1 Single Variable Calculus

References: Simon and Blume, chapters 2 to 5.

This section studies real functions of one variable, that is, functions f mapping $\mathbb{R}$ into $\mathbb{R}$. We write $f : \mathbb{R} \to \mathbb{R}$ or $x \mapsto f(x) = y$.

We work in the *metric space* $\mathbb{R}$ using the *Euclidean distance* on $\mathbb{R}$ as metric: $d(x, y) = |x - y|$ for any $x, y \in \mathbb{R}$. This is assumed in what follows if no other details are given.

1.1 Some definitions

1.1.1 Metric Space

A *metric space* is an ordered pair (M, d), where M is a non-empty set and d is a metric on M, i.e. a distance function

$$d : M \times M \to \mathbb{R}_+,$$

such that for any $x, y, z \in M$,

1. $d(x, y) \geq 0$ (non-negativity),
2. $d(x, y) = 0$ if and only if (iff) $x = y$ (identity of indiscernibles),
3. $d(x, y) = d(y, x)$ (symmetry),
4. $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

1.1.2 Sets

- A *set* is a collection of elements of some kind. In general, from now onwards capital letters will denote sets while small letters will denote elements of a set.
- The set B is a *subset* of a set A if every element in B also belongs to A (we write $B \subset A$ or $B \subseteq A$ and $a \in A$, for the singleton set $\{a\}$).
- The *empty set* $\emptyset = \{\}$ is the set that contains no elements at all.
- *Operations with sets:* intersection ($\bigcap$), union ($\bigcup$), difference ($-$) or ($\setminus$), complementation (^{c}).
- The set of *natural numbers*, $\mathbb{N}$, includes all non-negative whole numbers: $\mathbb{N} = \{0, 1, 2, 3, ...\}$.

- The set of integers, $\mathbb{Z}$, includes both positive and negative whole numbers:
  $\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$.
- The set of *rational numbers*, $\mathbb{Q}$, includes all fractional numbers:
  $\mathbb{Q} = \left\{ r : r = \frac{p}{q} \text{ where } p \in \mathbb{Z} \text{ and } q \in \mathbb{N} \setminus \{0\} \right\}$.
- An *irrational* number is one that cannot be reduced to a ratio between an integer and a non-zero natural. $\sqrt{2}$ is an irrational number.
- The set of *real numbers*, $\mathbb{R}$, includes all rational and irrational numbers. It is often called the *continuum* or the *real line*.
- The ϵ-*neighbourhood* of $x \in X$ and is denoted by $O_\epsilon(x)$ and defined as the set of $y \in X$ such that $d(x, y) < \epsilon$ where d is a measure of distance (in $\mathbb{R}$ this is typically the Euclidean distance): $O_\epsilon(x) = \{y \in X : d(x, y) < \epsilon\}$.
- A subset S of M is *open* if every point of S has a neighbourhood that lies entirely within S.
- A point $x \in M$ is a *limit point* of S if every neighbourhood of x contains at least one point of S other than x.
- A subset S of M is *closed* if it contains all its limit points.
- The complement of an open set is closed, and vice versa.
- Intervals in $\mathbb{R}$: $[a, b]$ is a closed interval, (a, b) is an open interval and a or b can be $+(-)\infty$.
- A subset S of M is *bounded* if there is a finite k such that $d(x, y) < k$ for all $x, y \in S$.
- A subset S of M is *compact* if it is closed and bounded.
- A subset S of M is *convex* if, for any $s_1, s_2 \in S$ and all $\alpha \in [0, 1]$, $\alpha s_1 + (1 - \alpha)s_2 \in S$.

1.2 Functions

- A *function* f is a relation which assigns to each element of a set A, called the domain, a unique element of a set B, called the range or codomain. It is denoted by $f : A \to B$.
- In the general case where A and B may not be real, f is typically called a *mapping* of A into B.
- For any $a \in A$, the corresponding element in B, $b = f(a)$, is called the *image* of a while a is the *preimage* of b. More generally, we can talk of images and preimages of sets of elements.

- A function f is said to be *surjective*, or to map A *onto* B, if $f(A) = B$; i.e. its values span the entire codomain; for each value $b \in B$, there exist $a \in A$ such that $f(a) = b$.
- A function f is said to map A *into* B if $f(A) \subset B$.
- If for every unique $a \in A$, the image $f(a) \in B$ is unique, then f is said to be *injective*.
- A function f of A onto B - i.e. it is surjective - is said *one-to-one* or a *one-to-one correspondence* or *bijection* if each element $b \in B$ has a unique preimage $a \in A$, i.e. if this mapping is also injective. We write $a = f^{-1}(b)$ where f^{-1} is called the *inverse* of f. Remark: If both $f(\cdot)$ and its inverse $f^{-1}(\cdot)$ are continuous (see below), then f is called a homeomorphism.
- The *graph of a real function* is the set $\{(x, y) \in \mathbb{R}^2 : y = f(x)\}$. It allows us to depict the function graphically by plotting the pairs (x, y) in a xy-plane.
- *Monotonicity*: a real function f is *strictly increasing* (*decreasing*) if

$$\forall x_1, x_2 \in \mathbb{R} : x_1 > x_2 \Rightarrow f(x_1) > (<) f(x_2)$$

- *Extrema*: the points at which a real function changes from increasing to decreasing. They can be *minima* or *maxima*, they can be *local* or *global*.
- Extrema are related to bounds on the codomain of a function $f : A \to B$: $\overline{b}$ is an *upper bound* on f if $f(a) \leq \overline{b}$ for all $a \in A$; $\underline{b}$ is a *lower bound* on f if $f(a) \geq \underline{b}$ for all $a \in A$.
- A number c is called a *least upper bound* for f if it is an upper bound and if $c \leq \overline{b}$ for each upper bound $\overline{b}$ of f. Such a least upper bound is called the *supremum* of f, denoted by $\sup_{a \in A} f(a)$. Similarly, the *infimum* of f is the *largest lower bound* on f, denoted by $\inf_{a \in A} f(a)$.
- When f achieves a maximum (minimum), then the maximum (minimum) is equal to the supremum (infimum) of f.

1.2.1 Some examples of functions and their properties

- The following are a few functions you should be familiar with (for some $a, b, c \in \mathbb{R}$): For $x \in \mathbb{R}$,

  $f(x) = ax + b, a \neq 0$, (bijective; $f^{-1}(y) = (y - b)/a$); unbounded

  $f(x) = ax^2 + bx + c, a \neq 0$, (surjective, but not injective; minimum is $c - \frac{b^2}{2a}$)

  $f(x) = x^n, n \in \mathbb{N}_+$ (bijective for n odd, $f^{-1}(y) = y^{\frac{1}{n}}$; surjective for n even, with minimum 0)

  $f(x) = \exp(x)$, codomain $\mathbb{R}_+$ ($a \neq 1$: surjective and injective, i.e. bijective; $f^{-1}(y) = \ln(y)$, $y > 0$, infimum zero; $a = 1$: not injective)

  $f(x) = a^x, a > 0$ (surjective and injective with codomain $\mathbb{R}_+$, i.e. bijective; $f^{-1}(y) = \ln(y)/\ln(a)$, infimum zero)

  $f(x) = \ln(x)$, with domain $\mathbb{R}_+$ (surjective and injective, i.e. bijective; $f^{-1}(y) = \exp(y)$, unbounded)

  $f(x) = |x|$ (surjective, but not injective, minimum zero)

  $f(x) = 1/(1 + \exp(x))$ (surjective and injective with codomain (0,1), $f^{-1}(y) = \ln(1 - y) - \ln(y)$, infimum=0, supremum=1)

- A *composite function* $(f \circ g)(x)$ defines the successive application of two functions: $(f \circ g)(x) = f(g(x))$.

1.3 Limits and continuity

1.3.1 Concepts

- *Continuity at a point*: A function f is continuous at a point x_0 if

$$\forall \epsilon > 0 \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon.$$

- *Continuous function*: a function that is continuous at every point of its domain (we say that a continuous function is of class $\mathcal{C}^0$, the class of pointwise continuous functions).

- *Limit of a function at a point*: $y \in \mathbb{R}$ is the limit of f at x_0 if

$$\forall \epsilon > 0 \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - y| < \epsilon$$

  and we write $\lim_{x \to x_0} f(x) = y$.

- Alternative definition of continuity at a point: f is continuous at x_0 if for any sequence x_n in $\mathbb{R}$ such that $\lim_{n \to \infty} x_n = x_0$ it is the case that $\lim_{n \to \infty} f(x_n) = f(x_0)$.

- If two functions f and g are continuous, then $|f|$, $|g|$, $f+g$, fg and $f \circ g$ are continuous.

- Extension: A function $f(x)$, $x \in X$, is *uniformly continuous* on X if

$$\forall \epsilon > 0 \exists \delta > 0 : \sup_{x, x_0 \in X : |x - x_0| < \delta} |f(x) - f(x_0)| < \epsilon$$

  Note that, unlike in the case of continuity at x_0 where δ may depend on ϵ and x_0, in the case of uniform continuity δ does not depend on x_0.

  Example: The function $f(x) = \exp(x)$, $x \in \mathbb{R}$ is continuous at every point $x \in \mathbb{R}$, but not uniformly continuous on $\mathbb{R}$. The reason is that $f(x) = \exp(x)$ grows exponentially as x tends to $+\infty$, so that for sufficiently large x_0, say $x_0 > \tilde{x}$, $\sup_{x, x_0 > \tilde{x} : |x - x_0| < \delta} |f(x) - f(x_0)| > \epsilon$; hence, in order to satisfy the ϵ bound, δ needs to shrink as $\tilde{x}$ increases and, therefore, δ is not uniform on $\mathbb{R}$.

- Extension: Let (X, d) and (Y, d) be two metric spaces, and F a family of functions from X to Y. The family F is *equicontinuous* at a point $x_0 \in X$ if, given any $\epsilon > 0$, there exists a $\delta > 0$, such that $d(f(x_0), f(x)) < \epsilon$ for all $f \in F$ and all x whenever $d(x_0, x) < \delta$. Formally,

$$\forall \epsilon > 0 \; \exists \delta > 0 : d(x, x_0) < \delta \Rightarrow \sup_{f \in F} d(f(x_0), f(x)) < \epsilon.$$

  The family is equicontinuous if it is equicontinuous at each point of X.

  The family F is *uniformly equicontinuous* if, given any $\epsilon > 0$, there exists a $\delta > 0$, such that $d(f(x_1), f(x_2)) < \epsilon$ for all $f \in F$ and all $x_1, x_2 \in X$ whenever $d(x_1, x_2) < \delta$. Formally,

$$\forall \epsilon > 0 \; \exists \delta > 0 : \sup_{x_1 x_2 \in X : d(x_1, x_2) < \delta} \sup_{f \in F} d(f(x_1), f(x_2)) < \epsilon.$$

  For comparison, the statement "all functions $f \in F$ are continuous" means that, given any $\epsilon > 0$, any $f \in F$, and any $x_0 \in X$, there exists a $\delta > 0$, such that $d(f(x_0), f(x)) < \epsilon$ for all $x \in X$ whenever $d(x_0, x) < \delta$. So, for continuity, δ may depend on ϵ, x_0 and f; for uniform continuity, δ may depend on ϵ and f, but not on x_0; for equicontinuity, δ must be independent of f, i.e. it must be the same for all $f \in F$, but it may depend on x_0; and for uniform equicontinuity, δ must be independent of both f and x_0.

- Examples:

  (1) Consider the family of functions $F_\tau = \{f(x; \tau) = x + \tau, x \in \mathbb{R}, \tau \in \mathbb{R}\}$, which is parameterized by τ. Given $\epsilon > 0$, for any $x_0 \in \mathbb{R}$ and any $\tau \in \mathbb{R}$, $|x_0 - x| < \epsilon$ implies $|f(x_0; \tau) - f(x; \tau)| = |x_0 - x| < \epsilon$. So, given $\epsilon > 0$, there exists $\delta > 0$, namely $\delta = \epsilon$, such that $|x_0 - x| < \delta = \epsilon$ implies $|f(x_0; \tau) - f(x; \tau)| < \epsilon$. Since this δ does not depend on $f \in F_\tau$ (i.e. it does not depend on τ), the family F_τ is equicontinuous; since it does not depend on x_0 either, the family F_τ is uniformly equicontinuous.

(2) Consider the family of functions $F_\alpha = \{f(x;\alpha) = \alpha x, x \in \mathbb{R}, \alpha \in \mathbb{R}\}$, which is parameterized by α. Given $\epsilon > 0$ and $x_0 \in \mathbb{R}$, $|x_0 - x| < \delta = \frac{\epsilon}{|\alpha|}$ implies $|f(x_0;\alpha) - f(x;\alpha)| = |\alpha||x_0 - x| < \epsilon$, for any $\alpha \in \mathbb{R} \setminus \{0\}$. So, given ϵ, there exists $\delta = \frac{\epsilon}{|\alpha|}$, $\alpha \neq 0$, such that $|x_0 - x| < \delta$ implies $|f(x_0;\alpha) - f(x;\alpha)| < \epsilon$. Note, however, that this δ depends on α (i.e. on $f \in F_\alpha$), so F_α is not equicontinuous (and hence not uniformly equicontinuous). Every member f of the family F_α is continuous, however, and in fact uniformly continuous.

(3) Consider the family of functions $F_\gamma = \{f(x;\gamma) = \cos(x + \gamma), x \in \mathbb{R}, \gamma \in \mathbb{R}\}$. Recall from trigonometry: $|\cos(x_0) - \cos(x)| \leq |x_0 - x|$. Then, given $\epsilon > 0$, $|x_0 - x| < \epsilon$ implies $|f(x_0;\gamma) - f(x;\gamma)| = |\cos(x_0+\gamma) - \cos(x+\gamma)| \leq |x_0 - x| < \epsilon$. So, given $\epsilon > 0$, there exists $\delta > 0$, namely $\delta = \epsilon$, such that $|x_0 - x| < \delta$ implies $|f(x_0;\gamma) - f(x;\gamma)| < \epsilon$. As in (1), this δ does not depend on $f \in F_\gamma$ (i.e. it does not depend on γ), so F_γ is equicontinuous; moreover, it does not depend on x_0, so F_γ is uniformly equicontinuous.

- Extension: *Lipschitz continuity* is a smoothness condition for functions which is stronger than regular continuity. Intuitively, a Lipschitz continuous function is limited in how fast it can change; a line joining any two points on the graph of this function will never have a slope steeper than a certain number called the Lipschitz constant of the function.

  Let (X, d) and (Y, d) be two metric spaces, and $f \in F$ a family of functions from X to Y. The function f is called Liptschitz continuous if if there exists a real, finite constant $K \geq 0$ such that, for all x_1 and x_2 in X,

  $$d(f(x_1), f(x_2)) \leq K d(x_1, x_2).$$

  The smallest such K is called the Lipschitz constant of the function f. If $K = 1$, the function is called a short map, and if $0 < K < 1$, the function is called a *contraction*.

  More generally, a function f defined on X is said to be Hölder continuous or to satisfy a Hölder condition of order $\alpha > 0$ on X if there exists a finite constant $M > 0$ such that

  $$d(f(x_1), f(x_2)) \leq M d(x_1, x_2)^\alpha = M|x_1 - x_2|^\alpha$$

  for all $x_1, x_2 \in X$.

  If there exists $K \geq 1$ such that

  $$\frac{1}{K} d(x_1, x_2) \leq d(f(x_1), f(x_2)) \leq K d(x_1, x_2)$$

  then f is called bi-Lipschitz. A bi-Lipschitz mapping is injective, and is in fact a homeomorphism onto its image.

A set of functions with the same Lipschitz constant is (uniformly) equicontinuous. This is in particular the case, for example, if the set consists of functions with derivatives (see below) bounded by the same constant.

Examples: (below)

- Suppose there exists a $y \in \mathbb{R}$ such that for any sequence of real numbers x_n satisfying $\lim_{n\to\infty} x_n = +(-)\infty$ it is the case that $\lim_{n\to\infty} f(x_n) = y$. Then y is the limit of f as x approaches $+(-)\infty$ and we write $y = \lim_{x\to\infty} f(x)$. Equivalently, $y = \lim_{x\to\infty} f(x)$ if and only if for all $\epsilon > 0$ there exists $s > 0$ such that $|f(x) - y| < \epsilon$ for all $x > s$.

- Extension: The limit inferior (lim inf) and limit superior (lim sup) can be thought of as eventual extreme bounds on the function, even when the function itself does not converge. If f is a real valued function defined for all x in an interval containing y, then the limit superior and limit inferior are defined as

$$\begin{aligned}\limsup_{x\to y} f(x) &= \inf_{\delta>0} \sup_{0<|x-y|<\delta} f(x) \\ \liminf_{x\to y} f(x) &= \sup_{\delta>0} \inf_{0<|x-y|<\delta} f(x),\end{aligned}$$

  satisfying $\liminf_{x\to y} f(x) \leq \limsup_{x\to y} f(x)$, and the limit $f(y) = \lim_{x\to y} f(x)$ exists if and only if $\liminf_{x\to y} f(x) = \limsup_{x\to y} f(x)$.

- There exists a similar definition of lim inf and lim sup in terms of sequences:

$$\begin{aligned}\limsup_{x_n\to\infty} f(x_n) &= \inf_n \sup_{m>n} f(x_m) \\ \liminf_{x_n\to\infty} f(x_n) &= \sup_n \inf_{m>n} f(x_m),\end{aligned}$$

  satisfying the analogous (in)equalities as above.

- Examples: (For the purpose of easy exposition, assume that the sequence $\{x_n, n = 1, 2, \cdots\}$ below satisfies $x_n < x_{n+1}$ for all $n \in \mathbb{N}$.)

  1. $f(x) = 1/(1+\exp(x))$, $x \in \mathbb{R}$; this functions satisfies $\lim_{x\to-\infty} f(x) = 1$, and $\lim_{x\to+\infty} f(x) = 0$. To confirm the existence of the limit, note that

$$\begin{aligned}\limsup_{x_n\to+\infty} f(x_n) &= \inf_n \sup_{m>n} f(x_m) = \inf_n 1/(1+\exp(x_n)) = 0 \\ \liminf_{x_n\to+\infty} f(x_n) &= \sup_n \inf_{m>n} f(x_m) = \sup_n 0 = 0,\end{aligned}$$

     and analogously for the limit when $x \to -\infty$.

  2. $f(x) = \cos(x)$, $x \in \mathbb{R}$; this function does not converge, but it satisfies $-1 = \liminf_{x\to\infty} f(x) < \limsup_{x\to\infty} f(x) = 1$. To confirm, note that

$$\begin{aligned}\limsup_{x_n\to+\infty} f(x_n) &= \inf_n \sup_{m>n} f(x_n) = \inf_n 1 = 1 \\ \liminf_{x_n\to+\infty} f(x_n) &= \sup_n \inf_{m>n} f(x_n) = \sup_n(-1) = -1.\end{aligned}$$

- Extension: o and O notation (Landau symbols)

  Let $f(x)$ and $g(x)$ be two functions defined on some subset of $\mathbb{R}$.

  Then, if $\limsup_{x\to a}\left|\frac{f(x)}{g(x)}\right| < \infty$, we write $f(x) = O(g(x))$. Read: "$f(x)$ is big O of $g(x)$".

  If $\lim_{x\to a}\left|\frac{f(x)}{g(x)}\right| = 0$, we write $f(x) = o(g(x))$. Read: "$f(x)$ is little o of $g(x)$".

