal{D}$ if and only if $-f$ is quasiconvex on $\mathcal{D}$. It is strictly quasiconcave on D if and only if $-f$ is strictly quasiconvex on $\mathcal{D}$.*

Proof. Let $x_1, x_2 \in \mathcal{D}$ and $\lambda \in (0,1)$ then

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \min\{f(x_1), f(x_2)\} \quad \Longleftrightarrow \quad -f(\lambda x_1 + (1-\lambda)x_2) \leq \max\{-f(x_1), -f(x_2)\}$$

Multiplying the min operator by a negative number switches to a max whose inner arguments are multiplied by the negative number. A similar proof applies to strict concavity, using a strict inequality instead of a weak inequality.

□

16.2 Derivative Characterization

Theorem 16.2.1. *Let $f : \mathcal{D} \to \mathbb{R}$ be a continuously differentiable function where $\mathcal{D} \subseteq \mathbb{R}^n$ is convex and open. Then f is a quasiconcave function on $\mathcal{D}$ if and only if it is the case that for any $x, y \in \mathcal{D}$,*

$$f(y) \geq f(x) \quad \Longrightarrow \quad Df_x(y - x) \geq 0$$

Proof. $\Longrightarrow$ First suppose that f is quasiconcave on $\mathcal{D}$ and let $x, y \in \mathcal{D}$ such that $f(y) \geq f(x)$. Let $t \in (0, 1)$. Since f is quasiconcave, we have

$$f(x + t(y - x)) = f((1 - t)x + ty) \geq \min\{f(x), f(y)\} = f(x).$$

Therefore, it is the case that for all $t \in (0, 1)$, we have

$$\frac{f(x + t(y - x)) - f(x)}{t} \geq 0$$

As $t \downarrow 0$ the left hand side converges to $Df_x(y - x)$, so $Df_x(y - x) \geq 0$.

$\Longleftarrow$ Now suppose that for all $x, y \in \mathcal{D}$ such that $f(y) \geq f(x)$, we have $Df_x(y - x) \geq 0$. Pick any $x, y \in \mathcal{D}$, and suppose without loss of generality that $f(x) = \min\{f(x), f(y)\}$. We will show that for any $t \in [0, 1]$, we must also have $f((1 - t)x + ty) \geq \min\{f(x), f(y)\}$, establishing the quasiconcavity of f. Let $z(t) = (1 - t)x + ty$.

Define $g(t) = f(x + t(y - x))$. Note that $g(0) = f(x) \leq f(y) = g(1)$; and that g is C^1 on $[0, 1]$ with $g'(t) = Df[x + t(y - x)](y - x)$. We will show that if $t^* \in (0, 1)$ is any point such that $f(z(t^*)) \leq f(x)$ we must have $g'(t^*) = 0$.

Suppose that $t^* \in (0, 1)$ we have $f(x) \geq f(z(t^*))$. Then by hypothesis, we must also have $Df_{z(t^*)}(x - z(t^*)) = -t^* Df_{z(t^*)}(y - x) \geq 0$. Since $t^* > 0$, this implies that $g'(t^*) \leq 0$. On the other hand it is also true that $f(y) \geq f(x) \geq f(z(t^*))$, so we must also have $Df_{(z(t^*))}(y - z(t^*)) = (1 - t^*)Df_{z(t^*)}[y - x] \geq 0$. Since $t^* < 1$, this implies in turn that $g'(t^*) \geq 0$. It follows that $g'(t^*) = 0$.

□

16.3 Conic Combinations Not Quasiconcave

In this section we present a counter-example showing than conic combinations of quasiconcave functions are not, in general, quasiconcave. This stands in contrast with concave functions, that were preserved under conic combinations. This finding has implications for decision theory, where expectations can be expresses as finite (or infinite) conic combinations.

Example 10. *Consider the following quasiconcave functions defined over* $\mathbb{R}$*,*

$$f(x) = x^3 \quad g(x) = 1 - x^2$$

Then the addition of the functions is not quasiconcave:

$$h(x) = x^3 + 1 - x^2$$

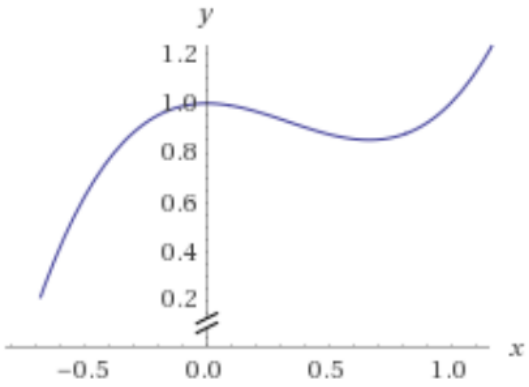

Figure 16.1: The figure depict the function $h(x)$.

16.4 Concavity and Quasi-Concavity

Theorem 16.4.1. *Let $f : \mathcal{D} \to \mathbb{R}$, $\mathcal{D} \subseteq \mathbb{R}^n$. If f is concave on $\mathcal{D}$, it is also quasiconcave. If f is convex on $\mathcal{D}$, it is also quasiconvex on $\mathcal{D}$.*

Proof. Suppose f is concave. Then for all $x, y \in \mathcal{D}$ and $\lambda \in (0, 1)$ we have

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\geq \lambda f(x) + (1-\lambda) f(y) \\ &\geq \lambda \min\{f(x), f(y)\} + (1-\lambda)\min\{f(x), f(y)\} \\ &\geq \min\{f(x), f(y)\} \end{aligned}$$

□

Theorem 16.4.2. *If $f : \mathcal{D} \to \mathbb{R}$ is quasiconcave on $\mathcal{D}$, and $\phi : \mathbb{R} \to \mathbb{R}$ is a monotone non-decreasing function, then the composition $\phi \odot f$ is a quasiconcave function from $\mathcal{D}$ to $\mathbb{R}$. In particular, any monotone transform of a concave function results in a quasiconcave function.*

Proof. Pick any $x, y \in \mathcal{D}$, and any $\lambda \in (0, 1)$. Since f is quasiconcave by hypothesis, we have

$$f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}$$

Since ϕ is non-decreasing, this implies that

$$\phi(f(\lambda x + (1-\lambda)(y)) \geq \phi(\min\{f(x), f(y)\}) = \min\{\phi(f(x)), \phi(f(y))\}$$

□

Part III

Answer Key

Chapter 17

Suggested Solutions

17.1 Overview of Linear Algebra

1. Suppose that $T(\boldsymbol{x}) = A\boldsymbol{x}$ and that $F(y) = B\boldsymbol{y}$, with $A_{m\times n}$ and $B_{k\times m}$.

   (a) Show that $G = F(T(\boldsymbol{x}))$ is also a linear map.

   Solution. $G = F(T(\boldsymbol{x})) = F(A\boldsymbol{x}) = B(A\boldsymbol{x}) = BA\boldsymbol{x} = C\boldsymbol{x}$, where C is a $k \times n$ matrix. Using Lemma 1.3.1 because G can be expressed as the multiplication of a constant matrix times a vector, it is a linear map. □

   (b) Show that $||G|| \leq ||F||\ ||T||$. Is the composite of two linear maps continuous?

   Solution. Let $\boldsymbol{x} \in \mathbb{R}^n$. Using the operator norm inequality twice.

$$\begin{aligned} ||F(T(\boldsymbol{x}))|| &\leq ||F||\ ||T(\boldsymbol{x})|| \\ &\leq ||F||\ ||T||\ ||\boldsymbol{x}|| \end{aligned}$$

   Restrict attention to vectors of unit length such that $||\tilde{\boldsymbol{x}}|| = 1$, then $||G(\tilde{\boldsymbol{x}})|| \leq ||F||\ ||T||$. Then right hand side does not depend on the input vector. Therefore we can take the supremum on the left-hand side to show that $||G|| \leq ||F||\ ||T||$. To prove continuity it suffices to use the fact that G is a linear map by using Theorem 1.5.1. □

   (c) Assume that P is a square matrix. Use part (b) to show that for any non-negative integer t, $||P^t|| \leq ||P||^t$.

Solution. If $t = 1$ then $||P^t|| = ||P||$. Suppose that the statement holds for some $t \geq 1$, then using the previous lemma $||P^{t+1}|| = ||P^t P|| \leq ||P^t||\ ||P||$, where $||P^t|| \leq ||P||^t$ by the induction hypothesis. Therefore, $||P^{t+1}|| \leq ||P||^{t+1}$. By the principle of induction, the result is proved. □

(d) Show that if $\boldsymbol{x}$ is a probability vector, then $||\boldsymbol{x}|| \geq a$ for some $a > 0$.

Solution. First, we show that $||\boldsymbol{x}|| > 0$. We proceed using proof by contradiction. Since $||\boldsymbol{x}|| \geq 0$, assume (by contradiction) that $||\boldsymbol{x}|| = 0$ then $\boldsymbol{x} = \mathbf{0}_n$, meaning all its entries are zero. However, the entries of a probability vector must add up to one. Thus we must have $||\boldsymbol{x}|| > 0$. But this is not sufficient for the statement of interest–we want to show $||\boldsymbol{x}||$ is bounded below by a positive constant.

Let $\mathcal{P}$ denote the space of probability vectors. Now we show it is compact. Consider a sequence of probability vectors $x_k = (x_{1k}, \ldots, x_{nk})$ with $\sum_i x_{ik} = 1$, $x_k \geq 0$ for all k, and $x_k \to x$ as $k \to \infty$. Since limits preserve equalities and weak inequalities, it follows that $\sum_i x_i = 1$ and $x_i \geq 0$, therefore the limit is still a probability vector and hence the set $\mathcal{P}$ is closed. It is also bounded because all its entries are non-negative and less that or equal to one. Therefore, it is compact.

The function $f(\boldsymbol{x}) = \sqrt{\boldsymbol{x}^t \boldsymbol{x}}$ is continuous because it is a polynomial of the entries of the vector. The extreme-value theorem states that if a function is continuous and the space is compact then it has a maximum and a minimum. Therefore, a minimum exists, and denote $\inf_{\boldsymbol{x} \in \mathcal{P}} \sqrt{\boldsymbol{x}^t \boldsymbol{x}} = \sqrt{\boldsymbol{x}_*^t \boldsymbol{x}_*}$ for some $x_* \in \mathcal{P}$. By our previous result $||x_*|| > 0$ and therefore $||x||$ is bounded away from zero.

In fact, $1/\sqrt{n} \leq ||\boldsymbol{x}|| \leq 1$, where n is the dimension of $||\boldsymbol{x}||$. The shortest probability vector has the value $1/n$ as each component of the vector, while the longest probability vector has the value 1 in a single component and 0 in all others. This constitutes an easier proof for the statement:

$$\begin{aligned} ||\boldsymbol{x}|| &= \sqrt{x_1^2 + x_2^2 + ... + x_n^2} = \sqrt{n}\sqrt{\frac{x_1^2 + x_2^2 + ... + x_n^2}{n}} \\ &\geq \sqrt{n}\frac{x_1 + x_2 + ... + x_n}{n} = 1/\sqrt{n}, \end{aligned}$$

where the inequality is due to the AM–GM inequality (the inequality of arithmetic and geometric means) and every element of a probability vector is non-negative with a sum of 1.

□

(e) If P is a stochastic matrix, could it be $||P|| < 1$? What would this imply for our migration example if it were true?

Solution. Let $\mathcal{P}$ denote the space of probability vectors. Then if $\boldsymbol{X} \in \mathcal{P}$ then $P\boldsymbol{x} \in \mathcal{P}$. By applying this argument recursively we know that $P^t\boldsymbol{x} \in \mathcal{P}$. Furthermore, using part (d), $||P^t\boldsymbol{x}|| \geq ||x_*|| \geq a$ where a is a positive constant. Furthermore, $||\boldsymbol{x}|| \leq 1$ because all the entries add-up to one and are non-negative. Using the operator norm inequality if $||P|| < 1$ then $||P^t\boldsymbol{x}|| \leq ||P||^t \, ||\boldsymbol{x}|| \to 0$. However, this contradicts the fact that $||P^t\boldsymbol{x}|| \geq a > 0$ for all integer t and probability vector $||\boldsymbol{x}|| \in \mathcal{P}$. Therefore, it cannot be that $||P|| < 1$. In our migration example, an implication of $||P|| < 1$ would be that some people go to other states apart from 1 and 2, i.e., the size of population in city 1 and 2 is shrinking (which is why the norm converges to zero). However, the population is not shrinking, but just changing location.

□

2. In this section you will expand some of the details of the proof of the Cauchy-Schwarz inequality. Let $\lambda \in \mathbb{R}, \boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^n$. We know that if $\boldsymbol{z} = \boldsymbol{v} - \lambda\boldsymbol{x}$, $||\boldsymbol{z}|| \geq 0$, then

$$\boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x} \geq 0$$

(a) Show that the condition in Equation 1.2 is equivalent to:

$$\inf_{\lambda\in\mathbb{R}^n} \{ \ \boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x} \ \} \geq 0, \quad \forall \boldsymbol{v}, \boldsymbol{x} \in \mathbb{R}^n$$

Solution. ($\implies$) Taking the infimum to the left hand side of Equation 1.2 implies the infimum inequality.

($\impliedby$) Suppose (by contradiction) that there exists a $\lambda \in \mathbb{R}$ such that $\boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x} < 0$ for some λ. Then that means that this λ produces a value strictly lower than the infimum, a contradiction. □

(b) Consider the case when $||\boldsymbol{x}|| > 0$. Use the fact that the function is quadratic in λ to show that a minimum exists and that is

$$\frac{\boldsymbol{v}^t\boldsymbol{x}}{\boldsymbol{x}^t\boldsymbol{x}} = \arg\min_{\lambda\in\mathbb{R}^n}\{ \ \boldsymbol{v}^t\boldsymbol{v} - 2\lambda\boldsymbol{v}^t\boldsymbol{x} + \lambda^2\boldsymbol{x}^t\boldsymbol{x} \ \}$$

Solution. The first order condition with respect to λ yields $-2\boldsymbol{v}^t\boldsymbol{x} + 2\boldsymbol{x}^t\boldsymbol{x}\lambda = 0$ which yields the solution. The second-order condition is $\boldsymbol{x}^t\boldsymbol{x} > 0$ hence it is indeed

a minimum. □

(c) Show that if $\boldsymbol{v} = \boldsymbol{x}$, then Cauchy-Schwarz attains equality.

Solution. If $\boldsymbol{v} = \boldsymbol{x}$ then $||\boldsymbol{v}^t\boldsymbol{x}|| = |\boldsymbol{v}^t\boldsymbol{v}| = ||\boldsymbol{v}||^2 = ||\boldsymbol{v}||\ ||\boldsymbol{x}||$. □

17.2 Image and Kernel

1. Suppose that X is a non-zero $m \times n$ rank deficient matrix. Suppose that we partition its columns $X = [X_1, X_2]$ in such a way that $Im(X_1) = Im(X)$ and X_1 is full rank. The block matrices have n_1, n_2 columns, respectively. This is equivalent to dropping redundant variables in a linear regression.

   (a) Show that Equation 2.1 can be written in block-partitioned form as:

   $$\begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix} \beta = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   Solution. The transpose of X in block-partition form is $\begin{bmatrix} X_1^t \\ X_2^t \end{bmatrix}$. That means that

   $$X^t X = \begin{bmatrix} X_1^t \\ X_2^t \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix}$$

   Similarly we can show that

   $$X^t Y = \begin{bmatrix} X_1^t \\ X_2^t \end{bmatrix} Y = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   □

   (b) Suppose that $\hat{\beta}_1 = (X_1^t X_1)^{-1}(X_1^t Y)$. Construct a vector $\beta^* = \begin{bmatrix} \hat{\beta}_1 \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix}$. Show that β^* is a solution to Equation 2.1 if and only if $X_2^t X_1 \hat{\beta}_1 = X_2^t Y$.

   Solution. Write the system of equations in block partition form:

   $$\begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \mathbf{0}_{n_2 \times 1} \end{bmatrix} = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   We can expand the terms in each block.

   $$\begin{bmatrix} X_1^t X_1 \hat{\beta}_1 + X_1^t X_2 \mathbf{0}_{n_2 \times 1} \\ X_2^t X_1 \hat{\beta}_1 + X_2^t X_2 \mathbf{0}_{n_2 \times 1} \end{bmatrix} = \begin{bmatrix} X_1^t X_1 \hat{\beta}_1 \\ X_2^t X_1 \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   By construction $(X_1^t X_1)\hat{\beta}_1 = (X_1^t Y)$. Therefore the only condition we need to verify is $X_2^t X_1 \hat{\beta}_1 = X_2^t Y$.

□

(c) Verify that the columns of X_2 belong in $Im(X_1)$. Use this fact to show that $X_2^t X_1 \hat{\beta}_1 = X_2^t Y$.

Solution. The columns in X_2 belong in $Im(X)$ which is equal to $Im(X_1)$ by assumption. Let x_{2l} denote the l^{th} column of X_2 which is contained in $Im(X_1)$. Then for $l \in \{1, \ldots, k_2\}$ there exists a vector c_l such that $x_{2l} = X_1 c_l$. We can stack this result in matrix form as $X_2 = X_1 C$. That means that $X_2^t X_1 \hat{\beta}_1 = C^t X_1^t X_1 \hat{\beta}_1$. On the other hand, substituting the definition of the estimator,

$$X_2^t X_1 \hat{\beta}_1 = C^t X_1^t X_1 (X_1^t X_1)^{-1} X_1^t Y.$$

Some terms cancel out and the expression simplifies to $C^t X_1^t Y = (X_1 C)^t Y = X_2^t Y$. This completes the proof. □

(d) Consider the data matrix,

$$X = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$$

Construct $X^t X$ and $X^t Y$. Now partition the matrix into X_1, X_2 and compute β^*. Verify that the results that you proved above are true for the following cases:

(i) Construct X_1 using columns 1 and 2.

Solution.

$$X_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \quad X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad X = [X_1, X_2] \quad X^t X = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}, \quad X^t Y = \begin{bmatrix} 15 \\ 3 \\ 12 \end{bmatrix}$$

$$X_1^t X_1 = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, \quad X_1^t Y = \begin{bmatrix} 15 \\ 3 \end{bmatrix} \quad \beta^* = \begin{bmatrix} 4 \\ -2.5 \\ 0 \end{bmatrix} \quad X^t X \beta^* = \begin{bmatrix} 15 \\ 3 \\ 2 \end{bmatrix}$$

□

(ii) Construct X_1 using columns 1 and 3.

Solution.

$$X_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad X = [X_1, X_2] \quad X^tX = \begin{bmatrix} 5 & 3 & 2 \\ 3 & 3 & 0 \\ 2 & 0 & 2 \end{bmatrix}, \quad X^tY = \begin{bmatrix} 15 \\ 12 \\ 3 \end{bmatrix}$$

$$X_1^tX_1 = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}, \quad X_1^tY = \begin{bmatrix} 15 \\ 2 \end{bmatrix} \quad \beta^* = \begin{bmatrix} 1.5 \\ 2.5 \\ 0 \end{bmatrix} \quad X^tX\beta^* = \begin{bmatrix} 15 \\ 3 \\ 2 \end{bmatrix}$$

□

Notice that we follow the convention to write a partition such that $X = [X_1, X_2]$. In this case we select columns 1 and 3, so the matrices X_1, X_2 are different than before in (i).

(e) Is β^* the same in both exercises? How can we interpret the result?

Proof. Typically β^* does not produce the same result. This is an example where there are two mutually exclusive categorical variables and an intercept. For example, column 1 of X presents a constant term, column 2 could represent a binary indicator for whether the individual is female and column 3 could represent a binary indicator for male. The interpretation of the coefficient changes. If we drop the last column, the "reference category" is male. If we drop the second column, the "reference category" is female. However, both models have the same ability to describe the data without loss of information, because their columns have the same image. □

You can use the fact that the inverse of a 2×2 matrix is given by:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \implies A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

17.3 Orthogonality

1. In this exercise you will prove a version of the Frisch-Waugh-Lovell Theorem (Greene, 2012) in the detrending example.

   (a) Prove that $\tilde{\beta}_1 = (\boldsymbol{X}_1^t M_2 \boldsymbol{X}_1)^{-1}(\boldsymbol{X}_1^t M_2 \boldsymbol{Y})$.