  In mathematics, computer science, and related fields, big O notation describes the limiting behavior of a function when the argument tends towards a particular value or infinity, usually in terms of simpler functions. Big O notation allows its users to simplify functions in order to concentrate on their growth rates: different functions with the same growth rate may be represented using the same O notation. The little o notation is used when $g(x)$ grows much faster than $f(x)$.

  Example:

  Consider the function $f(x) = \alpha x^4 + \beta x^3 + \gamma$, for some positive parameters α, β and γ. Then, for $|x| > 1$,

$$\begin{aligned} |f(x)| &\leq \alpha x^4 + \beta|x|^3 + \gamma \\ &\leq \alpha x^4 + \beta x^4 + \gamma x^4 \\ &= (\alpha + \beta + \gamma)x^4 \\ &= \kappa x^4 \end{aligned}$$

  where $\kappa = \alpha + \beta + \gamma > 0$. Similarly, for the lower bound, if $x > 0$,

$$\begin{aligned} |f(x)| &= |\alpha x^4 + \beta x^3 + \gamma| \\ &= \alpha x^4 + \beta x^3 + \gamma \\ &\geq \alpha x^4 + \beta x^3 \\ &\geq \alpha x^4. \end{aligned}$$

  If $x < 0$,

$$\begin{aligned} |f(x)| &\geq \alpha x^4 + \beta x^3 + \gamma \\ &\geq \alpha x^4 + \beta x^3 \\ &= (\alpha x + \beta)x^3; \end{aligned}$$

  hence, if $x < -\beta/\alpha$, then $(\alpha x + \beta)x^3 > 0$, with $\alpha x + \beta = O(x)$ and $x^3 = O(x^3)$, so that $(\alpha x + \beta)x^3 = O(x^4)$. Hence, $f(x) = O\left(x^4\right)$.

  Moreover, for any $\delta > 0$, $f(x) = o\left(|x|^{4+\delta}\right)$.

  The last two conclusions make use of the following rules for Landau symbols.

- Some rules for o and O notation
  - $f(x) = o(x^\alpha)$ and $\beta > \alpha$ implies $f(x) = o(x^\beta)$.
  - $f(x) = O(x^\alpha)$ and $\beta > \alpha$ implies $f(x) = o(x^\beta)$.
  - $f(x) = O(x^\alpha)$ and $g(x) = O(x^\beta)$ implies $f(x)g(x) = O(x^{\alpha+\beta})$.
  - $f(x) = O(x^\alpha)$ and $g(x) = o(x^\beta)$ implies $f(x)g(x) = o(x^{\alpha+\beta})$.
  - $f(x) = o(x^\alpha)$ and $g(x) = o(x^\beta)$ and $\beta \geq \alpha$ implies $f(x) + g(x) = o(x^\beta)$.
  - $f(x) = O(x^\alpha)$ and $g(x) = O(x^\beta)$ and $\beta \geq \alpha$ implies $f(x) + g(x) = O(x^\beta)$.
  - $f(x) = O(x^\alpha)$ and $g(x) = o(x^\beta)$ and $\beta > \alpha$ implies $f(x) + g(x) = o(x^\beta)$.
  - $f(x) = O(x^\alpha)$ and $g(x) = o(x^\beta)$ and $\beta < \alpha$ implies $f(x) + g(x) = O(x^\alpha)$.
  - $f(x) = O(x^\alpha)$ and $g(x) = o(x^\alpha)$ implies $f(x) + g(x) = O(x^\alpha)$.

1.3.2 Properties of limits

Consider two real functions f and g. Suppose $\lim_{x\to x_0} f(x) = a \in \mathbb{R}$ and $\lim_{x\to x_0} g(x) = b \in \mathbb{R}$. Let c be some real constant. Then,

- $\lim_{x\to x_0} c = c$
- $\lim_{x\to x_0} [f(x) \pm g(x)] = \lim_{x\to x_0} f(x) \pm \lim_{x\to x_0} g(x) = a \pm b$
- $\lim_{x\to x_0} [f(x)g(x)] = \lim_{x\to x_0} f(x) \lim_{x\to x_0} g(x) = ab$
- If $b \neq 0$, $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)} = \frac{a}{b}$
- $\lim_{x\to x_0} f(x)^n = [\lim_{x\to x_0} f(x)]^n = a^n$
- $\lim_{x\to x_0} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x\to x_0} f(x)} = \sqrt[n]{a}$
- $\lim_{x\to x_0} \exp\{f(x)\} = \exp\{\lim_{x\to x_0} f(x)\} = \exp\{a\}$
- If $a > 0$, $\lim_{x\to x_0} \ln f(x) = \ln[\lim_{x\to x_0} f(x)] = \ln a$
- In general, for any continuous functions f and g: $\lim_{x\to x_0} f[g(x)] = f[\lim_{x\to x_0} g(x)]$, provided the functions are defined at the limits.

1.3.3 Weierstrass Theorem

Let $f : \mathcal{C} \to \mathbb{R}$ be a continuous function whose domain is a compact subset $\mathcal{C}$ in $\mathbb{R}$. Then, there exist points x_m and x_M in $\mathcal{C}$ such that $f(x_m) \leq f(x) \leq f(x_M)$ for all $x \in \mathcal{C}$; that is, $x_m \in \mathcal{C}$ is the global minimizer of f in $\mathcal{C}$ ($f(x_m)$ is the global minimum of f on $\mathcal{C}$) and $x_M \in \mathcal{C}$ is the global maximizer of f in $\mathcal{C}$ ($f(x_M)$ is the global maximum of f on $\mathcal{C}$).

1.4 Slope of a function and derivatives

1.4.1 Concepts

- Suppose you have a linear function, $f(x) = ax + b$. The slope of the linear function is given by,

$$a = \frac{f(y) - f(x)}{y - x}$$

  where x and y are any two distinct real numbers.

- The slope of a non-linear function at a point x_0 can be approximated by taking a small increment to x_0, call it Δx, and computing the increment's ratio,

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

  This is the slope of a *secant line*.

- However, this is generally a rough measure of the slope of a non-linear function. Its level of accuracy depends on the size of Δx: the smaller it is, the more accurate is this measure of slope.

- A more accurate measure of the slope of a function at a given point x_0 is given by the slope of the *tangent line*,

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

  which is well defined if this limit exists.

- Denote a positive sequence tending to zero by $\searrow$, and a negative sequence tending to zero by $\nearrow$. A function is *differentiable at* x_0 if the following limits exist

$$\lim_{\Delta x \searrow 0} \left| \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \right| < \infty$$
$$\lim_{\Delta x \nearrow 0} \left| \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \right| < \infty$$

  and the two limits are the same, i.e.

$$f'(x_0) := \frac{d}{dx} f(x) \Big|_{x=x_0} = \lim_{\Delta x \searrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \lim_{\Delta x \nearrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

  In this case, $f'(x_0)$ is called the *derivative* of f at x_0.

- A function is *differentiable* if it is differentiable at every point of its domain.

- A function is said to be continuously differentiable (or of class $\mathcal{C}^1$) if it is differentiable and f' is a continuous function.
- Higher order derivatives can be obtained by sequential differentiation.
- A function is of class $\mathcal{C}^n$ if its nth derivative is a continuous function. Denote by f'' the second order derivative, f''' the third order derivative, and in general $f^{(n)}$ the derivative of degree n.

1.4.2 Continuity and differentiability

Theorem: If a function f is differentiable at a point x_0, then it is continuous at x_0.

To see why, suppose this is not true: f is differentiable at x_0, but not continuous at x_0. Then, e.g.

$$\lim_{\Delta x \searrow 0} f(x_0 + \Delta x) \quad = \quad f(x_0) + c, \text{ for some } c \neq 0.$$

But we know that

$$\lim_{\Delta x \searrow 0} \Delta x \quad = \quad 0$$

These two expressions then imply

$$\begin{aligned} \lim_{\Delta x \searrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} &= \frac{c}{0} \\ &= \pm\infty \end{aligned}$$

which means that f is not differentiable at x_0. An analogous argument applied if $\Delta x \nearrow 0$.

1.4.3 Some rules of differentiation

Let u and v be some real functions, u' and v' be the derivatives of u and v, c be a real constant, x be a real variable and n be an integer.

function	*derivative*
c (constant)	0
x^n	nx^{n-1}
$\exp(u(x))$	$u'(x)\exp(u(x))$
$\ln u(x)$	$u'(x)/u(x)$
$cu(x)$	$cu'(x)$
$u(x) \pm v(x)$	$u'(x) \pm v'(x)$
$u(x)v(x)$	$u'(x)v(x) + u(x)v'(x)$
$u(x)/v(x)$	$[u'(x)v(x) - u(x)v'(x)]/v^2(x)$
$u(x)^{v(x)}$	$u(x)^{v(x)}\left(\frac{u'(x)v(x)}{u(x)} + v'(x)\ln u(x)\right)$

1.4.4 Special differentiation rules

1. **The composite function differentiation rule - the chain rule**

   Let u and v be two arbitrary differentiable real functions. The function $f(x) = u(v(x))$ is a *composite function*. The derivative of the composite function can be obtained using the chain rule,

   $$f'(x) = \frac{du}{dv}\frac{dv}{dx} = u'(v(x))v'(x)$$

   Exercise: Use the chain rule to verify the last row in the preceding table of derivatives.

2. **The inverse function differentiation**

   Suppose f is one-to-one in an interval $U \subset \mathbb{R}$ so that $f : U \to V$ is invertible in U and $f^{-1} : V \to U$ is the uniquely defined inverse function of f on that subset.

   If, in addition, f is $\mathcal{C}^1$ in U (so that it is a homeomorphism) and $f'(x) \neq 0$ for all $x \in U$,

   $$f^{-1'}(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}$$

   where

   $$\begin{aligned} y &= f(x) \\ x &= f^{-1}(y) \end{aligned}$$

   since we can write

   $$1 = \frac{dx}{dy}\frac{dy}{dx} = \frac{df^{-1}}{dy}\frac{df}{dx},$$

   or, differentiating the identity $y \equiv f(f^{-1}(y))$, any $y \in V$, on both sides, using the chain rule,

   $$\begin{aligned} 1 &= f'(f^{-1}(y))f^{-1'}(y) \\ \Rightarrow f^{-1'}(y) &= \frac{1}{f'(f^{-1}(y))}. \end{aligned}$$

Exercise: Use this rule to find the derivative of the inverse function of $y = f(x) = 1/(1 + \exp(x))$, the logistic function, and directly verify the result, using $f^{-1}(y)$.

3. **de l'Hôpital's rule**

   In calculus, de l'Hôpital's rule (also called Bernoulli's rule) uses derivatives to help evaluate limits involving indeterminate forms.

   If $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$ or $\lim_{x\to c} g(x) = \pm\infty$ and $\lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists, then $\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$.

   Example: constant-elasticity-of-substitution utility function $U(x;\sigma) = \frac{x^{1-\sigma}-1}{1-\sigma}$, for $\sigma \in [0, 1]$. Using de l'Hôpital's rule, applied to $U(x;\sigma)$ as a function of σ,

$$\begin{aligned} U(x;1) &= \lim_{\sigma\to 1} \frac{x^{1-\sigma}-1}{1-\sigma} \\ &= \lim_{\sigma\to 1} \frac{\exp((1-\sigma)\ln(x)) - 1}{1-\sigma} \\ &= \lim_{\sigma\to 1} \frac{-\ln(x)x^{1-\sigma}}{-1} \\ &= \ln(x). \end{aligned}$$

1.5 Using derivatives to characterize functions

- The slope of a function is revealed by the sign of its first derivative: f is increasing (decreasing) in (a, b) iff $f'(x) > (<)0$ for any x in this interval.
- The convexity of a function f is revealed by the sign of its second derivative: f is convex (concave) in (a, b) iff $f''(x) > (<)0$ for any x in this interval.
- Necessary condition for interior local extrema: $f'(x) = 0$
- Recall *Weierstrass theorem* (paraphrase): Any continuous function defined on a compact interval $[a, b]$ attains its global maximum and minimum in such an interval.
- Sufficient conditions for an interior local maximum: $f'(x) = 0$ and $f''(x) < 0$.
- Sufficient conditions for an interior local minimum: $f'(x) = 0$ and $f''(x) > 0$.
- If $f'(x) = 0$ and $f''(x) = 0$, x may or may not be a local extremum. Examples:
  - $f(x) = x^3$
  - $f(x) = x^4$
  - $f(x) = -x^4$

- Other examples: Note, first, that for differentiable functions with bounded derivatives, the Lipschitz constant is the uniform bound on the derivative.
  - Lipschitz functions: The function $f(x) = \sqrt{x^2+5}$, $x \in \mathbb{R}$, is continuous, with derivative $f'(x) = \frac{1}{2}\frac{2x}{\sqrt{x^2+5}} = \frac{x}{\sqrt{x^2+5}}$, so $|f'(x)| = \frac{|x|}{\sqrt{x^2+5}} \in [0,1)$ and $K = \sup_{x\in\mathbb{R}} |f'(x)| = 1$ is the Lipschitz constant.
  - The function $f(x) = |x|$ is Lipschitz continuous with Lipschitz constant $K = 1$, but it is not differentiable at $x = 0$ because the right derivative at zero is 1, while the left derivative is -1.
  - The function $f(x) = \sqrt{x}$, $x \in \mathbb{R}_+$, is not Lipschitz continuous, even though it is differentiable with derivative $f'(x) = \frac{1}{2\sqrt{x}}$, because $\lim_{x\searrow 0} f'(x) = +\infty$, so there does not exist a finite Lipschitz constant. Since, for $x \in \mathbb{R}_+$, $h > 0$, and $\alpha > 0$,

$$\begin{aligned}\frac{\sqrt{x+h}-\sqrt{x}}{h^\alpha} &= \frac{\sqrt{x+h}-\sqrt{x}}{h}h^{1-\alpha} \\ &\simeq \frac{1}{2\sqrt{x}}h^{1-\alpha} = \frac{1}{2}\sqrt{\frac{h}{x}}h^{\frac{1}{2}-\alpha} \\ &< \infty \text{ for all } x \in \mathbb{R}_+ \text{ and } h > 0 \text{ if } \alpha \leq \tfrac{1}{2},\end{aligned}$$

    the function $f(x) = \sqrt{x}$ is Hölder continuous for $\alpha \leq \frac{1}{2}$.
  - Exercise: Consider the function $f : [a,b] \to [0,1]$, $a < b$, defined as $f(x) = \frac{\theta(x-a)^2}{(b-a)^2} + \frac{(1-\theta)(x-a)}{b-a}$, where $\theta \in (0,1)$. Show that, for any $x_1, x_2 \in [a,b]$,

$$\frac{1-\theta}{b-a}|x_1 - x_2| \leq |f(x_1) - f(x_2)| \leq \frac{1+\theta}{b-a}|x_1 - x_2|.$$

    For the special case where $b - a = 1$, conclude that the function $f(x)$ is bi-Lipschitz and show that $K = \frac{1+\theta}{1-\theta}$ is a bi-Lipschitz constant.

1.5.1 Monotonic functions

- A differentiable function is *monotonically increasing (decreasing)* in the real interval A iff its first derivative is always non-negative (non-positive) in A: $f'(x) \geqslant (\leqslant) 0$ for all $x \in A$.
- A differentiable function is *strictly monotonically increasing (decreasing)* in the real interval A iff its first derivative is always positive (negative) in A: $f'(x) > (<) 0$ for all $x \in A$.
- But, of course, a function does not need to be differentiable, or even continuous, to be monotonic.
- Suppose you want to study a monotonic function within an interval A. You know:

- If A is a closed interval $[a, b]$, then the function attains its global maximum and minimum in A and these will be at the edges of the interval (and maybe also at neighbouring points, if the first derivative is zero around a or b).
- If A is an open interval, then its global maximum and minimum may or may not be reached within A, it will depend on whether the interval is opened on both sides and whether the derivative is zero around the edges of the interval.

1.5.2 Convex and concave functions

- Alternative characterization of convexity (concavity): A function is convex (concave) if any point on a line uniting two points on the graph of the function does not lie below (above) the graph of the function.
- A twice differentiable function is convex (concave) in the interval A iff $f''(x)$ is always non-negative (non-positive): $f''(x) \geqslant 0$ ($f''(x) \leqslant 0$).
- As a consequence, any twice differentiable convex (concave) function has a non-decreasing (non-increasing) first derivative.
- Strict convexity (concavity) is defined by changing the definitions above to strict inequalities.
- In calculus, a critical point of a function of a real variable is any value in the domain where either the function is not differentiable or its derivative is 0. The value of the function at a critical point is a critical value of the function.
- *Result:* If a twice differentiable function is convex in a closed interval A then its interior critical point(s) is a (are) global minimum(s). If a twice differentiable function is strictly convex in a closed interval A then it has unique critical point and this is its global minimum.
- *The corresponding result to concave functions:* If a twice differentiable function is concave in a closed interval A then its interior critical point(s) is a (are) global maximum(s). If a twice differentiable function is strictly concave in a closed interval A then it has unique critical point and this is its global maximum.

1.6 Using derivatives to approximate functions at a given point: Taylor's series expansion

1.6.1 Linear approximation

Consider a real $\mathcal{C}^\infty$ function $f : S \subset \mathbb{R} \to \mathbb{R}$, where S is open.

Take a point $x \in S$ and consider a small deviation from x, say $x + h$. The definition of first derivative at x is

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Rearranging this expression we can approximate $f(x + h)$,

$$f(x+h) \simeq f(x) + hf'(x)$$

This can also be written as:

$$f(x+h) = f(x) + f'(x)h + R_1(x, x+h)$$

where R is the remainder and has the following property,

$$\lim_{h \to 0} \frac{|R_1(x, x+h)|}{h} = 0,$$

i.e. $R_1(x, x + h) = o(h)$.

1.6.2 Higher order approximations

- *The Mean Value Theorem:* If f is $\mathcal{C}^1$ in $[a, b]$ $(a < b)$ then there exists a $c \in [a, b]$ such that:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

- We can use this theorem to make higher order approximations:
  - Suppose f is $\mathcal{C}^2$ around x and consider the interval $[x, x + h]$.
  - From the mean value theorem we know there exists $c \in [x, x + h]$ such that:

$$f(x+h) = f(x) + hf'(c) \tag{1}$$

  - But since f' is C^1 we can use the linear Taylor approximation to approximate $f'(c)$:

$$\begin{aligned} f'(c) &\simeq f'(x) + (c - x)f''(x) \\ &\simeq f'(x) + \frac{h}{2}f''(x), \text{ since } c \in [x, x+h]. \end{aligned}$$

  - Replacing in (1) yields:

$$f(x+h) \simeq f(x) + hf'(x) + \frac{h^2}{2}f''(x)$$

- In general, suppose f is $\mathcal{C}^n$. The nth order approximation of $f(x+h)$ is:

$$f(x+h) = f(x) + f'(x)h + \frac{1}{2}f''(x)h^2 + ... + \frac{1}{n!}f^{(n)}(x)h^n + R_n(x, x+h)$$

where the remainder

$$R_n(x, x+h) = \frac{1}{n!}\left[f^{(n)}(x^\star) - f^{(n)}(x)\right]h^n, \text{ for } x^\star \in [x, x+h],$$

has the property that $R_n(x, x+h) = o(h^n)$, or equivalently

$$\lim_{h\to 0}\frac{|R_n(x, x+h)|}{h^n} = 0.$$

Examples:

1. The exponential function $f(x) = \exp(x)$ has an infinite Taylor's series expansion about zero. Since $f'(x) = f(x) = f^{(n)}(x)$ for any $n = 1, 2, \cdots$, $f^{(n)}(0) = 1$ for any n, and $h = x - 0 = x$, it follows that

$$\exp(x) = 1 + x + \frac{1}{2}x^2 + O(x^3) = \sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

2. For $|x| < 1$, the function $f(x) = \frac{1}{1-x}$ satisfies $f'(x) = \frac{1}{(1-x)^2}$ and $f^{(n)} = (n!)\frac{1}{(1-x)^{n+1}}$ $n = 1, 2\cdots$. Hence expanding about zero so that $h = x$ and $f^{(n)}(0) = n!$ yields the infinite expansion

$$\frac{1}{1-x} = \sum_{n=0}^{\infty}x^n.$$

1.6.3 Concavity/convexity of a function revisited

- We have seen that a sufficient condition for a maximum (minimum) at $x = a$ where $f'(a) = 0$ is that $f''(a) < (>)0$.
- The first order Taylor series approximation is uninformative about the nature of the extrema if a is a turning point.
- We should go to the second order approximation. For a small h

$$f(a+h) \simeq f(a) + h * 0 + \frac{h^2}{2!}f''(a)$$

- We can now see that,

$$f(a+h) - f(a) > 0 \qquad \text{if} \qquad f''(a) > 0$$
$$f(a+h) - f(a) < 0 \qquad \text{if} \qquad f''(a) < 0$$

- In the first case, a is a local minimum and in the second case a is a local maximum.

1.7 One simple economic application of differentiation: price elasticities of demand

- Demand for a good at price p is a decreasing function $q = q(p)$
- The price elasticity of demand measures the percentage change in q when p changes by 1%.

$$\frac{dq/q}{dp/p} = \frac{dq}{dp}\frac{p}{q}$$

- Generally,

$$\epsilon_{q,p} = -\frac{dq}{dp}\frac{p}{q} = -\frac{d\ln q}{d\ln p}$$

- The total revenue is $r(q) = p * q(p)$.
- The marginal revenue is,

$$\begin{aligned} r'(q) &= p + q\frac{dq}{dp} \\ &= p\left(1 + \frac{q}{p}\frac{1}{\frac{dq}{dp}}\right) \\ &= p\left(1 + \frac{1}{\frac{dq}{dp}\frac{p}{q}}\right) \\ &= p\left(1 - \frac{1}{\epsilon_{q,p}}\right) \end{aligned}$$

2 Integral Calculus

References: Chiang, chapter 13.

Two Perspectives on Integration

1. Integration is *approximately* the reverse operation of differentiation.

   We start with a function $f'(x)$ and want to recover $f(x)$. However, $f'(x)$ is not enough information to recover $f(x)$.

2. Integration calculates the area under the graph of a function such as $f'(x)$.

2.1 The indefinite integral

Let $f : S \subset \mathbb{R} \to \mathbb{R}$ be a continuous function in S. Then a differentiable function $F(x) : S \subset \mathbb{R} \to \mathbb{R}$ is called the integral of f if

$$F'(x) = f(x) \qquad \text{or} \qquad \frac{dF(x)}{dx} = f(x).$$

The main question of integral calculus is how to find such a function.

In general, there will be more than one solution to this problem:

$$\frac{d}{dx}\left(F(x)+c\right) \text{ does not depend on } c \text{ (a constant).}$$

Thus, it is possible to obtain arbitrarily many functions that are the integral of another function, unless some other condition is given.

The function $F(x)+c$ for $c=0$ is called the *primitive* function of f.

2.2 Some integration rules

- $\int c\,dx = cx + k$,
- $\int x^n\,dx = \frac{x^{n+1}}{n+1} + k$ for $n \neq -1$,
- $\int \frac{1}{x}\,dx = \ln x + k$ for $x > 0$,
- $\int \frac{1}{x}\,dx = \ln(-x) + k$ for $x < 0$,
- $\int e^x\,dx = e^x + k$,
- $\int c^x\,dx = \frac{c^x}{\ln c} + k$, where c is a constant,
- $\int cf(x)\,dx = c\int f(x)\,dx$, where c is a constant,
- $\int \left[f(x) \pm g(x)\right]\,dx = \int f(x)\,dx \pm \int g(x)\,dx$.

However, although every bounded and continuous function has an integral, this is often impossible to explicitly write down. Examples are the functions $f(x) = \exp\{-x^2\}$ (which is proportional to the standard normal (Gaussian) density function).

2.3 The definite integral

- One attempt at a definition of integral of a function f: it is a function F such that $\frac{dF(x)}{dx} = f(x)$.
- This means that: $dF(x) = f(x)dx$.
- But $f(x)dx$ is just the area of a narrow rectangle beneath $f(x)$ of size (length) dx.
- That is, a small change in F(x) is equal to the area of the rectangle of sides $f(x)$ and dx.

- Thus we might want to ask: can we quantify the area under the graph of f over some interval in x, say $[a, b]$?
- This is exactly the concept of definite integral: the Riemann interpretation.

2.3.1 The Riemann integral

- Suppose we want to compute $\int_a^b f(x)dx$.
- Start by making a partition of the interval $[a, b]$ into $n-1$ non-overlapping subintervals of equal length by selecting an increasing sequence of numbers $\{x_1 = a, x_2, ...x_n = b\}$ such that,
  - $\bigcup_{i=1,...n-1}[x_i, x_{i+1}] = [a, b]$
  - $(x_i, x_{i+1}) \cap (x_j, x_{j+1}) = \emptyset$ for any $i \neq j$.
- Define, for $i = 1, ..., n-1$,

$$M_{ni} = \max_{x\in[x_i,x_{i+1}]} f(x) \qquad \text{and} \qquad m_{ni} = \min_{x\in[x_i,x_{i+1}]} f(x)$$

  and,

$$\Delta_n = \frac{b-a}{n-1}\sum_{i=1}^{n-1} M_{ni} \qquad \text{and} \qquad \delta_n = \frac{b-a}{n-1}\sum_{i=1}^{n-1} m_{ni}.$$

- The Riemann integral exists iff $\lim_{n\to\infty}\Delta_n = \lim_{n\to\infty}\delta_n$ in which case we define,

$$I = \int_a^b f(x)dx = \lim_{n\to\infty}\Delta_n = \lim_{n\to\infty}\delta_n.$$

- From his interpretation of integral it results,

$$\int_a^b f(x)\,dx \;=\; \int_a^c f(x)\,dx + \int_c^b f(x)\,dx, \tag{2}$$

  for any $c \in [a, b]$. I.e., integration is a linear operation.
- Recall: Linear operators commute. E.g. $\int_a^b[\alpha f(x) + \beta g(x)]dx = \alpha\int_a^b f(x)dx + \beta\int_a^b g(x)dx$.
- Adopting the generally accepted convention,

$$\int_a^b f(x)\,dx \;=\; -\int_b^a f(x)\,dx,$$

  then equation (2) holds for any c, as long as f is integrable in $[\min(a, c), \max(b, c)]$.

2.4 The fundamental theorems of calculus

2.4.1 The first fundamental theorem of calculus

As presented, the Riemann integral is not of much use if we want to actually compute the integral of a function. To make it more operational, lets study how I changes with the upper bound of the integration set. Start by defining the function,

$$F(x) = \int_a^x f(t)\, dt.$$

As x changes by a small amount, $\Delta > 0$, $F(x)$ changes in the following way,

$$F(x + \Delta) - F(x) = \int_x^{x+\Delta} f(t)\, dt,$$

which, divided by Δ and taking the limits, yields as $\Delta \to 0$,

$$F'(x) = \lim_{\Delta \to 0} \int_x^{x+\Delta} \frac{f(t)}{\Delta}\, dt.$$

Now notice that, for a small Δ we can write $\int_x^{x+\Delta} f(t)dt \approx \Delta * f(x)$, which implies

$$F'(x) = \lim_{\Delta \to 0} \int_x^{x+\Delta} \frac{f(t)}{\Delta}\, dt = \lim_{\Delta \to 0} \frac{\Delta * f(x)}{\Delta} = f(x).$$

The following theorem establishes precisely this result.

Theorem: (*1st Fundamental Theorem of Calculus*) Let $f : [a, b] \to \mathbb{R}$ be $\mathcal{C}^0$. Then every function $F(x) = \int_a^x f(t)\, dt$ is differentiable on the interval $[a, b]$ and it holds,

$$F'(x) = \frac{d}{dx} \int_a^x f(t)\, dt = f(x).$$

Now notice that using the convention,

$$\int_a^b f(x)dx = -\int_b^a f(x)dx,$$

we have

$$\begin{aligned} \frac{d}{dx} \int_x^b f(t)dt &= -\frac{d}{dx} \int_b^x f(t)dt \\ &= -f(x). \end{aligned}$$

2.4.2 The second fundamental theorem of calculus

From the study of indefinite integrals, we have learned that every integrable function has more than one possible integral. Definite integrals, however, establish the exact function we work with:

- Denote by $F(x)$ the primitive function of f

$$F(x) \quad : \quad F'(x) = f(x) \text{ and the additive constant is 0.}$$

- Denote by $\bar{F}(x)$ the definite integral between a and x. We proved that,

$$\bar{F}'(x) = f(x).$$

- But then we know that

$$\bar{F}(x) = F(x) + c$$

  as all functions that have f has first derivative satisfy this condition.

- And we also know that

$$\overline{F}(a) = \int_a^a f(t)\, dt = 0 = F(a) + c.$$

- But then, there is only one c that defines the definite integral, namely $c = -F(a)$ and it follows that,

$$\int_a^x f(t)\, dt = F(x) - F(a).$$

This is formally established in the second fundamental theorem of calculus.

Theorem: (*2nd Fundamental Theorem of Calculus*) Let $f : [a, b] \to \mathbb{R}$ be $\mathcal{C}^1$ and F be an arbitrary integral of f (in particular, it might be its primitive). Then it holds that,

$$\int_a^b f(t)dt = F(b) - F(a) = F(x)\,\Big|_a^b$$

2.5 Special integration rules

Integration by parts (this holds for both the definite and indefinite cases)

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int f'(x)g(x)\, dx$$

Integration by substitution, change of variables - indefinite case Let $x = h(t)$, one-to-one, invertible and differentiable. Then, $dx = h'(t)dt$, and so

$$\int f(x)\, dx = \int f(h(t))h'(t)\, dt$$

where h is a strictly monotonic (one-to-one) transformation.

Integration by substitution change of variables - definite case From $x = h(t)$, it follows that $a = h(\underline{t})$ for some $\underline{t}$, and so, from the invertibility of h , $\underline{t} = h^{-1}(a)$. Similarly, $\overline{t} = h^{-1}(b)$.

$$\int_a^b f(x)\, dx = \int_{h^{-1}(a)}^{h^{-1}(b)} f(h(t))h'(t)\, dt.$$

Improper integrals - measuring integrals over unbounded domains

$$\begin{aligned}\int_a^{+\infty} f(t)\, dt &= \lim_{x\to+\infty}\int_a^x f(t)\, dt \\ \int_{-\infty}^a f(t)\, dt &= \lim_{x\to-\infty}\int_x^a f(t)\, dt\end{aligned}$$

2.6 Caveats

The Riemann approach to integration suffers from some problems.

- Consider the function

$$\begin{aligned}1_{\{x\in\{\mathbb{Q}\cap[0,1]\}\}} &= 1 \text{ if } x \text{ is a rational number in } [0,1] \\ &= 0 \text{ otherwise.}\end{aligned}$$

  The area under this function is non-zero, so one would expect its integral to exist. But for any partition of $[0,1]$ into subintervals, say $[x_i, x_{i+1}]$, $i = 1, \cdots, n$,

$$\begin{aligned}M_{ni} &= \max_{x\in[x_i,x_{i+1}]} 1_{\{x\in\{\mathbb{Q}\cap[0,1]\}\}} = 1 \text{ because } [x_i, x_{i+1}] \text{ contains a rational number} \\ m_{ni} &= \min_{x\in[x_i,x_{i+1}]} 1_{\{x\in\{\mathbb{Q}\cap[0,1]\}\}} = 0 \text{ because } [x_i, x_{i+1}] \text{ contains an irrational number,}\end{aligned}$$

  and hence the Riemann bounds do not yield the same limits, so this function is not Riemann integrable.

- The Riemann approach does not work well when attempting to integrate limits of functions, i.e. integrals of the form $\int \liminf_k f_k(x)dx$.

The Lebesgue approach to integration overcomes these problems. Its definition is based on a theory of sets, measures on these sets and measurable functions.

2.7 Basics of Measure Theory

2.7.1 Intuitive Framework

Measure theory deals with collections of sets and measures (numbers) assigned to them that can be interpreted as their respective size (length, area, volume, probability).

Main applications: foundations of (Lebesgue) integration; and probability theory.

Examples:

- *Lebesgue measure* on *Euclidean space* in $\mathbb{R}$ assigns the length to suitable subsets, e.g. intervals; e.g. the set $[0, 1]$ has Lesbegue measure 1.
- *Probability measure* on sets of states of the world assigns probabilities to them; e.g. fair coin toss: two possible states of the world (heads, tails), with probability measures $\frac{1}{2}$ assigned to each.

2.7.2 Definitions

- Let X be a set. A *σ-algebra* over X is a non-empty collection Σ of subsets of X (including X itself) that is closed under complementation (i.e. if the set $A \in \Sigma$, then its complement $A^c = X \setminus A \in \Sigma$) and and countable[1] unions of its members (i.e. if $A_i \in \Sigma$, $i = 1, \cdots$, then $\cup_{i=1}^{\infty} A_i \in \Sigma$).
- The pair (X, Σ) is called a *measurable space*.
- Suppose that X is a set and Σ is a σ-algebra over X. Then a *measure* μ is a function with domain Σ and codomain $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ (the extended real line) such that the following conditions are satisfied:
  - Non-negativity: $\mu(E) \geq 0$ for all $E \in \Sigma$;
  - Null empty set: $\mu(\emptyset) = 0$;
  - Countable additivity (or σ-additivity): if $\{E_i\}_{i \in I}$, where $I \subseteq \mathbb{N}$, is a countable collection of pairwise disjoint sets in Σ (i.e. $E_i \in \Sigma$, and $E_i \cap E_j = \emptyset$ for all $i, j \in I$, $i \neq j$), then the union of all E_i has measure equal to the sum of the E_i's measure:

$$\mu\left(\cup_{i \in I} E_i\right) = \sum_{i \in I} \mu\left(E_i\right).$$

[1] A set I is countable if there exists an injective function $g : I \to \mathbb{N}$. If g is also surjective, i.e. it is a bijection, then I is countably infinite.

- The triple (X, Σ, μ) is called a *measure space*.
- A *probability measure* is a measure with total measure 1, i.e. $\mu(X) = 1$. And a *probability space* $(\Omega, \mathcal{A}, P)$ is a measure space $(\Omega, \mathcal{A})$ with probability measure P, where Ω is the *sample space*, i.e. the set of all possible outcomes of a random experiment (so that $P(\Omega) = 1$), whose subsets are called *events*, and $\mathcal{A}$ is a σ-algebra on Ω.
- Consider the measure space (X, Σ, μ). A set E is *μ-measurable* if $E \in \Sigma$. There exists some (pathological) sets that are not measurable (e.g. Vitali sets, Hausdorff paradox).
- A measurable set $E \in \Sigma$ is called a *null set* (or a *set of measure zero*) if $\mu(E) = 0$.

  Example: Consider the measure space $(\mathbb{R}, \mathcal{B}, \mu)$, where $\mathcal{B}$ is the (Borel) σ-algebra of open and closed sets on $\mathbb{R}$, and μ is Lebesgue measure. Then $\mu([0,1]) = 1$, $\mu((0,1)) = 1$, $\mu([0,1)) = 1$, $\mu((0,1]) = 1$, and hence, by countable additivity, $\mu(\{0\}) = \mu([0,1] \setminus (0,1]) = 0$, and similarly $\mu(\{1\}) = 0$.

2.7.3 Further Properties of Measures

- *Monotonicity*: If E_1 and E_2 are measurable sets with $E_1 \subseteq E_2$, then $\mu(E_1) \leq \mu(E_2)$.
- *Sub-additivity*: If $E_i \in \Sigma$ for $i \geq 1$ and $E \subset \cup_{i=1}^{\infty} E_i$, then $\mu(E) \leq \sum_{i=1}^{\infty} \mu(E_i)$.
- *Continuity from below*: If $E_i \nearrow E$, i.e. $E_1 \subset E_2 \subset \cdots$ and $\cup_{i=1}^{\infty} E_i = E$, then $\mu(E_i) \nearrow \mu(E)$.
- *Continuity from above*: If $E_i \searrow E$, i.e. $E_1 \supset E_2 \supset \cdots$ and $\cap_{i=1}^{\infty} E_i = E$, then $\mu(E_i) \searrow \mu(E)$.
- *Finiteness and σ-finiteness*: A measure space (X, Σ, μ) is called *finite* if total measure $\mu(X) < \infty$. It is called σ-finite if X can be decomposed into a countable union of measurable sets of finite measure. A set in a measure space has σ-finite measure if it can be represented as a countable union of sets with finite measure.

  Example: The measure space $(\mathbb{R}, \mathcal{B}, \mu)$ above (i.e. Borel σ-algebra on $\mathbb{R}$ and Lebesgue measure) is σ-finite, but not finite, because the real line can be partitioned into non-overlapping intervals $[n, n+1)$, $n \in \mathbb{Z}$, each of which has finite measure 1, but $\mu(\mathbb{R}) = +\infty$.

2.7.4 Lebesque Integration

In Lebesque theory, integrals are defined for *measurable functions*. Consider the measure space (X, Σ, μ). An extended real-valued function f is called *Lebesgue measurable* if its domain is measurable and it satisfies one of the following properties:

(i) For each real α, the set $\{x : f(x) > \alpha\}$ is measurable;

(ii) For each real α, the set $\{x : f(x) \geq \alpha\}$ is measurable;

(iii) For each real α, the set $\{x : f(x) < \alpha\}$ is measurable;

(iv) For each real α, the set $\{x : f(x) \leq \alpha\}$ is measurable.

Example: Consider the function $f(x) = 1_{\{x\in[0,1]\}}$. This is an indicator function for the unit interval, with codomain $\{0, 1\}$. Its preimage consists of the sets

$$\begin{aligned} f^{-1}(0) &= \mathbb{R} \setminus [0, 1] \text{ with Lebesgue measure } +\infty, \\ f^{-1}(1) &= [0, 1] \text{ with Lebesgue measure 1.} \\ \{x : f(x) > \alpha\} &= \emptyset \text{ if } \alpha > 1 \\ &= [0, 1] \text{ if } 0 < \alpha \geq 1 \\ &= \mathbb{R} \text{ if } \alpha \leq 0. \end{aligned}$$

The Lebesgue integral of any Lebesgue measurable functions can then be built up in stages.

1. Indicator functions: Let 1_S denote the indicator function of the μ-measurable set S. Then, integration with respect to the measure μ is denoted by

$$\int 1_S d\mu = \mu(S).$$

   Note that the value of this integral may be $+\infty$, e.g. as in the case of the preceding example.

2. Simple functions: A finite linear combination of indicator functions

$$\psi = \sum_{k=1}^{K} \alpha_k 1_{S_k},$$

   where the coefficients α_k are real numbers and the sets S_k are μ-measurable, $k = 1, \cdots, K$, is called a μ-measurable simple function. Then,

$$\int \psi d\mu = \int \left[\sum_{k=1}^{K} \alpha_k 1_{S_k}\right] d\mu = \sum_{k=1}^{K} \alpha_k \int 1_{S_k} d\mu = \sum_{k=1}^{K} \alpha_k \mu(S_k).$$

   And if $E \in \Sigma$, then

$$\int_E \psi d\mu = \int 1_E \psi d\mu = \sum_{k=1}^{K} \alpha_k \mu(E \cap S_k).$$

   Simple functions are useful because any bounded μ-measurable function can be approximated by simple functions. This leads to the definition of the Lebesgue integral.