   Solution. $\tilde{\beta}_1 = (\hat{U}_1^t \hat{U}_1)^{-1}(\hat{U}_1^t \hat{U}_Y)$. Where $\hat{U}_1 = M_2 X_1$ and $\hat{U}_Y = M_2 Y$. Thus we can rewrite the estimator as

   $$\begin{aligned}
   \tilde{\beta}_1 &= ((M_2 X_1)^t M_2 X_1)^{-1}((M_2 X_1)^t (M_2 Y) && \text{Plugging-in Expressions } \hat{U}_1 \text{ and } \hat{U}_Y. \\
   &= (X_1^t M_2^t M_2 X_1)^{-1}(X_1^t M_2 M_2 Y) && \text{Distributing Transpose} \\
   &= (X_1^t M_2 X_1)^{-1}(X_1^t M_2 Y) && \text{Using idempotency and symmetry of } M_2
   \end{aligned}$$

   □

   (b) Show that the system in Equation 3.1 can be written in block-partition form as:

   $$\begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   Solution. The transpose of X in block-partition form is $\begin{bmatrix} X_1^t \\ X_2^t \end{bmatrix}$. Therefore

   $$X^t X = \begin{bmatrix} X_1^t \\ X_2^t \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^t X_1 & X_1^t X_2 \\ X_2^t X_1 & X_2^t X_2 \end{bmatrix}$$

   Similarly we can show that

   $$X^t Y = \begin{bmatrix} X_1^t \\ X_2^t \end{bmatrix} Y = \begin{bmatrix} X_1^t Y \\ X_2^t Y \end{bmatrix}$$

   □

   (c) Show that second row can be rewritten as $\hat{\beta}_2 = (\boldsymbol{X}_2^t \boldsymbol{X}_2)^{-1}(\boldsymbol{X}_2^t \boldsymbol{Y} - \boldsymbol{X}_2^t \boldsymbol{X}_1 \hat{\beta}_1)$.

   Solution. The formula for the second row is $(X_2^t X_1)\hat{\beta}_1 + (X_2^t X_2)\hat{\beta}_2 = X_2^t Y$. We can solve this equation in terms of the second coefficient as $\hat{\beta}_2 = (X_2^t X_2)^{-1}(X_2^t Y - X_2^t X_2 \hat{\beta}_1)$. □

(d) Plug the above result into the first row of equations and show that $(\boldsymbol{X}_1^t M_2 \boldsymbol{X}_1)\hat{\beta}_1 = (\boldsymbol{X}_1^t M_2 \boldsymbol{Y})$. Conclude that $\hat{\beta}_1 = \tilde{\beta}_1$.

Solution. The equation in the first row is $X_1^t X_1 \hat{\beta}_1 + X_1^t X_2 \hat{\beta}_2 = X_1^t Y$. Plugging-in the result in part (c) we get that

$$\begin{aligned}
X_1^t X_1 \hat{\beta}_1 + X_1^t X_2 (X_2^t X_2)^{-1} X_2^t Y - X_1^t X_2 (X_2^t X_2)^{-1} X_2^t X_1 \hat{\beta}_1 &= X_1^t Y \\
X_1^t X_1 \hat{\beta}_1 + X_1^t P_2 Y - X_1^t P_2^t X_1 \hat{\beta}_1 &= X_1^t Y && \text{(Definition } P2\text{)} \\
X_1^t (I - P_2) X_1 \hat{\beta}_1 &= X_1^t (I - P_2) Y && \text{(Grouping terms)} \\
X_1^t M_2 X_1 \hat{\beta}_1 &= X_1^t M_2 Y && \text{(Definition } M_2\text{.)}
\end{aligned}$$

□

2. In the detrending example:

(a) Show that $\boldsymbol{X}$ full rank implies that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are full rank. (Hint: Prove by contradiction)

Solution. (By contradiction) suppose that X_1 or X_2 are not full rank. Suppose WLOG that it is X_2. Then by Corollary 2.2.1 we can write one of the columns as a linear combination of the other columns in X_2. However, this implies that one of the columns of X can be written as a linear combination of other columns in X, implying that X is not full rank. This is a contradiction. □

(b) Define $B = M_2 \boldsymbol{X}_1$. Show that replacing $\boldsymbol{X}_1$ with the matrix B does not change the image, i.e. $Im(\boldsymbol{X}_1, \boldsymbol{X}_2) = Im(B, \boldsymbol{X}_2)$. (Hint: Modify Lemma 3.1.1)

Solution. First rewrite $B = M_2 X_1 = X_1 - X_2 (X_2^t X_2)^{-1} X_2^t X_1$ and define $\Theta := (X_2^t X_2)^{-1} X_2^t X_1$, which is a $k_2 \times k_1$ vector. Then $B = X_1 - X_2 \Theta$.

(i) $Im(B, X_2) \subseteq Im(X_1, X_2)$. Suppose that $z \in Im(B, X_2)$. Then there exists a vector $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, where $\beta_1 \in \mathbb{R}^{k_1}$ and $\beta_2 \in \mathbb{R}^{k_2}$ such that $z = [B, X_2]\beta$. This can be decomposed as $B\beta_1 + X_2\beta_2$, which is equal to $(X_1 - X_2\Theta)\beta_1 + X_2\beta_2$ and can be written in the form $X_1\beta_1 + X_2(-\Theta\beta_1 + \beta_2)$. Therefore, $x \in Im(X_1, X_2)$.

(ii) $Im(X_1, X_2) \subseteq Im(B, X_2)$. Suppose that $z \in Im(X_1, X_2)$. Then there exists a vector $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, where $\beta_1 \in \mathbb{R}^{k_1}$ and $\beta_2 \in \mathbb{R}^{k_2}$ such that $z = [X_1, X_2]\beta$. This can be decomposed as $X_1\beta_1 + X_2\beta_2$, which is equal to $(B + X_2\Theta)\beta_1 + X_2\beta_2$ and can be written in the form $B\beta_1 + X_2(\Theta\beta_1 + \beta_2)$. Therefore, $x \in Im(B, X_2)$.

□

(c) Show that if $\boldsymbol{X}$ is full rank then $(\boldsymbol{X}_1^t M_2 \boldsymbol{X}_1)$ is full rank. (Hint: Review Linear Regression Section)

Proof. The matrix can be rewritten as $X_1^t M_2 X_1 = (X_1^t M_2^t M_2 X_1)$ because M_2 idempotent and symmetric implies that $M_2 = M_2^t M_2$. Therefore the equation can be written as $(M_2 X_1)^t (M_2 X_1)$. By Lemma 2.4.1, the gram matrix is full rank if and only $B = M_2 X_1$ is full rank.

Now let us show that $B = M_2 X_1$ is indeed full rank. Suppose not. Then there exists some nonzero vector $c_1 \neq 0$ such that

$$(I - X_2(X_2^t X_2)^{-1} X_2^t) X_1 c_1 = 0.$$

Define $c_2 = (X_2^t X_2)^{-1} X_2^t X_1 c_1$, and hence the above equation could be rewritten as

$$X_1 c_1 - X_2 c_2 = 0.$$

Since (c_1, c_2) is a nonzero vector, this contradicts the condition that $\boldsymbol{X}$ is full rank. □

17.4 Convex Sets (I): Hyperplanes

1. For any $p \in \mathbb{R}^n \backslash \{0\}$ and $a \in \mathbb{R}$, let

$$h(p,a) \equiv \left\{x \in \mathbb{R}^n | p^T x \geq a\right\}$$

be the half space generated by the hyperplane $H(p,a)$. Assume D is a closed subset of $\mathbb{R}^n$. Let E be the intersection of all half spaces that contain D, i.e.

$$E \equiv \bigcap_{h(p,a) \supseteq D} h(p,a).$$

Prove D is convex if and only if $D = E$. This gives another characterization of convexity. (Hint: separating hyperplane theorem.)

Solution. If $D = E$, then D is clearly convex because each half space in the intersection is convex.

Now assume D is convex. Because D is contained in each of the half space in the intersection, $D \subseteq E$. Assume there is $x \in E$ but $x \notin D$. Then because D is convex and closed, there exists a hyperplane $H(p^*, a^*)$ that strictly separates x and D:

$$p^{*T} d > a^* > p^{*t} x \quad \forall d \in D.$$

The first inequality implies $D \subseteq h(p^*, a^*)$, implying $E \subseteq h(p^*, a^*)$. But $x \in E$ implies $p^{*T} x \geq a^*$, a contradiction. Hence $E = D$. □

2. Assume $U \subset \mathbb{R}^n$ is convex. Let $x^* \in U$ be a point. Prove the followings are equivalent:

   (a) there is no $x \in U$ such that $x_i > x_i^*$ for all $i = 1, \cdots, n$,

   (b) there exists $\lambda \in \mathbb{R}^n_+ \backslash \{0\}$ such that x^* solves

$$\max_{x \in U} \lambda^T x.$$

Solution. (a)$\Longrightarrow$(b): Define $W \equiv \left\{x \in \mathbb{R}^n \middle| x_i > x_i^* \ \forall i = 1, \cdots, n\right\}$. The set W is nonempty and convex, and $W \cap U = \emptyset$ by assumption. Then by supporting hyperplane theorem, there exists $\lambda \in \mathbb{R}^n \backslash \{0\}$ and real number c such that

$$\lambda^T y \geq c \geq \lambda^T x, \quad \forall y \in W, \ x \in U.$$

Because x^* is a limit point of W by construction, clearly $\lambda^T x^* \geq c \geq \lambda^T x$ for all $x \in U$. It remains to show $\lambda \geq 0$. Assume $\lambda_i < 0$ for some i. For any arbitrary $\tilde{y} \in W$, $\lambda^T(\tilde{y} + ne_i)$ tends to $-\infty$ as $n \to \infty$, where e_i is the ith unit vector in $\mathbb{R}^n$. But $\tilde{y} + ne_i \in W$ for all n, contradicting $\lambda^T y \geq c$ for all $y \in W$.

(b)$\Longrightarrow$(a): Suppose there exists $\lambda \in \mathbb{R}^n_+\{0\}$ such that $\lambda^T x^* \geq \lambda^T x$ for all $x \in U$. Pick any $x \in \mathbb{R}^n$ satisfying $x_i > x_i^*$ for all i. Because $\lambda_i \geq 0$ and $\lambda \neq 0$, we have $\lambda^T x > \lambda^T x^*$. Thus such x is not in U. □

3. Let D be a nonempty convex subset of $\mathbb{R}^n$. Prove its closure $\overline{D}$ is convex.

Solution. Pick any $x, x' \in \overline{D}$. There must exist sequences $\{x_n\} \subset D$ and $\{x'_n\} \subset D$ such that $x_n \to x$ and $x'_n \to x'$ (if $x \in D$, then let $x_n = x$). So $\lambda x_n + (1-\lambda)x'_n \in D$ for all $\lambda \in [0,1]$. Because $\lambda x_n + (1-\lambda)x'_n$ converges to $\lambda x + (1-\lambda)x'$, $\lambda x + (1-\lambda)x'$ is a point in $\overline{D}$. □

17.5 Convex Sets (II): Cones

1. There are several different characterizations of Farkas' Lemma. For example

   Lemma 17.5.1 (Farkas' Lemma V2)**.** *Let A be an $m \times n$ matrix and $b \in \mathbb{R}^m$. Then one and only one is true:*

   (i) *There exists $x \in \mathbb{R}^n$ such that $Ax \leq b$.*

   (ii) *There exists $y \in \mathbb{R}^m$ such that $y \geq \mathbf{0}_{m\times 1},\ y^tA = \mathbf{0}_{1\times n},\ y^tb < 0$.*

   In this exercise, you will prove the lemma.

   (a) Define $C = [A, -A, I_{m\times m}] \in \mathbb{R}^m \times \mathbb{R}^{2n+m}$. Show that condition (i) is equivalent to $b \in Cone(C)$ (Hint: Use properties of block-partitioned matrices and define a vector $z \in \mathbb{R}^{2n+m}_+$).

   Solution. Before we proceed with the proof we will analyze an object in $Cone(C)$. The vector $z \in Cone(C)$ if there exists a vector $\lambda \in \mathbb{R}^{2n+m}_+$ such that $z = C\lambda$. In block-partition form this means that:

   $$z = \begin{bmatrix} A & -A & I \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = A\lambda_1 - A\lambda_2 + \lambda_3 = A(\lambda_1 - \lambda_2) + \lambda_3$$

   ($\Longrightarrow$) We show that condition (i) implies that $b \in Cone(C)$. Suppose that there exists an $x \in \mathbb{R}^n$ such that $Ax \leq b$. Set $\lambda_3 = b - Ax \geq 0$. For each entry of x_j set $\lambda_{1j} = x_j$ if $x_j \geq 0$ and zero otherwise. Similarly, set $\lambda_{2j} = -x_j$ if $x_j < 0$ and zero otherwise. Then $x = \lambda_1 - \lambda_2$ and $\lambda_1, \lambda_2 \in \mathbb{R}^+_n$. Therefore, $b \in Cone(A)$.
   ($\Longleftarrow$) If $b \in Cone(C)$, there exists λ^{2n+m}_+ such that $b = C\lambda$. Set $x = \lambda_1 - \lambda_2 \in \mathbb{R}^n$. By definition $b = Ax + \lambda_3 \geq Ax$ since $\lambda_3 \geq 0_{m\times 1}$.

   □

   (b) Show that condition (ii) is equivalent to: There exists $y \in \mathbb{R}^m$ such that $y^tC \geq \mathbf{0}_{1\times(2n+m)}$ and $y^tb < 0$.

   Proof. Before we show the equivalence, let us express the y^tC in block form.

   $$y^tC = y^t \begin{bmatrix} A & -A & I \end{bmatrix} \geq 0 \quad \Longleftrightarrow \quad \begin{array}{c} y^tA \geq \mathbf{0}_{1\times n} \\ -y^tA \geq \mathbf{0}_{1\times n} \\ y^t \geq \mathbf{0}_{1\times m} \end{array}$$

Combining the two inequalities gives us $y^t A = \mathbf{0}_{1\times n}$ and $y^t \geq \mathbf{0}_{1\times m}$. We can transpose y^t to show that the two conditions are identical.

□

(c) Use the original Farkas' Lemma to prove (Version 2).

Solution. Apply Farkas' Lemma with the matrix C. Then either the statement in question (a) occurs or the statement in question (b). The proof is completed between these statements and those of the lemma that we want to prove. □

2. Consider an alternative restriction on asset prices.

Definition 17.5.1 (Pricing Restrictions)**.** Suppose that there does not exist an $x \in \mathbb{R}^n$ such that ($q^t x \leq 0$ and $Rx > \mathbf{0}_{m\times 1}$) or such that ($q^t x < 0$ and $Rx \geq \mathbf{0}_{m\times 1}$).

(a) Write down an economic interpretation of this condition.

Solution. It says that a market is arbitrage free if an investor cannot purchase a portfolio at (1) zero cost or lower and obtain a positive return in at least one state, or (2) get paid for the assets (negative costs) and receive a non-negative return. □

(b) Suppose that there exists a set of portfolio weights $x \in \mathbb{R}^n$ that yield positive returns in every state ($\Pi x \gg 0$). Show that $Rx > \mathbf{1}_{m\times 1} q^t x$. Give a simple example of a return matrix R, a price vector q and a portfolio x where this holds but the conditions in Definition 5.5.1 does not hold.

Proof. By definition, the expected profit from a portfolio is $\Pi x = Rx - \mathbf{1}_{n\times 1} q^t x$. Then $\Pi x \gg 0$ implies that $Rx \gg \mathbf{1}_{n\times 1} q^t x$. Consequently, $Rx > \mathbf{1}_{n\times 1} q^t x$. Let $n = m = 1$, and $q = 1$ and $R = 2$. Then $x = 1$ ensures that $\Pi x \gg 0$. However, $q^t x > 1$ and $Rx > 0$. □

(c) Suppose that there exists a probability vector $\alpha \in \mathbb{R}^m$ with **strictly positive** entries which satisfies $\alpha^t \Pi = \mathbf{0}_{1\times n}$. Show that Definition 5.5.1 is satisfied.

Proof. If there exists a vector $\alpha \in \mathbb{R}^m$ with strictly positive probabilities such that $\alpha^t \Pi = \mathbf{0}_{1\times n}$ then $\alpha^t Rx = \alpha^t \mathbf{1}_{m\times 1} q^t x$. Suppose that (i) $Rx > 0$ and $q^t x \leq 0$, then since $\alpha \gg 0$ then $\alpha^t Rx > 0$ and $\alpha^t \mathbf{1} q^t x \leq 0$. On the other hand if (ii) $Rx \geq 0$ and $q^t x < 0$ then $\alpha^t Rx \geq 0$ and $\alpha^t \mathbf{1} q^t x < 0$. This a contradiction because we should have $\alpha^t Rx = \alpha^t \mathbf{1}_{m\times 1} q^t x$. □

17.6 Quadratic Forms

1. Let A be an $n \times n$ square matrix. Assume:

$$x^T A x = 0, \qquad \forall x \in \mathbb{R}^n.$$

(a) Prove all diagonal components of A are $0 \in \mathbb{R}$.

Solution. Let $x = e_i$ be the i-th unit vector in $\mathbb{R}^n$. Then $a_{ii} = e_i^T A e_i = 0$. □

(b) Show by example that condition (6.1) does not imply $A = \mathbf{0}$.

Solution. For example

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

works. □

17.7 Determinants

A matrix $B_{n\times n}$ is positive definite if $\forall x \in \mathbb{R}^n, x^T Bx > 0$. An equivalent definition of positive definiteness can be formulated using the determinant:

$$B = \begin{bmatrix} b_{11} & ... & b_{n1} \\ ... & ... & ... \\ b_{n1} & ... & b_{nn} \end{bmatrix}$$

Define the leading principal minor k of B, as the matrix formed by taking the upper left $(k \times k)$ submatrix. In other words:

$$B_1 = \begin{bmatrix} b_{11} \end{bmatrix}, B_2 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, ..., B_n = \begin{bmatrix} b_{11} & ... & b_{n1} \\ ... & ... & ... \\ b_{n1} & ... & b_{nn} \end{bmatrix}$$

A matrix is positive definite if and only if $\forall i \in \{1, ..., n\}, \det(B_i) > 0$. (Take this as a given, you do not need to prove it).

1. Define a function $F : \mathcal{M}_{n\times n} \to \mathbb{R}^n$. $F(B) = (\det(B_1), ..., \det(B_n))$. Reformulate the definition of positive definiteness in terms of $F(B)$.

   Solution. The condition is: A matrix is positive definite if and only if $F(B) \in \mathbb{R}^n_{++}$.

   Remark As a reminder, the set $\mathbb{R}^n_{++}$ is the set in $\mathbb{R}^n$ that has strictly positive components for all dimensions.

   □

2. Define a metric for the distance between two matrices, $d(A, B)$. Show that it is a metric: that it is non-negative, symmetric and satisfied the triangle inequality.

   Solution. Let $vec(A), vec(B)$ be the vectorized versions of the matrices (A, B). Then let us define the distance between two matrices as:

   $$d(A, B) = |vec(B) - vec(A)|_{\mathbb{R}^{mn}}$$

   $$d(A, B) = \sqrt{(a_{11} - b_{11})^2 + ...(a_{n1} - b_{n1})^2 + ... + (a_{mn} - b_{mn})^2}$$

   where $|\cdot|_{\mathbb{R}^{mn}}$ is the vector norm in $\mathbb{R}^{mn}$. This metric satisfies the three properties of a metric (because the vector metric is a proper metric):

(a) It is non-negative and $A = B$ iff $d(A, B) = 0$.

(b) It is symmetric. $d(A, B) = d(B, A)$.

(c) It satisties the triangle inequality :

$$d(A, C) \leq d(A, B) + d(B, C)$$

□

3. Show that the function $F(B)$ is continuous.

Solution. Let $F_i(B)$ be the i^{th} coordinate of $F(B)$. A vector valued function is continuous if and only if all of its components are continuous functions. Therefore we only need to prove that $F_i(B)$ is continuous $\forall i \in \{1, ..., n\}$.