3. If f is a measurable non-negative function, defined on a μ-measurable set E with $\mu(E) < \infty$, $f : E \to \mathbb{R}_+ \cup \{+\infty\}$, the Lebesgue integral of f over E is defined by

$$\int_E f d\mu = \inf \left\{ \int_E \psi d\mu : \psi \geq f, \psi \text{ simple} \right\} = \sup \left\{ \int_E \psi d\mu : 0 \leq \psi < f, \psi \text{ simple} \right\}.$$

   i.e. the infimum and supremum are taken over all simple functions ψ bounding f from above and below, respectively.

4. Singed functions: Let f be a measurable function, defined on a μ-measurable set E with $\mu(E) < \infty$, $f : E \to \mathbb{R} \cup \{\pm\infty\}$. Then, f can be decomposed as

$$f = f^+ - f^-,$$

   where

$$\begin{aligned} f^+ &= f(x) \text{ if } f(x) > 0 \\ &= 0 \text{ otherwise} \\ f^- &= -f(x) \text{ if } f(x) < 0 \\ &= 0 \text{ otherwise,} \end{aligned}$$

   so that f^+ and f^- are both non-negative, measurable functions and satisfy

$$|f| = f^+ + f^-.$$

   If $\int |f| d\mu < \infty$, then it follows that $\int f^+ d\mu < \infty$ and $\int f^- d\mu < \infty$, so

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu < \infty,$$

   i.e. f is Lebesgue integrable (the integral is finite).

   Note: The Lebesgue integral does not distinguish between two functions that differ only on a set of μ-measure zero; i.e. two functions f and g are equal almost everywhere (a.e.) if $\mu(\{x : f(x) \neq g(x)\}) = 0$.

   Example: Re-consider the indicator function

$$\begin{aligned} 1_{\{x \in \{\mathbb{Q} \cap [0,1]\}\}} &= 1 \text{ if } x \text{ is a rational number in } [0,1] \\ &= 0 \text{ otherwise.} \end{aligned}$$

   Saw before: This function is not Riemann integrable. Now consider its Lebesgue integral:

$$\int_{x \in [0,1]} 1_{\{x \in \{\mathbb{Q} \cap [0,1]\}\}} d\mu = \mu\left(\mathbb{Q} \cap [0,1]\}\right) = 0,$$

   since there are countably infinite rational numbers in the unit interval.

2.7.5 Further Properties of Lebesgue Integrals

- Lebesgue integrability: If $\int |f| d\mu < \infty$, the function f is said to be *Lebesgue integrable*.
- More generally, consider the measure space (X, Σ, μ), where μ is Lebesgue measure, and μ measurable functions $f : X \to \mathbb{R}$. Then, f is said to belong to $\mathcal{L}^p(X, \mu)$ if $\left(\int |f|^p d\mu\right)^{\frac{1}{p}} < \infty$ for $0 \leq p < \infty$.
- The definition of $\mathcal{L}^p(X, \mu)$ spaces can be extended to $p = \infty$ by requiring $|f|_\infty = \inf\{c : |f(x)| \leq c \text{ for almost all } x\} < \infty$, i.e. f must have a finite supremum, except possibly on a set of μ-measure zero (so called finite *essential supremum*); no integration is part of the definition. An L^∞ space is also called *Banach space*.
- Monotonicity: If the μ-measurable functions f and g satisfy $f \leq g$, then $\int f d\mu \leq \int g d\mu$.

2.8 Convergence Theorems

1. *Fatou's Lemma*: If f_n is a sequence of non-negative μ-measurable functions and $\liminf_n f_n(x) = f(x)$ almost everywhere on a set E, then

$$\int_E \liminf_n f_n d\mu = \int_E f d\mu \leq \liminf_n \int_E f_n d\mu.$$

- Example: Consider the sequence of functions $f_n(x) = n1_{\{x \in (0,1/n)\}}$, $n = 1, 2, \cdots$, defined on $[0, 1]$. For any n, $f_n(x) \geq 0$, and

$$\limsup_n f_n(x) = \inf_n \sup_{m>n} f_m(x) = \inf_n \left[0 \cdot 1_{\{x>1/n\}} + \lfloor 1/x \rfloor 1_{\{x \leq 1/n\}}\right] = 0,$$

  where $\lfloor z \rfloor$ denotes the integer part of $z \in \mathbb{R}$; so $\liminf_n = \limsup_n = \lim_n f_n(x) = 0$. Then,

$$\int_0^1 \liminf_n f_n(x) dx = 0 < \liminf_n \int_0^1 n1_{\{x \in (0,1/n)\}} dx = \liminf_n \int_0^{1/n} n dx = 1.$$

- Counter-example: To see why non-negativity of $f_n(x)$ is important, consider the sequence of functions $f_n(x) = -\frac{1}{n} 1_{\{x \in [n,2n]\}}$, $n = 1, 2, \cdots$, defined on $\mathbb{R}_+$. For any n, $f_n(x) \leq 0$, and $\limsup_n |f_n(x)| = \inf_n \sup_{m>n} |f_n(x)| = \inf_n \sup_{m>n} \frac{1}{n} 1_{\{x \in [n,2n]\}} = \inf_n \frac{1}{n} 1_{\{x \in [n,2n]\}} = 0$, so $\liminf_n f_n(x) = \limsup_n f_n(x) = \lim_n f_n(x) = 0$ and

$$\int_{\mathbb{R}_+} \liminf_n f_n(x) dx = 0.$$

On the other hand,

$$\begin{aligned}\liminf_n \int_{\mathbb{R}_+} f_n(x)dx &= \liminf_n \int_{\mathbb{R}_+} \left(-\frac{1}{n}\right) 1_{\{x\in[n,2n]\}} dx \\ &= \liminf_n \int_n^{2n} \left(-\frac{1}{n}\right) dx \\ &= \liminf_n (-1) \\ &= -1.\end{aligned}$$

2. *Lebesgue Monotone Convergence Theorem* : Let (X, Σ, μ) be a measure space, and $f_1, f_2, \cdots$ be a non-decreasing sequence of positive, μ-measurable functions, i.e. $f_n : \Sigma \to \mathbb{R}_+ \cup \{+\infty\}$, $n \in \mathbb{N}$, and $0 \leq f_n(x) \leq f_{n+1}(x)$ for all $x \in X$ and $n \in \mathbb{N}$. Furthermore, let $f(x) = \lim_n f_n(x)$ almost everywhere on a set $E \in \Sigma$. Then,

   (i) f is μ-measurable;

   (ii) $\lim_n \int_E f_n d\mu = \int_E \lim_n f_n d\mu = \int_E f\, d\mu$.

- Proof of (i): Consider the set $[0, \alpha]$, $\alpha > 0$. It is sufficient to show that $A := f^{-1}([0, \alpha]) \in \Sigma$.

  Note, first, that, as the sequence of functions f_n is non-negative and non-decreasing, $A_{n+1} := f_{n+1}^{-1}([0, \alpha]) \subseteq f_n^{-1}([0, \alpha]) =: A_n$ for all $n \in \mathbb{N}$. Since f_n is measurable by assumption, $A_n \in \Sigma$ for all $n \in \mathbb{N}$. Furthermore, $A = \bigcap_{n\in\mathbb{N}} A_n$, and

$$A = \bigcap_{n\in\mathbb{N}} A_n = \left[\bigcup_{n\in\mathbb{N}} A_n^c\right]^c.$$

  Since the σ-algebra Σ is closed under complementation, $A_n^c \in \Sigma$ for all $n \in \mathbb{N}$; since it is also closed under countable unions, $\bigcup_{n\in\mathbb{N}} A_n^c \in \Sigma$, and hence $A \in \Sigma$, so f is measurable.

- Proof of (ii): Fatou's Lemma implies

$$\int_E \lim_n f_n d\mu = \int_E f\, d\mu \leq \liminf_n \int_E f_n d\mu. \tag{3}$$

  Recall the definition of Lebesgue integral,

$$\int_E f\, d\mu = \sup\left\{\int_E \psi d\mu : 0 \leq \psi \leq f, \psi \text{ simple}\right\}.$$

  Since $f_n(x) \leq f(x)$ a.e. on X, by monotonicity of Lebesgue integration,

$$\left\{\int_E \psi d\mu : 0 \leq \psi \leq f_n, \psi \text{ simple}\right\} \subseteq \left\{\int_E \psi d\mu : 0 \leq \psi \leq f, \psi \text{ simple}\right\}.$$

  Therefore,

$$\begin{aligned}\int_E f_n d\mu &= \sup\left\{\int_E \psi d\mu : 0 \leq \psi \leq f_n, \psi \text{ simple}\right\} \\ &\leq \sup\left\{\int_E \psi d\mu : 0 \leq \psi \leq f, \psi \text{ simple}\right\} \\ &= \int_E f\, d\mu.\end{aligned}$$

This implies

$$\limsup_n \int_E f_n d\mu \leq \int_E f\, d\mu. \tag{4}$$

Combining (23) and (24),

$$\lim_n \int_E f_n d\mu = \int_E f\, d\mu.$$

□

- Example: Consider the sequence of non-negative functions $f_n(x) = 1_{\{x \geq 1/n\}} \exp\left(-\frac{1}{2}x^2\right)$, for $x \in \mathbb{R}$. For any n, $0 \leq f_n(x) \leq f_{n+1}(x) \leq \exp\left(-\frac{1}{2}x^2\right) =: g(x)$, which is proportional to the standard normal density and hence integrable, with $\int_{\mathbb{R}} g(x)dx = \sqrt{2\pi} < \infty$. Also, $\lim_n f_n(x) = 1_{\{x \geq 0\}} \exp\left(-\frac{1}{2}x^2\right) = f(x)$. Therefore,

$$\begin{aligned}
\int_{\mathbb{R}} \lim_n f_n(x)dx &= \int_{\mathbb{R}} 1_{\{x \geq 0\}} \exp\left(-\frac{1}{2}x^2\right) dx \\
&= \int_0^\infty \exp\left(-\frac{1}{2}x^2\right) dx \\
&= \lim_n \int_{\mathbb{R}} 1_{\{x \geq 1/n\}} \exp\left(-\frac{1}{2}x^2\right) dx \\
&= \lim_n \int_{1/n}^\infty \exp\left(-\frac{1}{2}x^2\right) dx \\
&= \frac{1}{2}\sqrt{2\pi} = \sqrt{\frac{\pi}{2}}.
\end{aligned}$$

3. *Lebesgue Dominated Convergence Theorem*: Let g be non-negative and integrable over E, and let f_n be a sequence of μ-measurable functions such that $|f_n| \leq g$ on E and, for almost all x in E, $\lim f_n(x) = f(x)$. Then,

$$\int_E \lim_n f_n d\mu = \int_E f\, d\mu = \lim_n \int_E f_n d\mu.$$

This result is also sometimes referred to as Dominated Convergence Theorem, because the sequence of functions $|f_n|$ is dominated by the integrable function g.

- Proof (using Fatou's Lemma): It is sufficient to prove the theorem for sequences of non-negative functions f_n which are uniformly dominated by an integrable function g.

Fatou's Lemma implies

$$\liminf_n \int_E f_n d\mu \geq \int_E f\, d\mu.$$

Since $g - f_n \geq 0$ for all n, Fatou's Lemma also implies

$$\liminf_n \int_E (g - f_n) d\mu \geq \int_E (g - f) d\mu.$$

Since g does not depend on n and integration is a linear operation, subtracting its integral from both sides implies

$$\begin{aligned} \liminf_n \int_E (-f_n) d\mu &\geq \int_E (-f) d\mu \\ \Rightarrow \quad \limsup_n \int_E f_n d\mu &\leq \int_E f \, d\mu. \end{aligned}$$

Combining these results,

$$\liminf_n \int_E f_n d\mu \geq \int_E f \, d\mu \geq \limsup_n \int_E f_n d\mu,$$

but since the limit inferior is no larger than the limit superior, the two must be equal, proving the theorem. □

- Example: Consider the sequence of functions $f_n(x) = \exp\left((a - \frac{1}{n})x - \frac{1}{2}x^2\right)$, $n \in \mathbb{N}$ and $1 < a < \infty$, which satisfy $\lim_n f_n(x) = \exp\left(ax - \frac{1}{2}x^2\right) =: f(x)$. For $x \geq 0$, $f_{n+1}(x) \geq f_n(x)$, while for $x < 0$, $f_{n+1}(x) < f_n(x)$, for all $n \in \mathbb{N}$, so the Monotone Convergence Theorem cannot be applied. However, for any $n \in \mathbb{N}$,

$$\begin{aligned} f_n(x) &= \exp\left(\frac{1}{2}\left(a - \frac{1}{n}\right)^2\right) \exp\left(-\frac{1}{2}\left(x - \left(a - \frac{1}{n}\right)\right)^2\right) \\ &\leq \exp\left(\frac{1}{2}a^2\right) \left[1_{\{x \geq a\}} \exp\left(-\frac{1}{2}(x-a)^2\right) + \right. \\ &\qquad \left. 1_{\{x \leq a-1\}} \exp\left(-\frac{1}{2}(x-(a-1))^2\right) + 1_{\{a-1 \leq x \leq a\}} \sqrt{2\pi}\right] \\ &=: g(x). \end{aligned}$$

Hence, $0 \leq f_n(x) \leq g(x)$ for all $n \in \mathbb{N}$, and

$$\int g(x) dx = 2\sqrt{2\pi} \exp\left(\frac{1}{2}a^2\right) < \infty,$$

i.e. g is integrable. So the Dominated Convergence Theorem can be applied,

$$\begin{aligned} \int \lim_n f_n d\mu &= \int \exp\left(ax - \frac{1}{2}x^2\right) dx \\ &= \lim_n \int_E \exp\left(\left(a - \frac{1}{n}\right)x - \frac{1}{2}x^2\right) \\ &= \lim_n \exp\left(\frac{1}{2}\left(a - \frac{1}{n}\right)^2\right) \\ &= \sqrt{2\pi} \exp\left(\frac{1}{2}a^2\right). \end{aligned}$$

- Counter-example: Re-consider the example for Fatou's Lemma. There, the inequality implied by the lemma is strict, so the Lebesgue convergence theorem does not hold. The reason is that there does not exist an integrable function $g(x)$ with the property that $0 \leq f_n(x) \leq g(x)$ for all n.

2.9 Radon-Nikodym Derivaties

- Let (X, Σ, μ) be a σ-finite measure space f a non-negative measurable function. Define

$$\lambda(E) = \int_E f d\mu, \quad E \in \Sigma.$$

  Note that λ is a measure satisfying

$$\mu(A) = 0 \;\Rightarrow\; \lambda(A) = 0.$$

  In light of the latter implication, the measure λ is said to be absolutely continuous with respect to the measure μ, denoted by $\lambda << \mu$.

- *Radon-Nikodym Theorem*: Let μ and λ be two measures on the measurable space (X, Ω) and μ be σ-finite. If $\lambda << \mu$, then there exists a non-negative measurable function f such that

$$\lambda(E) = \int_E f d\mu, \;\; E \in \Sigma,$$

  and f is unique a.e., i.e. if $\lambda(E) = \int_A g d\mu$ for all $A \in \Sigma$, then $f = g$ a.e.

- The function f is called the Radon-Nikodym derivative or density of λ with respect to μ and is denoted by $d\lambda/d\mu$.

- If $\lambda(X) = \int f d\mu = 1$ for a function f, non-negative a.e., then λ is probability measure and f is called probability density with respect to μ. If μ is Lebesgue measure on $\mathbb{R}$ in the measure space $(\mathbb{R}, \mathcal{B}, \mu)$, i.e. $d\mu = dx$, then f is the usual derivative of $\lambda(x) = F(x) = \int_{-\infty}^{x} f dx$ in calculus and the aforementioned rules for derivatives and differentiation apply.

- Example: Let $F(x) = 1 - \exp(-\theta x)$, $x \in \mathbb{R}_+$ and $\theta > 0$. Since $F : \mathbb{R}_+ \to [0, 1]$ and $F(x)$ is monotonically increasing, F is a cumulative distribution function (cdf) on $\mathbb{R}_+$. Its derivative is $f(x) = \theta \exp(-\theta x)$, $x \in \mathbb{R}_+$, and so

$$F(x) = \int_0^x f(z) dz, \;\; x > 0.$$

  Let P be probability measure corresponding to F. Then, $P(E) = \int_E f(x) dx$ for any Borel set E on $\mathbb{R}_+$, and dx is Lebesque measure on $\mathbb{R}_+$.

  Notice that P is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}_+$: An atom x_0 has Lebesgue measures zero, as shown above, and the probability that a continuously distributed random variable X with density $f(x)$ equals x_0 is zero, $P(x_0) = \int_{x_0}^{x_0} f(x) dx = 0$.

3 Linear Algebra

References: Simon and Blume, chapters 6 to 11, 16 and 23.

The methods and concepts of linear algebra allow us to capture and analyze interlinked *linear* economic systems of arbitrarily many variables - however complex.

Proceed in two parts:

1. Part 1: Matrices.
2. Part 2: Linear Vector Spaces.

3.1 Some definitions

- A *matrix* of dimension $m \times n$ is a rectangular array of numbers consisting of m rows and n columns,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ik} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mk} & \dots & a_{mn} \end{pmatrix} = [a_{ik}]_{i=1,\dots m;k=1,\dots n}$$

- Each entry a_{ik} is an *element* of the matrix $A_{m\times n}$ where the first index is for the row and the second index is for the column.
- Two matrices of the same dimension, $A_{m\times n}$ and $B_{m\times n}$, are equal if all corresponding elements are equal, $a_{ik} = b_{ik}$, for all $i = 1, \dots m$ and $k = 1, \dots n$ and where a_{ik} and b_{ik} are the generic elements of A and B, respectively.
- A *vector* is a matrix in which one of the dimensions is one. If $m = 1$, the vector is a *row vector*. If $n = 1$ the vector is a *column vector*. Convention: Vectors are defined a column vectors.
- A matrix with $m = n$ is a *square matrix*.
- A square matrix with $a_{ik} = a_{ki}, \forall_{i=1,..n} \forall_{k=1,\dots m}$ is a *symmetric matrix*.
- A square matrix with $a_{ik} = 0$ whenever $i \neq k$ is a *diagonal matrix*.
- A diagonal matrix with $a_{ii} = 1$ for all $i = 1, \dots n$ is the *identity matrix* and we write I_n.

In what follows, and unless otherwise explicitly stated, capital letters represent matrices or vectors and the respective small letters represent their elements. Moreover, α is a scalar.

3.2 Operations with matrices

- *Sum:* $A_{m\times n} + B_{m\times n} = [a_{ik} + b_{ik}]_{i=1,\ldots m;k=1,\ldots n}$
- *Product by a scalar:* $\alpha A_{m\times n} = [\alpha a_{ik}]_{i=1,\ldots m;k=1,\ldots n}$
- *Matrix product:* $A_{m\times n}B_{n\times l} = \left[\sum_{j=1}^{n} a_{ij}b_{jk}\right]_{i=1,\ldots m;k=1,\ldots l}$
- *Transpose:* $A'_{m\times n} = [a_{ik}]'_{i=1,\ldots m;k=1,\ldots l} = [a_{ki}]_{i=1,\ldots m;k=1,\ldots l}$. Notice that

$$\begin{aligned}(A+B)' &= A'+B'\\ (AB)' &= B'A'\end{aligned}$$

3.3 Linear Vector Spaces

- A *linear vector space* $\mathbb{V}$ is a non-empty set of elements called vectors with two laws of combination, vector addition and scalar multiplication, satisfying the following axioms: If $u, v, w \in \mathbb{V}$ and a, b are scalars, then
  1. $\mathbb{V}$ is closed under vector addition: $u + v \in \mathbb{V}$;
  2. vector addition is commutative: $u + v = v + u$;
  3. vector addition is associative: $(u + v) + w = u + (v + w)$;
  4. there is a zero vector, $0 \in \mathbb{V}$, such that $u + 0 = 0$;
  5. $\mathbb{V}$ is closed under scalar multiplication: $av \in \mathbb{V}$;
  6. scalar multiplication is distributive: $a(u + v) = au + av$, and $(a + b)u = au + bu$;
  7. scalar multiplication is associative: $(ab)u = a(bu)$;
  8. $0u = 0, 1v = v$.
- Henceforth, the set of all real-valued $N \times 1$ vectors v will be denoted by $\mathbb{R}^N$.
- A non-empty subset $\mathbb{S}$ of a vector space $\mathbb{V}$ is called a *subspace* of $\mathbb{V}$ if for all $u, v \in \mathbb{S}$ and all scalars a, b, $au + bv \in \mathbb{S}$.
- Let $\mathbb{W}$ be a subset of a vector space $\mathbb{V}$. The *subspace spanned by* $\mathbb{W}$ is the set of all linear combinations of the vectors in $\mathbb{W}$.

- A particular *linear spanned subspace* is the *column space of a matrix* X of dimension $N \times k$, where $N > k$ and $\text{rk}(X) = k$:

$$\text{Col}(X) = \{z \in \mathbb{R}^N : z = X\alpha \text{ for some } \alpha \in \mathbb{R}^k\}.$$

- Let $\mathbb{S}_1$ and $\mathbb{S}_2$ be subspaces of a vector space $\mathbb{V}$. Then, the sum of the subspaces, $\mathbb{S}_1 + \mathbb{S}_2$, is defined to be the set of all vectors of the form $v_1 + v_2$, where $v_1 \in \mathbb{S}_1$ and $v_2 \in \mathbb{S}_2$. If $\mathbb{S}_1 \cap \mathbb{S}_2 = \{0\}$, then $\mathbb{S}_1 + \mathbb{S}_2$ is called the *direct sum* of $\mathbb{S}_1$ and $\mathbb{S}_2$, denoted by $\mathbb{S}_1 \oplus \mathbb{S}_2$.

- A finite set $\mathbb{W}$ of linearly independent vectors is a basis for the linear vector space $\mathbb{V}$ if $\mathbb{W}$ spans $\mathbb{V}$.

  Examples:

  1. Consider the matrix X of dimension $N \times k$, where $N > k$ and $\text{rk}(X) = k$. Its columns are a basis for $\text{Col}(X)$.
  2. The unit vectors in $\mathbb{R}^N$ - i.e. the $n \times 1$ vectors $u_1 = (1, 0, \cdots)'$, $u_2 = (0, 1, 0, \cdots)'$, $\cdots$, $u_n = (0, \cdots, 1)'$ - are a basis for the linear vector space $\mathbb{R}^N$. This space is therefore equivalent to $\text{Col}(I_n)$.

- The number of vectors forming a basis of a linear vector space is unique and called the *dimension of the vector space*.

  Example: The dimension of $\text{Col}(X)$ as defined above is $k = \dim(\text{Col}(X))$.

- *Cartesian Products* of vector spaces join their vectors. Example: $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x_1, x_2)' : x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$.

3.4 Inner products and orthogonality

- Let $\mathbb{V}$ be a real vector space. An *inner product* is a scalar function defined on $\mathbb{V} \times \mathbb{V}$. For every $u, v, w \in \mathbb{V}$, the inner product $u'v$ has four properties:

  1. $u'v = v'u$;
  2. $(u + w)'v = u'v + w'v$;
  3. $(au)'v = a(u'v)$;
  4. $u'u \geq 0$, and $u'u = 0$ iff $u = 0$.

- *Cauchy-Schwarz Inequality*: For every u, v in a real linear vector space that admits inner products,

$$(u'v)^2 \leq (u'u)(v'v);$$

equality holds iff $u = av$ for some scalar a or $u = 0$ or $v = 0$.

Exercise: Prove the Cauchy-Schwarz inequality.

- *Orthogonality*: Let u, v be vectors in $\mathbb{V}$. If $u'v = 0$, then u and v are said to be orthogonal. This is often denoted by $u \perp v$.
- *Orthogonality of a matrix*: An $n \times n$ matrix A is orthogonal if $A'A = D$, where D is a diagonal matrix with $d_{ii} = A'_{\cdot i}A_{\cdot i} = ||A_{\cdot i}||^2$, $i = 1, \cdots, n$.
- *Orthonormality of a matrix*: An $n \times n$ matrix A is orthonormal if $A'A = I_n$.
- *Orthogonal basis*: An orthogonal basis is a basis with mutually orthogonal vectors.

  Example: The unit vectors in $\mathbb{R}^N$ form an orthogonal basis for $\mathbb{R}^N$.
- *Orthogonal complement*: The linear subspace of vectors orthogonal to a k-dimensional subspace $\mathbb{S}$ of an N-dimensional vector space $\mathbb{V}$ is defined as

$$\mathbb{S}^{\perp} = \{u \in \mathbb{V} : u'v = 0 \; \forall v \in \mathbb{S}\}.$$

  The linear space $\mathbb{S}^{\perp}$ is the orthogonal complement of $\mathbb{S}$.

  Moreover, $\mathbb{V} = \mathbb{S} \oplus \mathbb{S}^{\perp}$, and $\dim(\mathbb{V}) = N = \dim(\mathbb{S}) + \dim(\mathbb{S}^{\perp}) = k + (N - k)$.

3.5 Normed Linear Vector Spaces

- A normed linear vector space is a vector space $\mathbb{V}$ and a real-valued scalar function on all the vectors $v \in \mathbb{V}$, denoted $\|v\|$ and called the norm of v, such that
  1. $\|av\| = |a|\|v\|$ for every scalar a;
  2. $\|v\| \geq 0$, and $\|v\| = 0$ iff $v = 0$;
  3. $\|u + v\| \leq \|u\| + \|v\|$ for every $u, v \in \mathbb{V}$ (triangle inequality).
- One common norm is the Euclidean length in $\mathbb{V} = \mathbb{R}^N$,

$$\|v\| = \sqrt{v'v} = \sqrt{v_1^2 + \cdots + v_N^2}.$$

3.6 Systems of equations

- A system of m equations in n variables can be concisely represented in a matrix format,

$$Ax \;\; = \;\; b$$

  where A is a $m \times n$ matrix, x is an $n \times 1$ vector and b an $m \times 1$ vector.

- How are systems of equations solved?
  - A possibility is to use the techniques of univariate calculus to solve sequentially each of the simultaneous equations.
  - This is equivalent to use Gaussian (or Gauss-Jordan) elimination.
- *Under what conditions does the system $Ax = b$ have a unique solution?*

  Given a system of m equations and n unknowns, there are three possible cases:
  - there is a unique solution;
  - there is no possible solution;
  - there are infinitely many solutions.

  We can use the concept of *linear dependence* to better understand the sort of system being studied:
  - Take a collection of vectors in $\mathbb{R}^n$, $V_1, ..., V_k$. A *linear combination* of these vectors is the sum of their scalar products using some scalars, $\alpha_1, ..., \alpha_k$: $\alpha_1 V_1 + ... + \alpha_k V_k$.
  - A collection of vectors in $\mathbb{R}^n$, $V_1, ..., V_k$, is said to be *linearly independent* (*linearly dependent*) if there is no (there is a) collection of scalars $\alpha_1, ..., \alpha_k$ with some $\alpha_i \neq 0$ such that the linear combination of $V_1, ..., V_k$ is the null vector.

  We can now state (part of) the answer above in terms of the linear dependence of the rows of A. A system of m equations and n unknowns will have:
  - a unique solution iff the number of linearly independent rows in A is exactly n (the number of unknowns);
  - otherwise the system may have either infinitely many solutions or none.
- *Note:* In a square matrix, all the columns (rows) are linearly independent iff all its diagonal elements are non-zero after Gaussian elimination.

3.6.1 Matrix inversion

- Take a *square* matrix $A_{n \times n}$. If it exists, the inverse of A is the matrix $A^{-1}_{n \times n}$ such that $AA^{-1} = A^{-1}A = I_n$.
- We can use Gaussian elimination to check whether A is invertible and to compute its inverse.
- *Result:* A square matrix A is invertible iff all its rows (columns) are linearly independent.

- The number of linearly independent columns (rows) of a matrix is the *rank* of the matrix. The number of linearly independent rows equals the number of linearly independent columns.
- If A is $m \times n$ then $\text{rk}(A) \leq \min(m, n)$.
- A square matrix A_n is full rank if $\text{rank}(A) = n$.
- *Result*: A matrix is full-rank iff it is invertible.
- An $n \times n$ matrix A is said to be *singular* if it is rank deficient, i.e. if $\text{rk}(A) < n$; it is *non-singular* if it has full rank, $\text{rk}(A) = n$.
- We now repeat the question: *Under what conditions does the system $Ax = b$ have a unique solution when there are n equations and n unknowns?*

  And the new insight is that for square systems of equations a unique solution exists when A is non-singular. If A is singular then either the system has no solution or it has infinitely many solutions.

  To see why, take a square system of n equations in n variables.

  – If A is invertible then,

  $$\begin{aligned} Ax = b &\Leftrightarrow A^{-1}Ax = A^{-1}b \\ &\Leftrightarrow I_n x = A^{-1}b \\ &\Leftrightarrow x = A^{-1}b \end{aligned}$$

  and this is the only solution.

  – If A is not invertible, then there exists a linear combination of the columns (and of the rows) of A that yields a null vector of dimension n,

  $$\exists c_1, ..., c_n \text{ not all zero} \quad : \quad \sum_{j=1}^{n} c_i A_{\cdot i} = 0$$

  where $A_{\cdot i}$ is the ith column of A. But then the system $Ax = 0$ has a non-zero solution $x = c$.

  Now suppose $Ax = b$ has one solution, $x = x_0$. Then it has an infinite number of solutions $x = x_0 + kc$ for any $k \in \mathcal{R}$.

  But it is also possible that $Ax = b$ has no solution.

- Aside: (*Partitioned Inverse*) Let A be a nonsingular matrix, and partition it as

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

where A_{11} and A_{22} are square, non-singular submatrices of A. Then,

$$\begin{aligned} A^{-1} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} W^{-1} & -W_{12}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}W^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}W^{-1}A_{12}A_{22}^{-1} \end{bmatrix}, \end{aligned}$$

where $W = A_{11} - A_{12}A_{22}^{-1}A_{21}$.

3.7 The determinant

3.7.1 Definitions

Lets consider a square matrix $A_{n\times n}$.

- One way of computing the determinant can be derived by induction.
- First need to define *minor* (i,j)th of A: $M_{ij} = (-1)^{i+j}|A_{ij}|$ where A_{ij} equals matrix A after excluding row i and column j.
- We can now compute $|A|$,

$$\begin{aligned} |A| &= \sum_{i=1}^{n} a_{ij}M_{ij} \quad \text{for any } j \\ &= \sum_{j=1}^{n} a_{ij}M_{ij} \quad \text{for any } i \end{aligned} \tag{5}$$

- Special case: If A is a 2×2 matrix, then $|A| = a_{11}a_{22} - a_{12}a_{21}$.

3.7.2 Some properties of the determinant

- $|A'| = |A|$.
- $|AB| = |A||B|$.
- $|I_n| = 1$ for any $n > 0$.
- $|\alpha A_{n*n}| = |\alpha|^n|A_{n*n}|$ where α is a scalar.
- If A has a row/column of zeros then $|A| = 0$. Note that, in this case, the matrix is rank deficient, i.e. not full rank. Hence, $|A| = 0$ if rk(A) is full (necessity). It is also the case that $|A| = 0$ only if rk(A)(sufficiency). To see this, not that, if A_n is not full-rank, then one of its

rows, say row i, can be expressed as a linear combination of A's other rows. By subtracting such linear combination from row i we obtain a matrix A^* for which $|A^*| = |A|$. But as the ith row of A^* is zero, it follows that $|A^*| = |A| = 0$.

On the contrary, if all rows in A are linearly independent, no row i will be replicated by a linear combination of the other rows of A and $|A| \neq 0$.

- If D is a diagonal matrix, i.e. $D = \text{diag}\{d_1, \cdots, d_n\}$, then $|D| = \prod_{i=1}^{n} d_i$. This is easy to verify by induction, using the formulae for determinants (5): For example,

$$\begin{aligned} |D_{3*3}| &= \sum_{i=1}^{3} d_{1j} M_{1j} \\ &= d_{11} \left| \begin{bmatrix} d_{22} & 0 \\ 0 & d_{33} \end{bmatrix} \right| \\ &= d_{11} d_{22} d_{33}. \end{aligned}$$

- If a row/column of A are multiplied by the constant α to obtain A^* then $|A^*| = \alpha|A|$. To see this , consider the $n \times n$ matrix A and the commensurate diagonal matrix $D(\alpha)_k = \text{diag}\{1, \cdots, 1, \alpha, 1, \cdots\}$, where $D_{kk} = \alpha$, i.e. α is the kth diagonal element. Then, $D(\alpha)_k A$ multiplies the kth row of A by α, and $AD(\alpha)_k$ multiplies the kth column by α. Hence, since $|D(\alpha)_k| = 1^{n-1}\alpha = \alpha$ by the preceding property of determinants, if follows from the second property above that

$$\begin{aligned} |D(\alpha)_k A| &= |D(\alpha)_k||A| = \alpha|A|, \\ |AD(\alpha)_k| &= |A||D(\alpha)_k| = \alpha|A|. \end{aligned}$$

- If a multiple of one row/column of A is added to another row/column of A to obtain A^* then $|A^*| = |A|$.

- If A has two identical rows/columns, it is rank deficient and, hence, $|A| = 0$.

3.7.3 Further results

- *Result*: a square matrix A is invertible iff $|A| \neq 0$.

  Suppose A_n is invertible. Then $\exists A^{-1} : A^{-1}A = I_n$. Using the properties of determinants,

$$|A^{-1}A| = |A^{-1}||A| = |I_n| = 1$$

  But then, $|A| \neq 0$ (and $|A^{-1}| \neq 0$).

  The reverse implication is obvious from one of the formulas to compute the inverse of A. However, because this is a rather impractical form of computing A^{-1}, we are not going to see it here.

- Thus: A invertible $\Leftrightarrow$ A full ranked (i.e. non-singular) $\Leftrightarrow$ $|A| \neq 0$.

3.8 Eigenvectors and Eigenvalues

Eigenvalues and eigenvectors are important for matrix decompositions and transformations, and they are informative about properties of matrices (such as rank, non-singularity etc.).

Again, lets consider a square matrix A of dimension $n \times n$.

- An *eigenvalue* is a number λ which, if subtracted from each diagonal element of A, transforms A into a singular matrix.
- That is, the λ's are the n not necessarily distinct solutions to the characteristic equation or polynomial: $|A - \lambda I_n| = 0$. Note: This is a polynomial equation of order n in λ, i.e. it has n not necessarily distinct solutions. Notice also that there might be complex solutions to this equation.
- *Result:* A square matrix A is singular iff $\lambda = 0$ is one of its eigenvalues. The reason is that $\lambda = 0$ being a solution to the determinantal equation $|A - \lambda I_n| = 0$ implies $|A| = 0$, so the result follows by from $|A| \neq 0 \Leftrightarrow A$ non-singular.
- An *eigenvector* V of A associated with its eigenvalue λ is a non-zero solution to $(A - \lambda I)V = 0$.
- Notice that $AV = \lambda V$ implies that βV is also an eigenvector associated with λ, for any $\beta \neq 0$. Hence, there will be a continuum of solutions to the system of equations $(A - \lambda I)V = 0$. One will need to choose one of these (different from the null vector) to be the eigenvector associated with eigenvalue λ. Typically, this is done by a normalization, e.g. to require that the n eigenvectors associated with the n eigenvalues have length one (see definition of vector length below).
- Note that, even if the non-zero eigenvalues of A are not distinct, the associated eigenvectors are. For example, the matrix $A = I_n$ has all eigenvalues equal to 1, and the associated eigenvectors are the unit vectors e_i in $\mathbb{R}^n$, where e_i is a vector with 1 in the ith position and zero elsewhere, $i = 1, \cdots, n$.
- Note also that, if the $n \times n$ matrix A is nonsingular, then the eigenvectors V_i, $i = 1, \cdots, n$ are linearly independent. To see this, suppose not, i.e. there exists an eigenvector V_k which can be expressed as linear combination of the remaining $n - 1$ eigenvectors, $V_k = \sum_{i \neq k} \alpha_i V_i$, for some

scalars α_i, $i \neq k$. Then,

$$\begin{aligned} AV_k &= \lambda_k V_k \\ &= \lambda_k \sum_{i \neq k} \alpha_i V_i \\ &= \sum_{i \neq k} \lambda_k \frac{\alpha_i}{\lambda_i} \lambda_i V_i \text{ (because } \lambda_i \neq 0 \text{ for } A \text{ nonsingular)} \\ &= \sum_{i \neq k} \lambda_k \frac{\alpha_i}{\lambda_i} AV_i \\ &= A \sum_{i \neq k} \beta_i V_i, \end{aligned}$$

where $\beta_i = \lambda_k \frac{\alpha_i}{\lambda_i}$, $i \neq k$. But this is a contradiction to the nonsingularity of A.

- Combining $AV_i = \lambda_i V_i$, $i = 1, \cdots, n$, as the matrix $AP = PD$, where $P = [V_1 \ldots V_n]$ is nonsingular (by the preceding result) and $D = \text{diag}\{\lambda_1, ..., \lambda_n\}$, it follows that $A = PDP^{-1}$, which is (a special case of) the *Schur decomposition* of a square matrix.

3.9 The relationship between the determinant of a matrix and its eigenvalues and eigenvectors

Under some conditions, one can diagonalise a square matrix. To arrive at this result, consider, first, the case where the $n \times n$ symmetric matrix A is nonsingular.

- Let $\lambda_1, ..., \lambda_n$ be the n eigenvalues of A, possibly not all different, but all non-zero.
- Let P be the $n \times n$ matrix composed of the n eigenvectors of A: $P = [V_1 \ldots V_n]$.
- Since V_i is such that,

$$(A - \lambda_i I)V_i = 0, \quad i = 1, \cdots, n$$

then,

$$AV_i = \lambda_i V_i, \quad i = 1, \cdots, n.$$

- Orthogonality of eigenvectors: The preceding equations imply

$$\begin{aligned} V_i' A' &= V_i' A = \lambda_i V_i', \quad i = 1, \cdots, n \\ V_i' A V_j &= \lambda_i V_i' V_j \\ &= \lambda_j V_i' V_j, \text{ (by symmetry)}, \end{aligned}$$

so, if $\lambda_i \neq \lambda_j$, then $V_i'V_j = 0$. If $\lambda_i = \lambda_j$, $i \neq j$, then the preceding equations and the normalization of $||V_i|| = 1$ for all i imply $V_i'AV_i = V_j'AV_j$, and it turns out that the associated eigenvectors are also mutually orthogonal. The general result follows from:

$$\begin{aligned} AP &= PD \Rightarrow P'AP = P'PD \\ P'A &= DP' \Rightarrow P'AP = DP'P \\ \Rightarrow \; P'PD &= DP'P = (P'PD)', \end{aligned}$$

i.e. the matrix $DP'P$ is symmetric. Since D is diagonal, this implies that $P'P$ must be diagonal, with the lengths of the eigenvectors as diagonal elements. Since the eigenvectors are normalized to have length 1, this implies $P'P = I_n$.