$F_i(B) = \det(h_i(B)) = \det(B_i)$, where $h_i(B)$ is a functions that selects the submatrix B_i. We discussed in class that the determinant is a continuous function because it is essentially a polynomial of the components of a matrix, and polynomial functions are always continuous. Furthermore $h_i(B)$ is also a continuous function (it only selects elements from B). Therefore the composite function $F_i(B)$ is also continuous.

Remark Continuity has to be defined within a metric space. We can choose the metric we selected in part (b).

□

4. Show that the set of positive definite matrices of size (n) is an open set in $\mathcal{M}_{n\times n}$.

Remark This shows that under small perturbations in the components of a positive definite matrix, the resulting matrix preserves the property of positive definiteness.

Solution. One definition of continuity that is very useful in the case is that a function is continuous if and only if the pre-image of an open set is also an open set in the domain. In this cases a matrix is positive definite if $F(B) \in \mathbb{R}^n_{++}$. The set $\mathbb{R}^n_{++}$ is an open set. Therefore, the set of matrices that satisfy the condition is also an open set.

We can also derive the proof using $\epsilon - \delta$ arguments. Suppose that a matrix is positive definite, then $F(B) \in \mathbb{R}^n_{++}$. There exists an $\epsilon > 0$ such that the all values in an open

ball around $F(B)$ also belong to $\mathbb{R}^n_{++}$. By the definition of continuity $\exists \delta > 0$ such that $\forall B' s.t. d(B, B') < \delta \implies d(F(B), d(F(B')) < \epsilon$. This means that a neighborhood around B is also positive definite. Thus the set of positive definite matrices is an open set.

□

17.8 Eigenvalues and Eigenvectors

1. (25 points) Let P be an $n \times n$ matrix.

   (a) (5 points) Define a *markov matrix* P as an $n \times n$ matrix that has non-negative entries where the entries of each column sum to one. Let π be a non-negative vector whose entries sum to one. Show that π does not belong to the kernel. Further show that $P\pi$ is a vector whose entries sum to one.

   Solution.

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{12} & p_{22} & \cdots & p_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}, \pi = \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_n \end{bmatrix}$$

   Then $P\pi$ can be written as:

$$P\pi = \pi_1 \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{bmatrix} + \cdots + \pi_n \begin{bmatrix} p_{1n} \\ p_{2n} \\ \vdots \\ p_{nn} \end{bmatrix} = \sum_i \pi_j p_j := \text{ A linear combination of the columns}$$

   Therefore, the sum of the entries on $P\pi$ is:

$$\sum_i \sum_j \pi_j p_{ij} = \sum_j \pi_j \sum_i p_{ij} = \sum_j \pi_j = 1$$

   The second equality changes the order of the sum. The third equality uses the fact that the elements of each column of P sum to 1. The fourth equality uses the fact that the entries of π_i sum to one.

   π belongs to the kernel $\iff P\pi = 0$. However, since its entries of $P\pi$ sum to one, $P\pi \neq 0$. Therefore, π is not part of the kernel.

   Using the fact that π is non-negative is not necessary to prove the above properties. However, it implies that $P\pi$ has non negative entries.

$$
\begin{aligned}
p_{ij} &\geq 0 && \forall i,j \in \{1,...,n\} \\
\implies \pi_j p_{ij} &\geq 0 && \text{since } \pi_j \geq 0, \forall j \\
\implies \sum_j \pi_j p_{ij} &\geq 0 && \text{which is the } i^{th} \text{ entry of } P\pi
\end{aligned}
$$

□

(b) (3 points) Now suppose that $\lim_{m\to\infty} P^m \to P^*$. Show that P^* is also a markov matrix and show that π does not belong to its kernel. (Hint: Show that every P^n is markov).

Solution. First we will show that P^m is markov. We will do this by induction. For $m = 2$:

$$
P^2 = PP = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ Pp_1 & \cdots & Pp_n \\ \downarrow & \cdots & \downarrow \end{bmatrix}
$$

Notice that the columns of P are non-negative vectors that sum to one. Therefore, Pp_i is a non-negative vector whose entries sum to one in each column, by the proof in the previous exercise. For $n > 2$:

$$
P^m = PP^{m-1}
$$

Since P^m is markov, its columns are non-negative entries whose entries sum to one in each column, it follows that P^m is also markov. Let p_{ijm} denote the i, j entry of P^m.

We can summarize the set of conditions that define a markov matrix for a matrix P^m:

$$
p_{ijm} \geq 0, \ \forall i,j \in \{1,...,n\}
$$

$$
\sum_j p_{ijm} = 1, \ \forall i,j \in \{1,...,n\}
$$

Notice that if $p_{ijm} \to p^*_{ij}$, then it still satisfies the first weak inequality, and the second equality.This means that the set of markov matrices is closed. Therefore, in the limit it still satisfies the restrictions of a markov matrix. This completes the proof of why P^* is markov.

□

(c) (2 points) Show that if P is symmetric, then P^* is symmetric.

Solution. Firs we will show that P^m is symmetric. We will prove this by induction. For $m = 2$:

$$(PP)^t = P^t P^t = P$$

For $n > 2$: $(P^m)^t = (PP^{m-1})^t = (P^{(m-1)})^t P^t = P^{(m-1)} P = P^m$. This shows that P^m is symmetric. Notice that a matrix is symmetric iff:

$$p_{ijm} - p_{jim} = 0 \ \forall i, j \in \{1, ..., n\}$$

If $p_{ijm} \to p^*_{ij}$, it will still satisfy this equality b

□

(d) (5 points) Suppose that P^* is such that for every π, $P^*\pi = \pi^*$, for a fixed π^*. Write down what the matrix P^* has to be for $\pi^* = (0.2, 0.3, 0.4, 0.1)$ if P^* is 4×4.

Solution. We will use the elementary basis to construct P^* :

$$P^* = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ Pe_1 & \cdots & Pe_n \\ \downarrow & \cdots & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ \pi^* & \cdots & \pi^* \\ \downarrow & \cdots & \downarrow \end{bmatrix}$$

This is a matrix with identical column vectors π^*. Since $P\pi$ is just a linear combination of the columns, with weights adding to one, then the resulting vector is just π^*, as desired. For the example 4×4 example suggested:

$$P^* = \begin{bmatrix} 0.2 & 0.2 & 0.2 & 0.2 \\ 0.3 & 0.3 & 0.3 & 0.3 \\ 0.4 & 0.4 & 0.4 & 0.4 \\ 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}$$

□

(e) (5 points) Under the previous property, for what set of vectors π^* will the implied P^* be symmetric. If it is symmetric, is it idempotent? If so, what is its rank? (Note that if P is symmetric, it implies very special restrictions on what P should converge to).

Solution. In the previous questions we established that under the previous property, $p^*_{ij} = \pi^*_i, \forall i, j \in \{1, ..., n\}$ (all columns are identical to π^*).

On the other hand, symmetry implies that $p^*_{ij} = p^*_{ji}$. Suppose that we take the first column: $\pi^*_i = p^*_{i1} = p^*_{1i} = \pi^*_1, \forall i \{1, ..., n\}$. Therefore all the entries of π^* are identical and equal to $1/n$ because they have to add up to 1. For the 4×4 case this means:

$$P^* = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix}$$

□

(f) (2 points) Construct an example of a 2×2 symmetric matrix P that doesn't converge. (Hint use zeros and ones only). Compute its eigenvalues. Use the spectral decomposition to give a reason why it doesn't converge.

Solution. An example of a 2×2 matrix that doesn't converge is:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The matrix P changes the order of the columns. It can be shown by induction that:

$$P^m = \begin{cases} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & m \text{ odd} \\ \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} & m \text{ even} \end{cases}$$

To compute its eigenvalues we need to compute the roots of:

$$\det(P - \lambda I) = \det \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 - 1 = 0$$

Then the roots are: $\lambda_1 = 1$, $\lambda_2 = -1$. Now we need to compute the eigenvectors

for each eigenvalue, respectively:

$$(P-I)v_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} v_1 = 0 \implies v_1 \in span\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\}$$

$$(P+I)v_2 = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} v_2 = 0 \implies v_2 \in span\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\}$$

Therefore we can construct a spectral decomposition of P. Notice that first we have to obtain orthogonal vectors from each span: $\tilde{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \tilde{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$.

$$\begin{aligned} P &= \begin{bmatrix} \uparrow & \uparrow \\ \tilde{v}_1 & \tilde{v}_2 \\ \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \leftarrow & \tilde{v}_1^t & \rightarrow \\ \leftarrow & \tilde{v}_2^t & \rightarrow \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \qquad C\Lambda C^t \end{aligned}$$

We can verify that this decomposition recovers the original matrix P. We can now use this decomposition to compute P^m:

$$\begin{aligned} P^m &= PP\cdots P \\ &= (C\Lambda C^t)(C\Lambda C^t)\cdots(C\Lambda C^t) \\ &= C\Lambda^m C^t \\ &= C \begin{bmatrix} (1)^m & 0 \\ 0 & (-1)^m \end{bmatrix} C^t \end{aligned}$$

The second to third line follow from the fact that $C^tC = I$, since the vectors are orthonormal. This means λ^m oscillates between -1 and 1, depending on m, and thus never converges. The example shows that symmetry does not guarantee convergence.

General remarks for other types of exercises: Notice that if one of the eigenvalues were strictly less than one in absolute value then λ^m would converge to zero. However, at least one of them has to be greater than or equal to zero, otherwise $\Lambda^m \to 0$ and P^* is not a markov matrix (which contradicts what we proved earlier). If $|\lambda| > 1$ then the values would be explosive and diverge to

infinity.

There are many more results for the eigenvalues of markov matrices that are not symmetric, even those that don't have a spectral decomposition. The key thing is to prove which and how many many eigenvalues are strictly less than one, equal to one and strictly greater than one. I encourage you to keep these concepts in mind in future work involving markov chains.