Hence, for $P = [V_1 \dots V_n]$ as above,

$$P'P = I_n \quad (\text{"orthonormality" of eigenvectors})$$

- Note that the orthonormality of the matrix of eigenvectors of a non-singular matrix implies $P^{-1} = P'$.
- It follows that $A = PDP^{-1} = PDP'$, the *eigenvalue decomposition* of the symmetric, nonsingular matrix A.
- Similarly, $AP = PD$ and $P^{-1} = P'$ imply $P'AP = D$, so the transformation P is said to diagonalize the matrix A.
- Aside: This generalizes to symmetric, singular $n \times n$ matrices A. Suppose that $\text{rk}(A) = m < n$. Then, $n - m$ eigenvalues of A are zero, and A has m nonzero eigenvalues and associated eigenvectors. The eigenvalue decomposition of A is then of the above form, $A = PDP'$, but now D is an $m \times m$ diagonal matrix and P is an $n \times m$ matrix containing the m eigenvectors of A, satisfying the orthonormality property $P'P = I_m$.
- For A nonsingular, from $AP = PD$ it follows that $|AP| = |A||P| = |PD| = |P||D|$, i.e. $|A| = |D|$ (because $|P| \neq 0$ for A nonsingular).
- This implies the *Result:* The determinant of A is the product of its eigenvalues: $|A| = \prod_{i=1}^{n} \lambda_i$ (*matrix spectrum*).
- This, in turn, implies the following *Result:* A square matrix A is full-rank iff all eigenvalues are different from zero.
- The trace of a square $n \times n$ matrix A is the sum of its diagonal elements eigenvalues: $\text{tr}(A) = \sum_{i=1}^{n} a_{ii}$.

- Aside: The trace operator is a linear operator. Hence, it commutes with other linear operators. In particular,

  (i) $\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$, for A and B square matrices of the same dimension;

  (ii) $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$ for matrices A, B and C for which the products inside the trace operator are permissible and yield a square matrix.

- *Result:* $\text{tr}(A) = \sum_{i=1}^{n} \lambda_i$.

  Exercise: Prove this.

- Let λ_i, $i = 1, \cdots, n$, be the eigenvalues of the $n \times n$ matrix A. Then, $\gamma_i = 1 - \lambda_i$, $i = 1, \cdots, n$, are the eigenvalues of the matrix $I_n - A$.

  This follows from

  $$0 = |A - \lambda_i I_n| = |A - I_n + (1 - \lambda_i) I_n| = (-1)^n |I_n - A - (1 - \lambda_i) I_n|,$$

  and therefore $0 = |I_n - A - (1 - \lambda_i) I_n|$, $i = 1, \cdots, n$.

3.10 Quadratic forms

3.10.1 Definition

- A *quadratic form* is a particular function in $\mathbb{R}^n$ defined as follows

  $$f_A(x) = x'Ax = \sum_{i,k=1}^{n} a_{ik} x_i x_k = x'A'x = \frac{1}{2} x'(A + A')x$$

  where A is a square $n \times n$ matrix.

- Notice that the above shows that, if A is not symmetric, one could always define exactly the same function in terms of a symmetric matrix. Thus, in what follows we always consider A to be a symmetric $n \times n$ matrix, and since f_A is completely specified by A, A is sometimes called a quadratic form.

3.10.2 Definiteness of quadratic forms

- A quadratic form A is said to be,

  - *Positive (negative) definite*: if $\forall x \in \mathbb{R}^n, x \neq 0, x'Ax > (<) 0$
  - *Positive (negative) semidefinite*: if $\forall x \in \mathbb{R}^n, x \neq 0, x'Ax \geqslant (\leqslant) 0$

- *Indefinite*: if neither of the above holds.

- *Result*: Let A be a symmetric (square) matrix,
  - A is positive (negative) definite iff all its eigenvalues are positive (negative).
  - A is positive (negative) semidefinite iff all its eigenvalues are non-negative (non-positive).
  - A is indefinite iff neither of the above holds.

  To see why, use the result on symmetric, nonsingular matrices. We know that since A is symmetric, then $\exists P$ orthogonal : $P'AP = D$, and P nonsingular.

  Take *any* $n \times 1$ vector y and define $x = Py$. Then,

$$\begin{aligned} 0 \leq x'Ax &= (Py)'APy \\ &= y'P'APy \\ &= y'Dy \\ &= \sum_{i=1}^{n} \lambda_i y_i^2, \end{aligned}$$

  for any y, which immediately leads us to the result.

3.11 Decompositions of matrices

- *Cholesky Decomposition*: Every $n \times n$, positive definite and symmetric (pds) matrix A can be factored into the matrix product CC', where C is lower - left triangular (and hence nonsingular), i.e.

$$C = \begin{pmatrix} c_{11} & 0 & \dots & 0 \\ c_{21} & c_{22} & 0 & \dots \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}$$

- Example: Consider the matrix

$$\Omega = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix},$$

  where $|\rho| \in [0, 1)$. Then,

$$\begin{aligned} C &= \Omega^{\frac{1}{2}} \\ &= \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}. \end{aligned}$$

- *Eigenvalue Decomposition* (see above): Every $n \times n$ real, symmetric matrix A with eigenvalues λ_i and eigenvectors V_i, $i = 1, \cdots, n$, can be factored into the matrix product $V'DV$, where $D = \text{diag}\{\lambda_1, \cdots, \lambda_n\}$ and $V = [V_1, \cdots, V_n]$ and $V'V = I_n$. This decomposition does not require A to be non-singular.

- Example: In the preceding example, for $\rho = 0$, the eigenvalues of Ω are $\lambda_1 = \sigma_1^2$ and $\lambda_2 = \sigma_2^2$, both positive as Ω is pds, and the associated eigenvectors are the unit vectors in $\mathbb{R}^2$. If ρ tends to 1, Ω becomes singular. In this case, $\lambda_1 = \sigma_1^2 + \sigma_2^2$ and $\lambda_2 = 0$, and the eigenvector $V_1 = \frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}}(\sigma_1, \sigma_2)'$.

  Note: If A is pds, so that $\text{rk}(A) = n$ and $\lambda_i > 0$ for $i = 1, \cdots, n$, then the eigenvalue decomposition yields another factorization of A of the type $A = CC'$, where now $C = V'D^{\frac{1}{2}}$ and $D^{\frac{1}{2}} = \text{diag}\{\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}\}$.

- *Singular Value Decomposition*: This is a generalization of the eigenvalue decomposition for any real matrix A of dimensions $m \times n$, i.e. not necessarily square. Suppose A is an $m \times n$ matrix with rank $r \leq \min\{m, n\}$. Then, the decomposition

$$A \;=\; UDV'$$

  holds, where U is an $m \times r$ matrix, D is a $r \times r$ diagonal matrix with positive, non-increasing elements down the diagonal, and V' is an $r \times n$ matrix; U and V are orthonormal, i.e, $U'U = V'V = I_r$.

- Matrix decompositions as above are often useful in establishing rank properties of matrices.

  Example: A square $n \times n$ matrix A is idempotent if $AA = A$. *Result:* If the $n \times n$ matrix A is idempotent, then $\text{rk}(A) = \text{tr}(A)$.

  To prove this, note that $A = UDV'$, where U and V are $n \times r$ orthonormal matrices and D is an $r \times r$ diagonal matrix, with $r = \text{rk}(A) \leq n$ and positive diagonal elements. Then, idempotency of A implies

$$A = UDV' = AA = UDV'UDV',$$

  and orthonormality of U and V implies the diagonalization

$$D = DV'UD.$$

  Define the $r \times r$ matrix $B = DV'U$, so that the last equation can be written as $D = BD$. Since D is diagonal, this implies $B_{ii}D_{ii} = D_{ii}$, $i = 1, \cdots, r$. Since $D_{ii} > 0$ for $i = 1, \cdots, r$, it

must be that $B_{ii} = 1$ for $i = 1, \cdots, r$. Hence,

$$\begin{aligned} \text{tr}(A) &= \text{tr}(UDV') \\ &= \text{tr}(DV'U) \\ &= \text{tr}(B) \\ &= \sum_{i=1}^{r} B_{ii} \\ &= r \\ &= \text{rk}(A). \end{aligned}$$

□

3.12 Orthogonal Projection

This section provides some supplementary material that is useful for the geometric interpretation of linear regression.

For simplicity, the following exposition refers to Euclidean space $\mathbb{V} = \mathbb{R}^N$.

- *Pythagorean Theorem*: If $v_1, v_2 \in \mathbb{R}^N$ and $v_1 \perp v_2$, then

$$\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2.$$

- Consider again the $N \times k$ matrix X with $\text{rk}(X) = k$. Saw above: X is a basis of $\text{Col}(X)$, which is a k-dimensional linear subspace of $\mathbb{R}^N$, which, in turn, can be decomposed as $\mathbb{R}^N = \text{Col}(X) \oplus \text{Col}(X)^\perp$.

- What is $\text{Col}(X)^\perp$? From the direct sum property of linear vector spaces, $\text{Col}(X)^\perp = \mathbb{R}^N \ominus \text{Col}(X) = \{v \in \mathbb{R}^N : v \neq X\alpha \text{ for any } \alpha \in \mathbb{R}^k\}$. Hence, elements of $\text{Col}(X)^\perp$ cannot be represented as linear combinations of the columns of X, so they must be orthogonal to them: $\text{Col}(X)^\perp = \{v \in \mathbb{R}^N : X'v = 0\}$, since then $\alpha'X'v = 0 = v'X\alpha$ and so v cannot lie in $\text{Col}(X)$ (unless it is the zero vector).

  Then, any $w \in \mathbb{R}^N$ can be decomposed as $w = v_1 + v_2$, where $v_1 \in \text{Col}(X)$ and $v_2 \in \text{Col}(X)^\perp$, and $\|w\|^2 = \|v_1\|^2 + \|v_2\|^2$.

- Consider the $k \times k$ matrix $P_X = X(X'X)^{-1}X'$, where the inverse exists because $\text{rk}(X'X) = k$. Notice that P_X satisfies $P_X P_X = P_X$, so P_X is said to be *idempotent*. It is also symmetric, $P_X = P_X'$.

- Since $v \in \text{Col}(X)$ satisfies $v = X\alpha$ for some $\alpha \in \mathbb{R}^k$, for any $v \in \text{Col}(X)$ it follows that $P_X v = P_X X\alpha = X(X'X)^{-1}X\alpha = X\alpha = v$, so pre-multiplication by P_X preserves every $v \in \text{Col}(X)$.
- Since $v \in \text{Col}(X)^\perp$ satisfies $X'v = 0$, for any $v \in \text{Col}(X)^\perp$ it follows that $P_X v = X(X'X)^{-1}X'v = 0$, so pre-multiplication by P_X annihilates every $v \in \text{Col}(X)^\perp$.
- Therefore, P_X is called the *orthogonal projector* onto $\text{Col}(X)$. Next to symmetry and idempotency, orthogonal projectors possess the property that they are unique.

4 Multivariate Calculus

References: Simon and Blume, chapters 13 to 15.

We will now study functions of several variables. In general, we are interested in functions f mapping $\mathbb{R}^n$ in $\mathbb{R}^m$ for some $n, m > 0$. We write $f : \mathbb{R}^n \to \mathbb{R}^m$ where $x \in \mathbb{R}^n \mapsto f(x) = y \in \mathbb{R}^m$.

We continue to work in the framework of metric spaces as defined before, with the Euclidean distance as metric. We will now consider sets in $\mathbb{R}^n$. For $x, y \in \mathbb{R}^n$, the Euclidean metric it is defined as

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

Thus, the notion of continuity applies to functions of several variables in exactly the same way as it applies to single variable functions, just by taking into account the specific metric that should be considered in each case.

4.1 Partial derivatives

- Consider a function $f : S \subset \mathbb{R}^n \to \mathbb{R}^m$. We can write it as $f(x) = [f_1(x) \quad f_2(x) \quad \dots \quad f_m(x)]'$ where $f_i : \mathbb{R}^n \to \mathbb{R}$ is a real-valued function.
- The *partial derivative* of f_i with respect to x_j is obtained by holding x_k fixed for all $k \neq j$ and differentiating f_i as if it was a single variable function. We write,

  $$\frac{\partial f_i(x)}{\partial x_j} = \lim_{\Delta \to 0} \frac{f_i(x + \Delta e_j) - f_i(x)}{\Delta}$$

  where e_j is the n-vector with a 1 in position j and zeros everywhere else (the jth unit vector), Δ is a small scalar increment, so that

  $$x + \Delta e_j = [x_1, \cdots, x_{j-1}, x_j + \Delta, x_{j+1}, \cdots, x_n]'.$$

- The partial derivative $\partial f_i/\partial x_j$ measures the slope of the curve in the x_j direction.
- Df, which is also sometimes denoted by $\nabla_x f(x)$, is the $n \times m$ matrix of all partial derivatives of f with respect to each argument. It is called the *Jacobian matrix* (or the *Gradient vector* if f is real-valued with $m = 1$).

  The generic term of $Df(x)$ is $a_{ij} = \partial f_i(x) \big/ \partial x_j$.
- f is $\mathcal{C}^1$ in S iff all the partials exist and are continuous in S.

4.2 Higher order derivatives

Let's consider a differentiable real-valued function $f : S \subset \mathbb{R}^n \to \mathbb{R}$, so that Df exists and is a function, $Df : S \subset \mathbb{R}^n \to \mathbb{R}^n$, i.e. it is an $n \times 1$ vector.

Suppose there exists an $x \in \mathbb{R}^n$ such that Df is differentiable at x. Then we say f is twice differentiable at x and the second derivative of f is $D^2 f$ and is given by:

$$D^2 f(x) = \nabla_{xx'} f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

- The matrix $D^2 f(x)$ is called the *Hessian* of f at x and typically represented by $H(x)$.
- The second partial derivative $\partial^2 f/\partial x_i^2$ measures the curvature of f in the x_i-direction.
- The second partial derivative $\partial^2 f/\partial x_i \partial x_j$ measures the rate at which the slope in direction j changes as we move in direction i.
- *Result:* if f is $\mathcal{C}^2$, then $D^2 f$ is a symmetric matrix, so that $\partial^2 f(x) \big/ \partial x_i \partial x_j = \partial^2 f(x) \big/ \partial x_j \partial x_i$.
- Higher order derivatives can be obtained applying differentiation sequentially.

4.3 Taylor series approximation of n-dimensional real-valued functions

- Suppose we know how to evaluate a function $f : \mathbb{R}^n \to \mathbb{R}$ at a point $a = (a_1, ..., a_n)$ and wish to approximate it at a nearby point $a + \Delta a$ for a small Δa.
- As for the one dimensional case, we can use the Taylor expansion,

$$f(a + \Delta a) \quad = \quad f(a) + \sum_{i=1}^{n} \frac{\partial f(a)}{\partial a_i} \Delta a_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial a_i \partial a_j} \frac{\Delta a_i \Delta a_j}{2!} + ...$$

- In matrix notation we can write,

$$f(a+\Delta a) \quad = \quad f(a) + \Delta a' \nabla f(a) + \frac{\Delta a' H(a) \Delta a}{2!} + ...$$

- The first order approximation of f around point a is,

$$f(a+\Delta a) \quad \simeq \quad f(a) + \sum_{i=1}^{n} \Delta a_i \frac{\partial f(a)}{\partial a_i}$$

- Writing $\Delta f(a) = f(a+\Delta a) - f(a)$ and rearranging we obtain,

$$\Delta f(a) \quad \simeq \quad \frac{\partial f(a)}{\partial a_1} \Delta a_1 + ... + \frac{\partial f(a)}{\partial a_n} \Delta a_n$$

  This expression measures the change in f when all its arguments change infinitesimally. This is the *total differential of* f and is usually written as:

$$df \approx \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} dx_i$$

4.4 The chain rule

- Suppose you have a function $f : \mathbb{R}^n \to \mathbb{R}$ where the arguments $(x_1, \ldots, x_n)$ are themselves functions of another variable, t.

- The *chain rule* states that:

$$\frac{df}{dt} \quad = \quad \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \ldots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt}$$

- An important application of the chain rule is when the arguments $(x_2, \ldots, x_n)$ can be written as functions of x_1 and we wish to know how f changes with, say, x_1 allowing for all the indirect effects through the impact of x_1 on the remaining arguments. The chain rule yields:

$$\frac{df}{dx_1} \quad = \quad \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \frac{dx_2}{dx_1} + \ldots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dx_1}$$

4.5 Turning points

- As in the one dimensional case, a turning point in the multidimensional case is identified by the first derivative being zero.

- In a n-dimensional real-valued function, we must ensure that all partial derivatives are zero,

$$\frac{\partial f}{\partial x_i} \quad = \quad 0 \quad \text{for all } i = 1, ...n$$

- This means that the function reached a *stationarity point* in all directions.
- As for the one dimensional case, however, this is not a sufficient condition, it is only a necessary condition.
- If a is a turning point then $Df(a) = 0$. Using a second order Taylor series approximation we now have,

$$\begin{aligned} \Delta f(a) &= f(a + \Delta a) - f(a) \\ &\simeq \frac{\Delta a' H(a) \Delta a}{2} \end{aligned}$$

  for a small change Δa.
- This is a quadratic form in Δa, giving us the means to identify the nature of the turning point.
- The sufficient conditions for local extrema will then be established by the nature of the Hessian matrix $H(a)$:
  - If $H(a)$ is positive/negative definite then a is a local unique minimum/maximum.
  - If $H(a)$ is positive/negative semi-definite then a is a local non-unique minimum/maximum: there is at least one direction on which the function is flat and other nearby minimums/maximums exist.
  - If $H(a)$ is indefinite then a is either an inflection or a saddle point.
- Aside: Negative (positive) [semi-] definiteness of the Hessian imply [strict] (convexity) concavity of f. Formally, f is (convex) concave if, for any $x_1, x_1 \in \mathbb{R}^n$and any $\alpha \in [0, 1]$, $\alpha f(x_1) + (1 - \alpha) f(x_2)(\geq) \leq f(\alpha x_1 + (1 - \alpha) x_2)$; strict (convexity) concavity is defined with the respective strict inequality. The function is said to be [strictly] (quasi-convex) quasi-concave if the "better-than" set $\{x \in \mathbb{R}^n : f(x) \geq \bar{f}\}$ is [strictly] (convex) concave.

4.6 The Implicit Function Theorem

Let's assume the following relationship cannot be explicitly solved for y as a function of x,

$$f(x, y) = 0 \qquad \text{where} \qquad (x, y) \in S \subset \mathbb{R}^{n+m} \text{ and } f : S \to \mathbb{R}^m,$$

and we would like to know how y changes with x. The *Implicit Function Theorem* tells us how to do it.

Note that one can interpret f as a system of m equations with m endogenous variables y and n exogenous variables x. One then might be interested in determining how changes in the exogenous variables x translate into changes of the endogenous variables y.

Notation: We will denote the $n \times m$ matrix of partial derivatives of f with respect to x by $D_x f$, and the $m \times m$ matrix of partial derivatives of f with respect to y by $D_y f$.