□

(g) (3 points) Show that the following asymmetric P converges to a P^* such that $P^*\pi = \pi^*$. Compute P^* and π^*.

$$P = \begin{bmatrix} 0.5 & 0 \\ 0.5 & 1 \end{bmatrix}$$

Solution. I will prove that $P^m = \begin{bmatrix} (0.5)^m & 0 \\ 1-(0.5)^m & 1 \end{bmatrix}$ by using induction. The result holds trivially for $m = 1$, then for $m > 1$:

$$\begin{aligned} P^m &= PP^{m-1} \\ &= \begin{bmatrix} 0.5 & 0 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} (0.5)^m & 0 \\ 1-(0.5)^m & 1 \end{bmatrix} \\ &= \begin{bmatrix} (0.5)^{(m+1)} & 0 \\ (0.5)(0.5)^m + (1-(0.5)^m) & 1 \end{bmatrix} \\ &= \begin{bmatrix} (0.5)^{(m+1)} & 0 \\ 1-(0.5)^{(m+1)} & 1 \end{bmatrix} \end{aligned}$$

Then $P^* = \lim_{n\to\infty} P^m = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ and $\pi^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. This means that regardless of the initial vector π, $P^m\pi$ will converge to π^*. This highlights that both symmetric and non-symmetric matrices can converge. □

2. This questions asks you to analyze the eigenvalues of stochastic matrices:

(a) (3 points) Let $v \in \mathbb{R}^n$. Show that the entries of the vector Pv add up to $\sum_{j=1}^n v_j$.

Solution. Let $1_n \in \mathbb{R}^n$ be a vector with only 1s in each entry. Then the sum of the entries of Pv can be represented as $1_n^t Pv$. This is a quadratic form that can be represented as a double sum.

$$\begin{aligned} 1_n^t Pv &= \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} v_j \\ &= \sum_{j=1}^{n} v_j \sum_{i=1}^{n} P_{ij} \\ &= \sum_{j=1}^{n} v_j \end{aligned}$$

□

(b) (9 points) Let $v^* \in \mathbb{R}^n, v \neq 0$ be an eigenvector of P, with corresponding eigenvalue λ. Prove the following statements:

i. (1 point) $P^s v^* = \lambda^s v^*, s \in \mathbb{N}$.

Solution. We can prove this by induction. For $s = 1$, by definition of an eigenvector. $P^1 v^* = Pv^* = \lambda^1 v^*$.
Suppose it holds for s. Then $P^{s+1} v^* = P(P^s v^*) = P(\lambda^n) v^* = \lambda^s P v^* = \lambda^s \lambda v^* = \lambda^{s+1} v^*$.

□

ii. (4 points) Show that if $\sum_{j=1}^{n} v_j^* \neq 0$, then $\lambda = 1$. [Hint: show that P^s is also markov].

Solution. P is a matrix whose columns $(p_1, ..., p_n)$ sum to one. Let P' be another markov matrix. Then PP' is a matrix with columns $(Pp'_1, ..., Pp'_n)$. By the result in part (a) the columns of PP' must sum to one. Furthermore, since P and P' have non-negative entries, PP' has to have non-negative entries. Now we can show that P^s is markov by induction. For $s = 2$, if $P = P'$ then $PP' = P^2$, which is markov. Now suppose that it holds for s. Then $P' = P^s$. Then $P^{s+1} = PP^s = PP'$ which is also markov.

Therefore, by the result in part (a) the entries of $P^s v^*$ have to sum up to $\sum_{j=1}^{n} v_j^*$ for all s. From the result in part (b)(i) we know that $P^s v^* = \lambda^s v^*$. Therefore the entries sum up to $\lambda^s \sum_{j=1}^{n} v_j^*$. This means that $\lambda^s = 1, \forall s \implies$

$\lambda = 1$.

Additional result: It is also possible to show that there exists at least one vector with eigenvalue 1. Since $det(B) = det(B^t), \forall B$ then $det(P - \lambda I) = det(P^t - \lambda I)$. This means that the eigenvalues of P and P^t are the same (although the eigenvectors can be different). Since the rows of P^t sum to one, it can be shown that 1_n is an eigenvector of P^t with eigenvalue $\lambda = 1$. Therefore, P has at least one eigenvector with $\lambda = 1$, which is not necessarily 1_n.

□

iii. (4 points) Show that if $\sum_{j=1}^{n} v_j^* = 0, v^* \neq 0$, then $|\lambda| \leq 1$. [Hint: show that for any fixed $v \neq 0$ (not necessarily an eigenvector), $\sup_P ||Pv|| \leq M < \infty$, P markov].

Solution. Let $w = Pv$. Notice that $||Pv|| = \sqrt{\sum_{i=1}^{n} w_i^2} = \sqrt{\sum_{i=1}^{n} (\sum_{j=1}^{n} P_{ij} v_j)^2}$. This is a continuous function of the P_{ij}. Suppose that we represent P as $vec(P) \in \mathbb{R}^{n^2}$. If P is markov, then each entry is bounded $P_{ij} \in [0, 1]$ and $\sum_{i=1}^{n} P_{ij} = 1, \forall j$, which is a closed set. Then for fixed v the norm $||Pv||$ is a continuous function from a compact space in $\mathbb{R}^{n^2}$ (the set of markov matrices) into $\mathbb{R}$. By the maximum theorem, there exists a markov matrix P^* such that $||P^* v|| = \max_P ||Pv|| = \sup_P ||Pv|| = M < \infty$. Consequently $||Pv|| \leq M$, for all P markov.

Notice that P^n is also markov, therefore $||P^n v^*|| \leq M$. By part (b)(i), this implies that $||\lambda^n v^*|| = |\lambda|^n ||v^*|| \leq M, \forall n$. Since $v^* \neq 0$, $||v^*|| > 0$ and $|\lambda|^n \leq \frac{M}{||v^*||}, \forall n$. If $|\lambda| > 1$, there exists an n large enough that $|\lambda|^n > \frac{M}{||v||}$. This is a contradiction, therefore $|\lambda| \leq 1$.

□

17.9 Introduction to Differentiation

1. Let $f(x) = \begin{cases} x^\alpha \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$. For what values of α is $f(x)$ differentiable at $x = 0$?

   Solution. The function is not well defined some some non-integer values of α. For example, if $\alpha = 0.5$, $x^{\alpha-1} = 1/\sqrt{x}$. Therefore, I will restrict this proof to integer values of α.

   $$S(x) = \frac{f(x) - f(0)}{x - 0} = \frac{x^\alpha \sin(1/x)}{x} = x^{\alpha-1} \sin(1/x)$$

   - If $\alpha = 1$, $S(x) = \sin(1/x)$, which oscillates around for x close to 0.
   - If $\alpha < 1$ then $x^{\alpha-1}$ is not defined for some values of α. For example the sequence $x_n = 1/(2\pi n)$ has the property that $S(x_n) = 0$ and if $x_n = 1/(2\pi n + (\pi/2))$, then $S(x_n) = (2\pi n + (\pi/2))^{\alpha-1} \to \infty$. Therefore, it doesn't satisfy the sequential definition of convergence.
   - If $\alpha > 1$ it does converge:

     $$-x^{\alpha-1} \leq x^{\alpha-1} \sin(1/x) \leq x^{\alpha-1}$$

     Since $\lim_{x\to 0} x^{\alpha-1} = 0$, then $\lim_{x\to 0} S(x) = 0$. Then the derivative exists for integer values of α strictly greater than one but not for other integer values of α.

   □

2. Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions. Let $y_0 = g(x_0)$ for some $x_0 \in \mathbb{R}$. Find examples for the following cases when:

   (a) g is differentiable at x_0 and f is not differentiable at y_0;

   (b) g is not differentiable at x_0 and f is differentiable at y_0;

   (c) g is not differentiable at x_0 and f is not differentiable at y_0,

   but $f \circ g(x)$ is differentiable.

   Solution. (a) Consider $f(y) = |y|$, $g(x) = x^2$. Consider $x_0 = 0$ and $y_0 = 0$. Then $f \circ g(x) \equiv x^2$, hence differentiable at x_0.

   (b) Consider $f(y) = y^2$, $g(x) = |x|$ and $x_0 = 0$.

(c) Consider

$$f(x) = g(x) = \begin{cases} \frac{1}{x} & x \neq 0, \\ 0 & x = 0. \end{cases}$$

Then neither f nor g is continuous at 0. But $f \circ g(x) \equiv x$ which is differentiable.

□

3. (Exercise 11 on page 186, Pugh) Assume that $f : (-1, 1) \to \mathbb{R}$ and $f'(0)$ exists. If $\alpha_n, \beta_n \to 0$ as $n \to \infty$, define the different quotient

$$D_n = \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n}.$$

(a) Prove that $\lim_{n\to\infty} D_n = f'(0)$ under each of the following conditions (Hint: First rewrite this expression in terms of $\frac{f(\beta_n)-f(0)}{\beta_n}$ and $\frac{f(\alpha_n)-f(0)}{\alpha_n}$ and use the sequential definition of the limit.

i. $\alpha_n < 0 < \beta_n$.

Solution. Rewrite

$$\begin{aligned} D_n &= \frac{f(\beta_n) - f(0)}{\beta_n}\frac{\beta_n}{\beta_n - \alpha_n} + \frac{f(\alpha_n) - f(0)}{\alpha_n}\frac{-\alpha_n}{\beta_n - \alpha_n} \\ &= \frac{f(\beta_n) - f(0)}{\beta_n} + \Big(\frac{f(\alpha_n) - f(0)}{\alpha_n} - \frac{f(\beta_n) - f(0)}{\beta_n}\Big)\frac{-\alpha_n}{\beta_n - \alpha_n}. \end{aligned}$$

Because $0 \leq \frac{-\alpha_n}{\beta_n - \alpha_n} \leq 1$, as $n \to \infty$, the right hand side tends to $f'(0)$. □

ii. $0 < \alpha_n < \beta_n$ and $\frac{\beta_n}{\beta_n - \alpha_n} \leq M$.

Solution. The proof is similar to previous one. Rewrite

$$\begin{aligned} D_n &= \frac{f(\beta_n) - f(0)}{\beta_n}\frac{\beta_n}{\beta_n - \alpha_n} + \frac{f(\alpha_n) - f(0)}{\alpha_n}\frac{-\alpha_n}{\beta_n - \alpha_n} \\ &= \frac{f(\alpha_n) - f(0)}{\alpha_n} + \Big(\frac{f(\beta_n) - f(0)}{\beta_n} - \frac{f(\alpha_n) - f(0)}{\alpha_n}\Big)\frac{\beta_n}{\beta_n - \alpha_n}. \end{aligned}$$

Because $\frac{\beta_n}{\beta_n - \alpha_n}$ is bounded, the limit exists and is equal to $f'(0)$. □

iii. $f'(x)$ exists and is continuous for all $x \in (-1, 1)$.