Theorem: Suppose $f : S \to \mathbb{R}^m$ is $\mathcal{C}^1$ in $S \subset \mathbb{R}^{n+m}$ and there exists vector (x^*, y^*) such that $f(x^*, y^*) = 0$ and $D_y f(x^*, y^*)$ is invertible, i.e. full rank when evaluated at (x^*, y^*). Then, there exists a neighborhood of x^*, $U_{x^*} \subset \mathbb{R}^n$, and a $\mathcal{C}^1$ function, $g : U_x \to \mathbb{R}^m$, such that

- $(x, g(x)) \in S$ for all $x \in U_{x^*}$.
- $g(x^*) = y^*$.
- $f(x, g(x)) = 0$ for all $x \in U_{x^*}$.
- And the derivative of g at any $x \in U_{x^*}$ is given by,

$$Dg(x)' = -(D_y f(x, g(x)))^{-1} D_x f(x, g(x))'.$$

To understand the last result, we can apply the total differential to the equation $f(x, y) = f(x, g(x)) = 0$ to obtain,

$$\frac{\partial f}{\partial y} g'(x) dx + \frac{\partial f}{\partial x} dx = 0$$

Assuming that $\frac{\partial f}{\partial y} \neq 0$, re-arranging the terms in this equation yields,

$$g'(x) = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

which is precisely the result in the implicit function theorem for the case of $m = n = 1$.

For the higher dimensional case, totally differentiating

$$f(x, g(x)) = 0$$

with respect to x yields

$$D_x f(x, g(x))' dx + D_y f(x, g(x)) Dg'(x) dx = 0,$$

which is an $m \times 1$ dimensional system. Assuming the $m \times m$ matrix $D_y f$ has full rank for all $x \in U_{x^*}$, this can be solved

$$Dg(x)' dx = -\left[D_y f(x, g(x))\right]^{-1} D_x f(x, g(x))' dx.$$

Since this has to hold for any small $dx \in \mathbb{R}^n$, it follows that

$$Dg(x)' = -\left[D_y f(x, g(x))\right]^{-1} D_x f(x, g(x))'.$$

Note that the Implicit Function Theorem provides a local, not a global result. For global results, see Gale-Nikkaido (1965) and Mas-Colell (1975).

Notice also the different requirements in the case $m = n = 1$ compared to the higher-dimensional case: For global invertibility in the former, it suffices that $f(x, \cdot)$ be strictly monotone; in the latter, there is no obvious analogue to monotonicity: f might increase if one argument is increased, even though another argument is decreased, due to differential curvature in the two dimensions. Hence, in higher dimensional problems, the matrix $D_y f$ has to satisfy stronger conditions to ensure that the theorem's result extends globally.

4.6.1 Example: Consumer Choice

Suppose a consumer makes consumption choices by maximizing a utility function $U(x)$, $x \in \mathbb{R}^J_+$, subject to a budget constraint $p'x \leq y$, given prices $p \in \mathbb{R}^J_{++}$ and income $y > 0$. Assume that $U(\cdot) \in \mathcal{C}^2$ is strictly monotonically increasing, i.e. $\nabla_x U(x) \in \mathbb{R}^J_{++}$ for any $x \in \mathbb{R}^J_+$.

Then, as a consequence of strict monotonicity, the consumer's UMP is

$$\max_{x \in \mathbb{R}^J_+ : p'x \leq y} U(x) = \max_{x \in \mathbb{R}^J_+ : p'x = y} U(x).$$

The first-order conditions of the UMP at the solution $x^\star \in \mathbb{R}^J_+$ and $\lambda^\star > 0$ are

$$\begin{aligned} p\lambda^\star &= \nabla_x U(x^\star) \\ y &= p'x^\star, \end{aligned}$$

where $\lambda^\star$ is the Lagrange multiplier on the budget constraint and can be interpreted as the shadow value of income (mathematical details in the section of Static Optimization).

The FOCs can be re-arranged as

$$0 = z(x^\star, \lambda^\star; y, p) = \begin{bmatrix} \nabla_x U(x^\star) - p\lambda^\star \\ p'x^\star - y \end{bmatrix}.$$

Hence, $0 = z(x^\star, \lambda^\star; y, p)$ implicitly characterizes the solution $(x^{\star\prime}, \lambda^\star)$ to the UMP.

To apply the Implicit Function Theorem in order to obtain a local characterization of the solution, it is required that the matrix

$$\begin{aligned} D &= \nabla_{(x',\lambda)'} z(x^\star, \lambda^\star; y, p) \\ &= \begin{bmatrix} \nabla_{xx'} U(x^\star) & -p \\ p' & 0 \end{bmatrix} \end{aligned}$$

be non-singular, i.e. that $|D| \neq 0$. Since $p\lambda^\star = \nabla U(x^\star)$, letting

$$\bar{D} = \begin{bmatrix} \nabla_{xx'} U(x^\star) & \nabla_x U(x^\star) \\ \nabla_x U(x^\star)' & 0 \end{bmatrix}$$

implies that $|-\lambda^\star||\lambda^\star||D| = |\bar{D}|$, and since $\lambda^\star > 0$, it follows that non-singularity of D is equivalent to non-singularity of $\bar{D}$. The latter requirement, $|\bar{D}| \neq 0$, is referred to as *smoothness of U in the sense of Gérard Debreu*. By virtue of the Implicit Function Theorem, it guarantees that, in a neighborhood to the solution $(x^{\star'}, \lambda^\star)$ of the UMP, there exists a system of demand functions $d(y, p)$ and a shadow value function of income $v(y, p)$ that is continuous and continuously differentiable in income and prices, satisfying $0 = z(d(y, p), v(y, p); y, p)$.

4.7 Integration

The theory of measures and Lebesgue integration extends to multivariate measurable functions in a straightforward manner.

4.7.1 Fubini's Theorem

There is one further result that is often useful: Fubini's Theorem (below).

- *Preliminaries*: Recall the Cartesian product of linear spaces[2], e.g. $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$. Let (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) be two measure spaces. Denote by $\Sigma_1 \times \Sigma_2$ the σ-algebra on the Cartesian product $X_1 \times X_2$ generated by subsets of the form $B_1 \times B_2$, where $B_1 \in \Sigma_1$ and $B_2 \in \Sigma_2$. The *product measure* $\mu_1 \times \mu_2$ is defined to be the unique measure on the measurable space $(X_1 \times X_2, \Sigma_1 \times \Sigma_2)$, with the property

  $$(\mu_1 \times \mu_2)(B_1 \times B_2) = \mu_1(B_1)\mu_2(B_2),$$

  for all $B_1 \in \Sigma_1$ and $B_2 \in \Sigma_2$. Lebesgue measure on $\mathbb{R}^{n+m}$ satisfies this property.

- *Fubini's Theorem*: Consider a μ-measurable function $f : S \to \mathbb{R}$, where $S \subset \mathbb{R}^{n+m}$ and $\mu = \mu_x \times \mu_y$. If $f \geq 0$ or $\int |f| d\mu < \infty$, then

  $$\int_{\mathbb{R}^n \times \mathbb{R}^m} f d\mu = \int_{x \in \mathbb{R}^n} \int_{y \in \mathbb{R}^m} f(x, y) d\mu_y d\mu_x = \int_{y \in \mathbb{R}^m} \int_{x \in \mathbb{R}^n} f(x, y) d\mu_x d\mu_y.$$

  The theorem, if applicable, permits to change the order of integration.

- Examples:

[2] The Cartesian product of two sets A and B (also called the product set, set direct product, or cross product) is defined to be the set of all points (a, b) where $a \in A$ and $b \in B$. It is denoted $A \times B$, and is called the Cartesian product. The main example is n copies of $\mathbb{R}$, $\mathbb{R} \times \cdots \times \mathbb{R} = \mathbb{R}^n$.

1. Consider the function $f(x,y) = x^2 + y^2$, where $x, y \in S = [-1,1]$ and μ is Lebesgue measure. Then,

$$\begin{aligned}\int_{-1}^{1}\int_{-1}^{1} f(x,y)dxdy &= \int_{-1}^{1}\left[\frac{1}{3}x^3 + xy^2\right]_{-1}^{1} dy \\ &= \int_{-1}^{1}\left[\frac{2}{3} + 2y^2\right] dy \\ &= \left[\frac{2}{3}y + \frac{2}{3}y^3\right]_{-1}^{1} \\ &= \frac{8}{3},\end{aligned}$$

   and, by symmetry, the reverse order of integration produces the same result.

2. The function $f(x,y) = \exp(-xy) - 2\exp(-2xy)$, for $S = [0,1] \times [1,\infty)$, with Lebesgue measure μ, is not integrable. To see this, let $x = 0$ so that $f(0,y) = -1$, which is not integrable on $[1,\infty)$. Hence, Fubini's Theorem is expected to fail. In fact,

$$\begin{aligned}\int_0^1\int_1^\infty f(x,y)dydx &= \int_0^1 \frac{1}{x}\left(\exp(-x) - \exp(-2x)\right)dx > 0 \\ \int_1^\infty\int_0^1 f(x,y)dxdy &= \int_1^\infty \frac{1}{y}\left(\exp(-2y) - \exp(-y)\right)dy < 0.\end{aligned}$$

4.7.2 Differentiation under the integral sign (Leibniz rule)

- Suppose that $f(x,y)$ as above satisfies that, for each $x \in X$, $f(x,\cdot) \in \mathcal{L}^1(Y,\mu_y)$,

  i.e. $\int_{y\in Y} |f(x,y)|d\mu_y < \infty$ for each x. Let $F(x) = \int_{a(x)}^{b(x)} f(x,y)d\mu_y$.

  If there exists a function $g \in \mathcal{L}^1(Y,\mu_y)$ such that $|f(x,y)| \leq g(y)$ for all x and y, then, if $f(x,y)$ is continuous in x for each y, then $F(x)$ is continuous.

  If $\frac{\partial}{\partial x}f(x,y)$, $a'(x)$ and $b'(x)$ exist and there exists a function $g \in \mathcal{L}^1(Y,\mu_y)$ such that $\left|\frac{\partial}{\partial x}f(x,y)\right| \leq g(y)$ for all x and y, then $F(x)$ is differentiable in x and

$$F'(x) = f(x,b(x))b'(x) - f(x,a(x))a'(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x}f(x,y)d\mu_y.$$

- Remark: The Leibniz rule is a generalization of the 1st Fundamental Theorem of Calculus.

- To see why the dominance condition on the partial derivative function is required, note that, for $x_n \to x$ as $n \to \infty$,

$$\frac{\partial}{\partial x}f(x,y) = \lim_n \frac{f(x_n,y) - f(x,y)}{x_n - x}.$$

  By the mean value theorem and the dominance condition on the partial derivative function,

$$\left|\frac{f(x_n,y) - f(x,y)}{x_n - x}\right| \leq \sup_{x\in X}\left|\frac{\partial}{\partial x}f(x,y)\right| \leq g(y).$$

Then, by the Lebesgue Dominated Convergence Theorem, for limits of integration that do not depend on x,

$$F'(x) = \lim_n \frac{F(x_n) - F(x)}{x_n - x} = \int_Y \frac{\partial}{\partial x} f(x, y) d\mu_y.$$

- Example: Suppose a consumer lives two periods and has temporal utility $u(c_t)$, $t = 1, 2$, where $u(\cdot)$ is strictly concave (i.e. $u'(c) > 0$ and $u''(c) < 0$ for all $c > 0$) and satisfies $\lim_{c\to 0} u'(c) = +\infty$. The consumer has a given period 1 endowment $y_1 > 0$ and faces an uncertain period 2 endowment $Y_2 \geq 0$, whose distribution is denoted by μ and has support $Y \subseteq \mathbb{R}_+$, i.e. $\mu(Y) = 1$ and $\mu([0, \epsilon)) > 0$ for any $\epsilon > 0$. The consumer's choice problem is to decide how much of y_1 to consume in period 1 when the remainder can be transferred to period 2 and then be consumed with Y_2. This problem is equivalent to the problem of how much to save from the period 1 endowment:

$$\begin{aligned} V(y_1) &= \max_{0\leq c_1\leq y_1} \{u(c_1) + \beta \mathbb{E}[u(Y_2 + y_1 - c_1)]\} \\ &= \max_{s:0\leq s\leq y_1} \{u(y_1 - s) + \beta \mathbb{E}[u(Y_2 + s)]\} \\ &= \max_{s:0\leq s\leq y_1} \left\{u(y_1 - s) + \beta \int_Y u(y + s) d\mu(y)\right\}, \end{aligned}$$

where $\beta \in (0, 1)$ discounts period 2 expected utility.

Since $\lim_{c\to 0} u'(c) = +\infty$ and the probability of arbitrarily small Y_2 is positive, the solution $s^\star$ to this maximization problem will be an interior one, i.e. $0 < s^\star < y_1$. Such an interior solution can be characterized using differential calculus, i.e. by the FOC

$$u'(y_1 - s^\star) = \beta \frac{d}{ds} [\mathbb{E}[u(Y_2 + s)]]_{s=s^\star} = \beta \frac{d}{ds} \left[\int_Y u(y + s) d\mu(y)\right]_{s=s^\star}.$$

Since $s^\star$ is an interior solution, it must be the case that $0 < u'(y + s^\star) < \infty$ with probability one, i.e. for all realizations y of the random period 2 endowment Y_2. Hence, the integrand on the right-hand side is bounded. Therefore,

$$u'(y_1 - s^\star) = \beta \mathbb{E}\left[u'(Y_2 + s^\star)\right] = \beta \int_Y u'(y + s^\star) d\mu(y).$$

This is the so-called Euler equation for stochastic intertemporal consumer choice problems.

4.8 Homogeneity of Functions

4.8.1 Definition

Consider a function $f : S \to \mathbb{R}^m$, $S \subseteq \mathbb{R}^n$. The function f is homogeneous of degree k if $f(\alpha x) = \alpha^k f(x)$ for all $\alpha > 0$ and $x, \alpha x \in S$.

4.8.2 Examples

4.8.3 Linear Functions

Linear functions are homogeneous of degree 1, since by definition

$$f(\alpha x) = \alpha f(x), \quad \text{for any } \alpha \in \mathbb{R} \text{ and any } x \in \mathbb{R}^n.$$

4.8.4 Homogeneous Polynomials

- Monomials in $S \subseteq \mathbb{R}^n$: These are defined as

  $$f(x;\beta) = \prod_{i=1}^{n} x_i^{\beta_i},$$

  where $b = \sum_{i=1}^{n} \beta_i$ is the degree of the monomial.

  Since, for $\alpha > 0$,

  $$f(\alpha x;\beta) = \prod_{i=1}^{n} (\alpha x_i)^{\beta_i} = \alpha^{\sum_{i=1}^{n} \beta_i} \prod_{i=1}^{n} x_i^{\beta_i} = \alpha^b f(x),$$

  monomials of degree b are homogeneous of degree b.

- Homogeneous Polynomials: These are defined as the sum of monomials of the same degree,

  $$g(x;\theta) = \sum_{j=1}^{m} f(x;\theta_j) = \sum_{j=1}^{n} \prod_{i=1}^{n} x_i^{\theta_{ij}},$$

  where $\theta' = (\theta_1', \cdots, \theta_m') = (\theta_{11}, \cdots, \theta_{mn}) \in \mathbb{R}^{nm}$ and $\sum_{i=1}^{n} \theta_{ij} = c$ for all $j = 1, \cdots, m$ and some c.

  It follows from the constituent monomials being homogeneous of degree c that $g(\alpha x;\theta) = \alpha^c g(x;\theta)$, i.e. the homogeneous polynomial inherits the constituent monomials' degree of homogeneity.

- Other functions

  E.g. the determinant operator, applied to $n \times n$ matrices,

  $$|\alpha A| = |\alpha|^n |A|,$$

  is homogeneous of degree n.

4.8.5 Properties

4.8.6 Euler's Theorem

Consider $f : \mathbb{R}^n \to \mathbb{R}$, $f \in \mathcal{C}^\infty$. Then, f is homogeneous of degree k if, and only if,

$$x'\nabla f(x) = kf(x) \quad \forall x \in \mathbb{R}^n. \tag{6}$$

Proof:

- Homogeneity of degree k implies (6): Homogeneity of f of degree k implies

  $$f(\alpha x) \equiv \alpha^k f(x), \quad \forall \alpha > 0, x \in \mathbb{R}^n.$$

  Differentiating this identity on both sides with respect to α yields

  $$\nabla f(\alpha x)' x = x'\nabla f(\alpha x) = k\alpha^{k-1} f(x), \quad \forall \alpha > 0, x \in \mathbb{R}^n.$$

  For $\alpha = 1$, (6) follows.

- (6) implies homogeneity of degree k: Fix $x \in \mathbb{R}^n$, and define the scalar function

  $$h(\alpha) = \alpha^{-k} f(\alpha x) - f(x),$$

  which maps $h : \mathbb{R}_+ \to \mathbb{R}$, for x given. Since $f \in \mathcal{C}^\infty$, h is differentiable, with derivative

  $$\begin{aligned} h'(\alpha) &= -k\alpha^{-k-1} f(\alpha x) + \alpha^{-k}\nabla f(\alpha x)' x \\ &= \alpha^{-k-1}\left[-kf(\alpha x) + \nabla f(\alpha x)'\alpha x\right] \\ &= \alpha^{-k-1}\left[-kf(\tilde{x}) + \nabla f(\tilde{x})'\tilde{x}\right], \end{aligned}$$

  where $\tilde{x} = \alpha x \in \mathbb{R}^n$. By (6), the expression in square brackets is zero. Hence, $h'(\alpha) = 0$ for all $\alpha > 0$, so h is seen to be constant. Since $h(1) = 0$, it follows that $h(\alpha) = 0$ for all $\alpha > 0$, i.e.

  $$\alpha^k f(x) = f(\alpha x),$$

  i.e. f is homogeneous of degree k. □

4.8.7 Homogeneity of Derivatives

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is homogeneous of degree k. Then, the partial derivatives of f w.r.t. x_i, $\frac{\partial}{\partial x_i} f(x)$, $i = 1, \cdots, n$, are homogeneous of degree $k - 1$.

This follows from

$$\begin{aligned} f(\alpha x) &\equiv \alpha^k f(x) \quad \forall \alpha > 0, x \in \mathbb{R}^n \\ \Rightarrow \quad \frac{\partial}{\partial x_i} f(\alpha x)\alpha &= \alpha^k \frac{\partial}{\partial x_i} f(x) \\ \Rightarrow \quad \frac{\partial}{\partial x_i} f(\alpha x) &= \alpha^{k-1} \frac{\partial}{\partial x_i} f(x) \quad \forall \alpha > 0, x \in \mathbb{R}^n, i = 1, \cdots, n. \end{aligned}$$

4.8.8 Applications

4.8.9 Consumer Choice: UMP

Suppose a consumer makes consumption choices by maximizing a utility function $U(x)$, $x \in \mathbb{R}^J_+$, subject to a budget constraint $p'x \leq y$, given prices $p \in \mathbb{R}^J_{++}$ and income $y > 0$. Assume that $U(\cdot) \in \mathcal{C}^2$ is strictly monotonically increasing, i.e. $\nabla_x U(x) \in \mathbb{R}^J_{++}$ for any $x \in \mathbb{R}^J_+$.