Solution. For each n, the mean value theorem implies that there exists $\theta_n \in (0, 1)$ such that

$$D_n = f'\big(\alpha_n + \theta_n(\beta_n - \alpha_n)\big).$$

Taking limits on both sides, the continuity of f' implies $\lim D_n = f'(0)$. $\square$

(b) Set $f(x) = x^2 \sin(1/x)$ for $x \neq 0$ and $f(0) = 0$. Observe that f is differentiable everywhere in $(-1, 1)$ and $f'(0) = 0$. Find α_n and β_n that tend to 0 in such a way that D_n converges to a limit unequal to $f'(0)$.

Solution. Let $\beta_n = \frac{1}{n} + \frac{1}{n^2}$ and $\alpha_n = \frac{1}{n}$. $\square$

17.10 Mean Value Theorems

1. In the auctions example.

   (a) Assume in addition that $\sigma(v)$ is a function such that $\forall v \in [0,1], b^*(v) = \sigma(v)$ (there is a symmetric equilibrium). Use Equation 10.5 to show that:

   $$\sigma(v) = v - \sigma'(v)\frac{F(v)}{F'(v)}$$

   The right hand side is called the virtual value.

   Solution. Substituting $b(v) = \sigma(v)$, then $\sigma^{-1}(b) = v$. The equation simplifies to:

   $$(v - \sigma(v))F'(\sigma^{-1}(\sigma(v)))\frac{1}{\sigma'(\sigma^{-1}(\sigma(v)))} - F(\sigma^{-1}(\sigma(v))) = 0$$

   $$(v - \sigma(v))F'(v)\frac{1}{\sigma'(v)} - F(\sigma^{-1}(\sigma(v))) = 0$$

   $$(v - \sigma(v))F'(v)\frac{1}{\sigma'(v)} - F(v) = 0$$

   Rearranging the equation,

   $$\sigma(v) = v - \sigma'(v)\frac{F(v)}{F'(v)}$$

   □

   (b) Using the above equation and the signs of the derivatives, show that if $\forall v \in [0,1], b^*(v) = \sigma(v)$ then $\forall v \in [0,1], \sigma(v) \leq v$ (this show that in a symmetric equilibrium everyone bids weakly below their valuation).

   Solution. Rearrange the above formula:

   $$\sigma(v) = v - \sigma'(v)\frac{F(v)}{F'(v)}$$

   .

   Since the second term is negative, then $\sigma(v) \leq v$.

   □

2. Assume f function is continuous on $[0, \infty)$ and differentiable on $(0, \infty)$. Suppose $f(0) = 0$ and f' is increasing on $(0, \infty)$. Prove

$$g(x) = \frac{f(x)}{x}$$

is increasing on $(0, \infty)$.

Solution. Consider $x_2 > x_1 > 0$. Then by the mean value theorem, there exists $\xi_1 \in (0, x_1)$ and $\xi_2 \in (x_1, x_2)$ such that

$$f(x_1) = f'(\xi_1)(x_1 - 0) + f(0) = f'(\xi_1)x_1$$

and

$$f(x_2) = f'(\xi_2)(x_2 - x_1) + f(x_1) = f'(\xi_2)(x_2 - x_1) + f'(\xi_1)x_1 \geq f'(\xi_1)x_2,$$

where the inequality comes from the fact that $\xi_2 > \xi_1$ and f' is increasing. Therefore

$$\frac{f(x_2)}{x_2} \geq \frac{f(x_1)}{x_1}.$$

□

17.11 Taylor Expansion

1. Suppose $f : \mathbb{R} \to \mathbb{R}$ is twice differentiable. Assume $f(0) > 0$, $f'(0) < 0$ and $f''(x) < 0$ for all $x \in \mathbb{R}$. Prove there exists $\xi \in \left(0, -\frac{f(0)}{f'(0)}\right)$ such that $f(\xi) = 0$.

   Solution. By Taylor's theorem, we have

   $$f(x) = f(0) + f'(0)x + \frac{f''(\eta)}{2}x^2 \quad \text{for some } \eta \text{ between } 0 \text{ and} x.$$

   Then

   $$f\big(-\frac{f(0)}{f'(0)}\big) = \frac{f''(\eta)}{2}\big(-\frac{f(0)}{f'(0)}\big)^2 < 0.$$

   Because $f(0) > 0$, there exists $\xi \in \left(0, -\frac{f(0)}{f'(0)}\right)$ such that $f(\xi) = 0$. □

2. Assume $f : [a, b] \to \mathbb{R}$ is twice differentiable and $f'(a) = f'(b) = 0$. Prove there exists $\xi \in (a, b)$ such that

   $$\big|f''(\xi)\big| \geq \frac{4}{(b-a)^2}\big|f(b) - f(a)\big|.$$

   (Hint: expand $f\big(\frac{a+b}{2}\big)$ at a and b respectively)

   Solution. By Taylor's theorem, we have

   $$f\big(\frac{a+b}{2}\big) = f(a) + f'(a)\frac{b-a}{2} + \frac{1}{2}f''(\xi_1)\big(\frac{b-a}{2}\big)^2 \quad \text{for some } \xi_1 \in \big(a, \frac{a+b}{2}\big),$$

   and

   $$f\big(\frac{a+b}{2}\big) = f(b) - f'(b)\frac{b-a}{2} + \frac{1}{2}f''(\xi_2)\big(\frac{b-a}{2}\big)^2 \quad \text{for some } \xi_2 \in \big(\frac{a+b}{2}, b\big).$$

   Then

   $$\frac{4}{(b-a)^2}\big|f(b)-f(a)\big| = \frac{1}{2}\big|f''(\xi_1)-f''(\xi_2)\big| \leq \frac{1}{2}\big(\big|f''(\xi_1)\big|+\big|f''(\xi_2)\big|\big) \leq \max\big\{\big|f''(\xi_1)\big|, \big|f''(\xi_2)\big|\big\}.$$

   □

3. Let $f : [a, b] \to \mathbb{R}$ be twice differentiable. Assume $\sup_{x\in[a,b]} |f''(x)| \leq M$ for some constant M. Assume also f achieves its global maximum at some point x^* in (a, b). Prove

   $$\big|f'(a)\big| + \big|f'(b)\big| \leq M(b-a).$$

Solution. Because $x^* \in (a, b)$, we know $f'(x^*) = 0$. Now apply the mean value theorem to f': there exists $\xi_1 \in (a, x^*)$ and $\xi_2 \in (x^*, b)$ such that

$$f'(a) = f'(x^*) + f''(\xi_1)(a - x^*),$$

and

$$f'(b) = f'(x^*) + f''(\xi_2)(b - x^*).$$

Hence

$$f'(a) = f''(\xi_1)(a - x^*) \quad \text{and} \quad f'(b) = \frac{f''(\xi_2)}{2}(b - x^*).$$

Thus,

$$|f'(a)| + |f'(b)| \leq |f''(\xi_1)|(x^* - a) + |f''(\xi_2)|(b - x^*) \leq M(b - a).$$

□

17.12 First-Order Differentiation in $\mathbb{R}^n$

1. (Euler's Equations) Assume $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable. Fix $(x, y) \in \mathbb{R}^2$. Define $g(t) = f(tx, ty)$ for all $t > 0$. Show g is differentiable and

$$g'(t) = x\frac{\partial f}{\partial x}(tx, ty) + y\frac{\partial f}{\partial y}(tx, ty).$$

Assume in addition, there exists $\alpha > 0$ such that

$$f(tx, ty) = t^\alpha f(x, y) \quad \forall t > 0 \quad \text{and} \quad \forall (x, y) \in \mathrm{R}^2. \tag{17.1}$$

Show for all $(x, y) \in \mathbb{R}^2$,

$$x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) = \alpha f(x, y). \tag{17.2}$$

A function with the property (17.1) is said to be homogeneous of degree α. The equation (17.2) is called Euler's formula.

Proof. Fix $(x, y) \in \mathbb{R}^2$. Let $h : \mathbb{R} \to \mathbb{R}^2$ be the linear mapping $t \mapsto t \begin{pmatrix} x \\ y \end{pmatrix}$. So $g(t) = f(h(t))$. Since both f and h are differentiable, we know g is differentiable. By chain rule,

$$g'(t) = Dg(t) = Df(h(t))Dh(t) = (\frac{\partial f}{\partial x}(tx, ty), \frac{\partial f}{\partial y}(tx, ty)) \begin{pmatrix} x \\ y \end{pmatrix}.$$

If in addition, (17.1) holds, then we know $g'(t) = \alpha t^{\alpha-1} f(x, y)$ for all t. This implies

$$x\frac{\partial f}{\partial x}(tx, ty) + y\frac{\partial f}{\partial y}(tx, ty) = \alpha t^{\alpha-1} f(x, y), \quad \forall t.$$

Evaluating both sides at $t = 1$ yields the desired result. □

2. (Exercise 16 on page 347, Pugh) Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f = (x, y, z)$ and $g = w$ where

$$\begin{aligned} w &= w(x, y, z) = xy + yz + zx \\ x &= x(s, t) = st \quad y = y(s, t) = s\cos t \quad z = z(s, t) = s\sin t \end{aligned}$$

(a) Find the matrices that represent the linear transformations $(Df)_p$ and $(Dg)_q$ where $p = (s_0, t_0) = (0, 1)$ and $q = f(p)$.

Proof. The representation matrix of $(Df)_p$ is

$$\left(\begin{array}{cc} t & s \\ \cos t & -s \sin t \\ \sin t & s \cos t \end{array} \right) \Bigg|_{(s,t)=(0,1)} = \left(\begin{array}{cc} 1 & 0 \\ \cos 1 & 0 \\ \sin 1 & 0 \end{array} \right).$$

The representation matrix of $(Dg)_q$ is

$$(y + z, x + z, x + y)|_{(x,y,z)=(0,0,0)} = (0, 0, 0).$$

□

(b) Use the Chain rule to calculate the 1×2 matrix $[\partial w/\partial s, \partial w/\partial t]$ that represents $(D(g \circ f))_p$.

Proof. It is simply

$$(0, 0, 0) \left(\begin{array}{cc} 1 & 0 \\ \cos 1 & 0 \\ \sin 1 & 0 \end{array} \right) = (0, 0).$$

□

(c) Plug the functions $x = x(s, t)$, $y = y(s, t)$ and $z = z(s, t)$ directly into $w = w(x, y, z)$ and recalculate $[\partial w/\partial s, \partial w/\partial t]$, verifying the answer given in (b).

Proof. Plugging x, y, z into w yields

$$w(s, t) = s^2 t \cos t + s^2 \cos t \sin t + s^2 t \sin t.$$

Hence

$$\begin{aligned} & (\frac{\partial w}{\partial s}, \frac{\partial w}{\partial t})|_{(s,t)=(0,1)} \\ = \; & (2st \cos t + 2s \cos t \sin t + 2st \sin t, s^2 \cos t - s^2 t \sin t - s^2 \sin^2 t + s^2 \cos^2 t + s^2 \sin t + s^2 t \text{ co} \\ = \; & (0, 0). \end{aligned}$$

□

17.13 Second-Order Differentiation in $\mathbb{R}^n$

1. We showed that a matrix representation exists for a linear map. Why does it have to be unique?

2. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} p_1^3 + p_2^3 \end{pmatrix}.$$

Prove for any $p \in \mathbb{R}^2$, the matrix that represents $(D^2 f)_p$ is

$$\begin{pmatrix} 6p_1 & 0 \\ 0 & 6p_2 \end{pmatrix}.$$

Proof. We will take it as given that we know $(Df)_p$ is represented by

$$\begin{pmatrix} 3p_1^2 & 3p_2^2 \end{pmatrix}.$$

Therefore

$$\begin{aligned} ||R(v)(u)|| &= \Big|\Big|(Df)_{p+v}(u) - (Df)_p(u) - T(v,u)\Big|\Big| \\ &= \Big|\Big| \begin{bmatrix} 3(p_1+v_1)^2 & 3(p_2+v_2)^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - \begin{bmatrix} 3p_1^2 & 3p_2^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{pmatrix} 6p_1 & 0 \\ 0 & 6p_2 \end{pmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \Big|\Big| \\ &= \begin{bmatrix} 3v_1^2 & 3v_2^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 3v_1^2 u_1 + 3v_2^2 u_2 \end{aligned}$$

Notice that residual is linear in the second argument by construction. In this case it is possible to directly compute the operator norm $||R(v)(\cdot)||$,

$$||R(v)(\cdot)|| = \sup_{u \in \mathbb{R}^n, ||u||=1} ||R(v)(u)|| = \sup_{u \in \mathbb{R}^n, ||u||=1} \{3v_1^2 u_1 + 3v_2^2 u_2\}$$

where $u_1^2 + u_2^2 = 1$. To compute the sup, it is without loss to consider $u_1, u_2 \geq 0$. Note that then $u_2 = \sqrt{1-u_1^2} \leq 1 - u_1$, where the equality holds when $u_1 = 0$ or $u_1 = 1$. Thus we have

$$3v_1^2 u_1 + 3v_2^2 u_2 \leq 3v_1^2 u_1 + 3v_2^2(1-u_1) \leq \max\{3v_1^2, 3v_2^2\}$$

The equality could be achieved bt taking $u_1 = 1, u_2 = 0$ if $v_1^2 \geq v_2^2$, and $u_1 = 0, u_2 = 1$

if $v_1^2 < v_2^2$. Therefore, we have shown that $||R(v,\cdot)|| = \max\{3v_1^2, 3v_2^2\}$ and hence

$$\frac{||R(v)(\cdot)||}{||v||} = \frac{\max\{3v_1^2, 3v_2^2\}}{||v||} = \max\{3(v_1^2/||v||), 3(v_2^2/||v||)\}.$$

$|v_i| \leq ||v||$ is bounded by construction. Therefore,

$$\lim_{v \to 0_{2\times 1}} v_1^2/||v|| = 0$$

$$\lim_{v \to 0_{2\times 1}} v_2^2/||v|| = 0$$

Therefore, $R(v)(\cdot)$ is sublinear and therefore we have show that our candidate matrix is second derivative.

□

3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as

$$f(x) = x^T A^T A x$$

where A is an $n \times n$ matrix. Calculate the matrices that represent $(Df)_x$.

Proof. Let $g : \mathbb{R}^n \to \mathbb{R}$ be such that $g(y) = y^T y = \sum_{i=1}^m y_i^2$. By calculating the first order partials, it is easy to see

$$2(y_1, \cdots, y_m) = 2y^T$$

represents $(Dg)_y$. Let $h : \mathbb{R}^n \to \mathbb{R}^n$ be such that $h(x) = Ax$. Then A represents $(Dh)_x$. Since $f(x) = g(h(x))$, by chain rule $(Df)_x = (Dg)_{h(x)} \circ (Dh)_x$, and

$$2h(x)^T A = 2x^T A^T A$$

is the representation matrix.

□

4. Assume that X is an $n \times k$ full rank matrix and that $Y \in \mathbb{R}^n$. Show that $\hat{\beta} = (X^t X)^{-1} X^t Y$ is the solution to the least squares criterion function by computing the first order conditions of

$$(Y - X\beta)^t (Y - X\beta)$$

17.14 Comparative Statics

1. Consider the Auctions Example in previous chapters. Show that $b^*(v)$ is increasing in v. [Hint: Use the implicit function theorem].

   Solution. Because this $b^*(v)$ is an interior maximum, $\frac{\partial^2}{\partial b^2}U_v(b^*) < 0$. Furthermore, since σ is strictly increasing then σ^{-1} is also strictly increasing (by using the inverse mapping theorem we can further show that $\sigma' > 0$).

   For applying the implicit function theorem, let $H(b, v) = \frac{\partial U_v}{\partial b} = 0$. Then

$$B = \frac{\partial H}{\partial b} = \frac{\partial^2}{\partial b^2}U_v(b^*) < 0. \text{ Then } B^{-1} < 0$$

$$A = \frac{\partial H}{\partial v} = \frac{\partial^2}{\partial b \partial v}U_v(b^*) = F'(\sigma^{-1}(b))\frac{\partial}{\partial b}\sigma^{-1}(b) > 0$$

   Using the implicit function theorem:

$$\frac{\partial b^*(v)}{\partial v} = -B^{-1}A > 0$$

   Therefore $b^*(v)$ is an increasing function of v.

□

2. Consider the following Keynesian IS-LM model. Suppose

$$\begin{aligned} Y &= C(Y - T) + I(r) + G \\ M &= L(Y, r) \end{aligned}$$

   where Y is GDP, T is taxes, r is interest rate, G is government spending and M is money supply. The functions $C(\cdot)$, $I(\cdot)$ and $L(\cdot, \cdot)$ are consumption function, investment function and money supply function respectively. Assume they are continuously differentiable and

$$0 < C'(x) < 1, \quad I'(r) < 0, \quad \frac{\partial L}{\partial Y} > 0, \quad \text{and} \quad \frac{\partial L}{\partial r} < 0.$$

   Suppose G, M and T are independent variables which can be controlled, Y and r are dependent variables determined by G, M and T. Analyze the relationships between $\{Y, r\}$ and $\{G, M, T\}$.

Solution. Define

$$\begin{aligned} f(T,G,M,Y,r) &= Y - C(Y-T) - I(r) - G \\ h(T,G,M,Y,r) &= L(Y,r) - M. \end{aligned}$$

Then

$$\begin{pmatrix} \frac{\partial f}{\partial Y} & \frac{\partial f}{\partial r} \\ \frac{\partial h}{\partial Y} & \frac{\partial h}{\partial Y} \end{pmatrix} = \begin{pmatrix} 1 - C'(Y-T) & -I'(r) \\ \frac{\partial L}{\partial Y} & \frac{\partial L}{\partial r} \end{pmatrix}.$$

This matrix is invertible because its determinant $\Delta = (1 - C'(Y-T))\frac{\partial L}{\partial r} + I'(r)\frac{\partial L}{\partial Y} < 0$. Therefore

$$\begin{aligned} \begin{pmatrix} \frac{\partial Y}{\partial T} & \frac{\partial Y}{\partial G} & \frac{\partial Y}{\partial M} \\ \frac{\partial r}{\partial T} & \frac{\partial r}{\partial G} & \frac{\partial r}{\partial M} \end{pmatrix} &= -\begin{pmatrix} 1 - C'(Y-T) & -I'(r) \\ \frac{\partial L}{\partial Y} & \frac{\partial L}{\partial r} \end{pmatrix}^{-1} \begin{pmatrix} C'(Y-T) & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\ &= -\frac{1}{\Delta}\begin{pmatrix} \frac{\partial L}{\partial r} & I'(r) \\ -\frac{\partial L}{\partial Y} & 1 - C'(Y-T) \end{pmatrix}\begin{pmatrix} C'(Y-T) & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\ &= -\frac{1}{\Delta}\begin{pmatrix} \frac{\partial L}{\partial r}C'(Y-T) & -\frac{\partial L}{\partial r} & -I'(r) \\ -\frac{\partial L}{\partial Y}C'(Y-T) & \frac{\partial L}{\partial Y} & -1 + C'(Y-T) \end{pmatrix} \end{aligned}$$

Therefore $\frac{\partial Y}{\partial T} < 0$, $\frac{\partial Y}{\partial G} > 0$, $\frac{\partial Y}{\partial M} > 0$, $\frac{\partial r}{\partial T} < 0$, $\frac{\partial r}{\partial G} > 0$ and $\frac{\partial r}{\partial M} < 0$. □

Bibliography

Greene, W. H. (2012). *Econometric analysis*. Pearson Education India.

Naiman, D. Q. and E. R. Scheinerman (2017). Arbitrage and geometry. *arXiv preprint arXiv:1709.07446*.

Pugh, C. C. and C. Pugh (2002). *Real mathematical analysis*, Volume 2011. Springer.

Rudin, W. et al. (1964). *Principles of mathematical analysis*, Volume 3. McGraw-hill New York.

Sundaram, R. K. et al. (1996). *A first course in optimization theory*. Cambridge university press.