Then, as above, the consumer's UMP satisfies

$$\max_{x \in \mathbb{R}^J_+ : p'x \leq y} U(x) = \max_{x \in \mathbb{R}^J_+ : p'x = y} U(x) = \max_{x \in \mathbb{R}^J_+ : \alpha p'x = \alpha y} U(x),$$

for any $\alpha > 0$. Therefore,

$$\begin{aligned} d(p, y) &= \arg \max_{x \in \mathbb{R}^J_+ : p'x \leq y} U(x) \\ &= \arg \max_{x \in \mathbb{R}^J_+ : p'x = y} U(x) \\ &= \arg \max_{x \in \mathbb{R}^J_+ : \alpha p'x = \alpha y} U(x) \\ &= d(\alpha p, \alpha y), \quad \alpha > 0, \end{aligned}$$

so that the Marshallian demand functions are seen to be homogeneous of degree zero in p and y.

4.8.10 Consumer Choice: EMP

The dual to the consumer's UMP is the consumer's expenditure minimization problem (EMP), i.e. for $\alpha > 0$,

$$e(p, u) = \min_{x \in \mathbb{R}^J_+ : U(x) \geq u} p'x = \frac{1}{\alpha} \left\{ \min_{x \in \mathbb{R}^J_+ : U(x) \geq u} \alpha p'x \right\} = \frac{1}{\alpha} e(\alpha p, u).$$

Hence, the consumer's expenditure function $e(p, u)$ is homogeneous of degree 1 in p.

4.8.11 Production

Consider a firm whose technology is characterized by the production function $f : \mathbb{R}^n_+ \to \mathbb{R}_+$, which transforms n inputs $x \in \mathbb{R}^n_+$ into output $y = f(x) > 0$. Assume f satisfies $\nabla f(x) \in \mathbb{R}^J_{++}$ for all $x \in \mathbb{R}^J$.

Assume that the final output and the factor markets are perfectly competitive, so that the firm acts as a price-taker on these markets (i.e. output and factor prices are exogenous). The firm will choose inputs such as to maximize its profit, given the unit price $p > 0$ for the final output y and factor prices $w \in \mathbb{R}^J_{++}$,

$$\begin{aligned} x^\star &= \arg\max_{x \in \mathbb{R}^J_+} \left\{ pf(x) - w'x \right\} \\ \text{FOCs: } p\nabla f(x^\star) &= w. \end{aligned}$$

Suppose, furthermore, that $f(x)$ is homogeneous of degree 1 in x; technologies characterized by production functions which are homogeneous of degree 1 are said to exhibit constant returns to scale (CRTS). Then, the FOCs imply

$$p\nabla x^{\star\prime} f(x^\star) = x^{\star\prime} w,$$

which, together with Euler's Theorem (6), implies in turn that

$$pf(x^\star) = x^{\star\prime} w,$$

i.e. a firm operating under a CRTS technology makes zero profits, given any positive output and factor prices.

5 Static optimisation in $\mathbb{R}^n$

References: Simon and Blume, chapter 17 - 19; Chiang, chapters 9, 11 and 12.

- An optimization problem is one where a given function $f : \mathbb{R}^n \to \mathbb{R}$ is to be maximized or minimized over a given set $S \subset \mathbb{R}^n$.
- The function f is called the *objective function*.
- Optimisation problems are typically represented as

$$\begin{aligned} \max_x f(x) \qquad & \text{subject to} \quad x \in \mathcal{D} \\ \min_x f(x) \qquad & \text{subject to} \quad x \in \mathcal{D}, \end{aligned}$$

where the set $\mathcal{D}$ is the *constraint set*.

- The set of solutions to the optimization problem is given by,

$$\begin{aligned} \arg\max\{f(x) : x \in \mathcal{D}\} &= \{x \in \mathcal{D} : f(x) \geqslant f(y) \text{ for all } y \in \mathcal{D}\} \\ \arg\min\{f(x) : x \in \mathcal{D}\} &= \{x \in \mathcal{D} : f(x) \leqslant f(y) \text{ for all } y \in \mathcal{D}\} \end{aligned}$$

- Recall *Weierstrass Theorem:* Let $S \subset \mathbb{R}^n$ be a compact set and let $f : S \to \mathbb{R}$ be a continuous function on S. Then f attains a maximum and minimum on S. These maxima and minima may not be unique, and they may be on the boundary of S, denoted by ∂S, i.e. they need not be interior solutions (the interior of S is denoted by $\overset{\circ}{S}$).

5.1 Unconstrained Optimisation

Consider a real-valued function in $\mathbb{R}^n$, $f : S \subset \mathbb{R}^n \to \mathbb{R}$.

5.1.1 First order conditions for unconstrained optimisation

- A *critical point* is a value of $x \in S$ where $Df(x) = 0$.
- *First order conditions* (FOCs) for a maximum or minimum: Let f be differentiable at $x^* \in \overset{\circ}{S}$. If x^* is a local maximum or minimum of f on S then x^* is a critical point of f ($Df(x^*) = 0$).
- The FOCs are *necessary* conditions for an *interior* local optimum, but they *are not sufficient*.
- For local optima x^* at the *boundary* of S, it need not be the case that the (one-sided) derivative of f at x^* be zero.
- Multivariate functions may have critical points that are not extrema: these are the either *saddle points* or *inflection points*.
- Univariate functions may have critical points that are not extrema: these are *inflection points*.

5.1.2 Second order conditions for unconstrained optimisation

How can we assess if a given critical point is an interior local maximum, minimum or a saddle point?

- A local interior maximum (minimum), x^* is called a strict local maximum (minimum) if there exists an $\epsilon > 0$ such that $f(x^*) > (<) f(y)$ for all y in the small neighbourhood of radius ϵ around x^* (that is, $y \in O_\epsilon(x^*) \cap S$).
- *Result*: Suppose that f is $\mathcal{C}^2$ in S and that $x^* \in \overset{\circ}{S}$ is a critical point of f. Then,
  - If x^* is an interior local maximum (minimum), then $D^2 f(x^*)$ is negative (positive) semidefinite.
  - If $D^2 f(x^*)$ is positive (negative) definite, then x^* is a strict (locally unique) interior local minimum (maximum).
  - If $D^2 f(x^*)$ is positive (negative) semi-definite, then x^* is a (possibly non-unique) interior local minimum (maximum).
  - If $D^2 f(x^*)$ is indefinite, then x^* is neither a local maximum or or a local minimum.
- Notice that if the Hessian matrix is not positive or negative definite, then the critical point may be a maximum, a minimum or neither. If the Hessian is neither positive nor negative semi-definite, then the critical point is a saddle point for sure.

5.1.3 Convexity and concavity

- We have seen that a univariate function f defined on an open set S is convex (concave) when its second derivative is non-negative (non-positive). In this case, a point x^* is a interior global minimum (maximum) iff it is a critical point. The function f is strictly convex (concave) when its second derivative is strictly positive (negative). In this case, only one critical point may exist at the most and the function has a unique interior minimum (maximum) at x^* iff x^* is the unique critical point; when x^* is the unique extremum and $x^* \in \partial S$, then it need not be a critical point.
- Similar results apply to multivariate functions. Consider a real-valued function $f : \mathbb{R}^n \to \mathbb{R}$. Then:
  - f is convex (concave) iff the Hessian matrix $D^2 f(x)$ is positive (negative) semi-definite on its entire domain.
  - If f is convex (concave), then a point x^* is an interior global minimum (maximum) iff x^* is a critical point.
  - f is strictly convex (concave) iff the Hessian matrix $D^2 f(x)$ is positive (negative) definite on its entire domain.
  - If f is strictly convex (concave), then a point x^* is the unique interior global minimum (maximum) iff x^* is the (unique) critical point.

5.2 Optimization With Equality Constraints

We will now study optimization problems of the form,

$$\begin{aligned} \max_{x \in S \subset \mathbb{R}^n} \quad & f(x) && (7) \\ \text{s.t.} \quad & g_i(x) = c_i \text{ for } i = 1, ..., k && (8) \end{aligned}$$

$$\begin{aligned} \min_{x \in S \subset \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) = c_i \text{ for } i = 1, ..., k, \end{aligned}$$

which is the same as considering the constraint set $\mathcal{D}$ to be

$$\mathcal{D} = S \cap \{x \in \mathbb{R}^n : g_i(x) = c_i \text{ for } i = 1, ..., k\}$$

Without any loss of generality, we can consider restrictions of the form, $g_i(x) = 0$, $i = 1, \cdots, k$.

5.2.1 Example: the utility maximization problem

A consumer decides how much to consume of two different consumption goods, x_1 and x_2, while facing a budget constraint, determined by the consumer's income I and the prices for the two consumption goods p_1 and p_2. We write (abstracting for now from the non-negativity constraints and the possibility of consuming less than the total income),

$$\begin{aligned} \max_{x_1, x_2} \quad & f(x_1, x_2) \\ \text{s.t.} \quad & g(x_1, x_2) = p_1 x_1 + p_2 x_2 = I \end{aligned}$$

- To solve this problem we first start by noticing that the constraint $g(x_1, x_2) - I = 0$ allows as to determine x_2 as a function of x_1 (for fixed I): $x_2 = y(x_1)$.
- But then we can replace it in the objective function to rewrite the optimization problem as:

$$\max_{x_1} \quad h(x_1) = \max_{x_1} f(x_1, y(x_1))$$

- The first order condition for this maximization problem is (chain rule):

$$\begin{aligned} h'(x_1) &= \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \frac{dy}{dx_1} \\ &= 0 \end{aligned}$$

- But since y is the solution to the budget constraint equation: $g(x_1, y(x_1)) = 0$ we can use the implicit function theorem to compute its derivative:

$$\frac{dy}{dx_1} = -\frac{\frac{\partial g}{\partial x_1}}{\frac{\partial g}{\partial x_2}},$$

provided $\frac{\partial g}{\partial x_2} \neq 0$.

- Replacing in the FOCs yields:

$$\frac{\partial f}{\partial x_1} - \frac{\partial g}{\partial x_1}\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial g}{\partial x_2}} = 0$$

- Let $\lambda = \frac{\partial f}{\partial x_2}\left(\frac{\partial g}{\partial x_2}\right)^{-1}$. Then the above equation can be written as:

$$\frac{\partial f}{\partial x_1} - \lambda\frac{\partial g}{\partial x_1} = 0 \tag{9}$$

and rearranging the equation defining λ yields:

$$\frac{\partial f}{\partial x_2} - \lambda\frac{\partial g}{\partial x_2} = 0 \tag{10}$$

- (9) and (10) are the first order conditions for the utility maximization problem above.
- However, notice that we have 2 equations and 3 unknowns. So we need an extra condition to have a unique solution: this is the budget constraint:

$$g(x_1, x_2) = p_1x_1 + p_2x_2 = I$$

- There are other forms of arriving at the same FOCs; e.g. see below.

5.2.2 The general n-dimensional case

- In the general n-dimensional case we notice that, at the optimum $x^\star \in \mathbb{R}^n$, infinitesimal changes in the variables $x_1, \ldots, x_n$ which satisfy the constraint must have no impact on the value of the objective function.
- That is, if $x^\star$ is a (interior) solution to the optimization problem (7)-(8) then

$$df = \frac{\partial f(x^\star)}{\partial x_1}dx_1 + \ldots + \frac{\partial f(x^\star)}{\partial x_n}dx_n = 0$$

for all $dx_1, \ldots, dx_n$ satisfying

$$dg = \frac{\partial g(x^\star)}{\partial x_1}dx_1 + \ldots + \frac{\partial g(x^\star)}{\partial x_n}dx_n = 0$$

- Solving with respect to dx_1 in the last equation yields (where it is being assumed that $\partial g/\partial x_1 \neq 0$),

$$dx_1 = \frac{-\sum_{i=2}^{n}\frac{\partial g(x^\star)}{\partial x_i}dx_i}{\frac{\partial g(x^\star)}{\partial x_1}},$$

- which, when replaced in the total differential of f and after re-arranging, yields,

$$\sum_{i=2}^{n}\left[\frac{\partial f(x^\star)}{\partial x_i}-\frac{\frac{\partial f(x^\star)}{\partial x_1}\frac{\partial g(x^\star)}{\partial x_i}}{\frac{\partial g(x^\star)}{\partial x_1}}\right]dx_i \;=\; 0$$

  for *any* infinitesimal changes $dx_2,\ldots,dx_n$

- Since this has to hold for any $dx_1,\ldots,dx_n$, the solution is to set all terms in square brackets to zero, which therefore yields,

$$\frac{\partial f/\partial x_i}{\partial g/\partial x_i} \;=\; \frac{\partial f/\partial x_1}{\partial g/\partial x_1}=\lambda \;\; \forall i=2,\cdots,n,$$

  provided that $\partial g/\partial x_i \neq 0$ for $i=2,\cdots,n$.

- These are analogous to (9) and (10) above.

5.2.3 The Theorem of Lagrange

- A convenient way to derive the system of FOCs above from an *unconstrained* optimization problem is to write the *Lagrangian equation*,

$$\mathcal{L}(x;\lambda)=f(x)-\lambda'\left[g(x)-c\right]$$

  where λ is a vector of k *Lagrange multipliers* and

$$f:\mathbb{R}^n\to\mathbb{R} \quad , \quad g:\mathbb{R}^n\to\mathbb{R}^k \quad \text{and} \quad \lambda,c\in\mathbb{R}^k, k<n.$$

- *Theorem*: Consider the following assumptions,

  1. S is an open set.
  2. The functions $f:S\subset\mathbb{R}^n\to\mathbb{R}$ and $g:\mathbb{R}^n\to\mathbb{R}^k$ are $\mathcal{C}^1$.
  3. x^* is an interior local maximum or minimum of f on $\mathcal{D}=S\cap\{x\in\mathbb{R}^n : g_i(x)=0 \text{ for } i=1,...,k\}$.
  4. $\text{rank}\left(Dg\left(x^*\right)\right)=k$.

  Then there exists a vector $\lambda^*=\left(\lambda_1^*,...,\lambda_k^*\right)\in\mathbb{R}^k$ such that,

$$Df(x^*)-\sum_{i=1}^{k}\lambda_i^* Dg_i(x^*) \;=\; 0 \qquad (11)$$

- Notice that this theorem provides necessary conditions only: if there is an interior solution, then it is captured by condition (11). But points satisfying this condition may not be extrema.

- The 4th condition is called the *constraint qualification*. It ensures Dg contains a $k\times k$ submatrix that is invertible. It plays the role of the requirement detailed within the utility maximization problem that the partial derivatives of g are non-zero.

- To appreciate the role of the constraint qualification, notice it implies that, when $k < n$, w.l.o.g. the $k \times n$ matrix $Dg(x^\star)$ can be partitioned as

$$\underset{k\times n}{Dg(x^\star)} = \left[\underset{k\times k}{D_1(x^\star)} : \underset{k\times(n-k)}{D_2(x^\star)}\right],$$

and $\text{rk}\,(D_1(x^\star)) = k$. Partition x accordingly as $\underset{1\times n}{x'} = [\underset{1\times k}{x_1'}, \underset{1\times(n-k)}{x_{-1}'}]$. Then, the Implicit Function Theorem implies that, for any x in an ϵ-neighborhood of $x^\star$, there exists a function $\gamma : \mathbb{R}^{n-k} \to \mathbb{R}^k$ such that $x_1 = \gamma(x_{-1})$; moreover, $g(\gamma(x_{-1}), x_{-1}) - c = 0$ for all x in this neighborhood implies $\nabla\gamma(x_{-1}) = -D_1(\gamma(x_{-1}), x_{-1})^{-1} D_2(\gamma(x_{-1}), x_{-1})$. Then,

$$\begin{aligned} &\max_{\{x:g(x)-c=0\}} f(x) \\ \Leftrightarrow\ &\max_{x_{-1}} f(\gamma(x_{-1}), x_{-1}). \end{aligned}$$

The FOCs of the latter optimization problem at the solution $x^\star$ are as usual,

$$\begin{aligned} 0 &= \nabla_{x_{-1}}\gamma(x_{-1}^\star)'\nabla_{x_1} f(x^\star) + \nabla_{x_{-1}} f(x^\star) \\ &= -D_2(x^\star)' D_1(x^\star)^{-1'}\nabla_{x_1} f(x^\star) + \nabla_{x_{-1}} f(x^\star), \end{aligned}$$

which, in turn, implies

$$\nabla_{x_{-1}} f(x^\star) = D_2(x^\star)' D_1(x^\star)^{-1'}\nabla_{x_1} f(x^\star). \tag{12}$$

Compare this to (11): The system of n equations (11) can be written as

$$0 = \begin{bmatrix} \nabla_{x_1} f(x^\star) \\ \nabla_{x_{-1}} f(x^\star) \end{bmatrix} - Dg(x^\star)'\lambda = \begin{bmatrix} \nabla_{x_1} f(x^\star) \\ \nabla_{x_{-1}} f(x^\star) \end{bmatrix} - \left[[D_1(x^\star) : D_2(x^\star)]'\lambda\right].$$

This implies

$$\begin{aligned} \nabla_{x_1} f(x^\star) &= D_1(f(x^\star))'\lambda \quad \Rightarrow \lambda = D_1(x^\star)^{-1'}\nabla_{x_1} f(x^\star) \\ \nabla_{x_{-1}} f(x^\star) &= D_2(x^\star)'\lambda, \end{aligned}$$

which can be combined to yield $\nabla_{x_{-1}} f(x^\star) = D_2(x^\star)' D_1(x^\star)^{-1'}\nabla_{x_1} f(x^\star)$, which is the same as (12).

5.2.4 The Lagrange multipliers

- The Lagrange multipliers measure the sensitivity of the optimal value of the objective function f to small (infinitesimal) relaxations in the respective constraint.
- To see this, consider $g(x) - c = 0$ as a function of c, $c \in \mathbb{R}^k$.

- Denote the optimal policy as a function of c by $(x^*(c), \lambda^*(c))$. Write $v(c) = f(x^*(c))$, v being the optimal value of f or the *maximum value function*.

- Under the conditions of the theorem and assuming $x^*(c)$ is $\mathcal{C}^1$ we can apply the composite function differentiation rule (chain rule) to $v(c) = f(x^*(c))$,

$$\underset{k\times 1}{Dv(c)} = \underset{k\times n}{Dx^*(c)'} \; \underset{n\times 1}{Df\,(x^*\,(c))}$$

  and to $c = g(x^*(c))$,

$$\begin{aligned} \underset{k\times 1}{e_i} &= \underset{k\times n}{Dg\,(x^*\,(c))} \underset{n\times 1}{\frac{\partial}{\partial c_i} x^*(c)} \quad i = 1, \cdots, k \\ \Rightarrow I_k &= Dg(x^*(c))Dx^*(c), \end{aligned}$$

  where $e_i = \partial c/\partial c_i$, a k-dimensional vector of 0's with a 1 in the ith position.

- Pre-multiplying the Lagrange optimality condition in (11) by $\underset{k\times n}{Dx^*(c)'}$ and using the above conditions yields,

$$\begin{aligned} & \underset{k\times n}{Dx^*(c)'} \underset{n\times 1}{Df(x^*(c))} - \underset{k\times n}{Dx^*(c)'} \underset{n\times k}{Dg(x^*(c))'} \underset{k\times 1}{\lambda^*} = \underset{k\times 1}{0} \\ \Leftrightarrow \quad & Dv(c) - I_k\lambda^* = 0 \\ \Leftrightarrow \quad & Dv(c) = \lambda^* \end{aligned}$$

- We can interpret λ_j now: it measures the value of relaxing restriction j as measured by the objective function.

5.2.5 Procedure to identify the optimum

- Given an optimisation problem

$$\begin{aligned} \max_x \;\; & f(x) \\ \text{s.t. } & x \in \mathcal{D} = S \cap \{x \in \mathbb{R}^n : g(x) = c\} \end{aligned}$$

  write the Lagrangian equation,

$$\mathcal{L}(x; \lambda) = f(x) - \sum_{i=1}^{k} \lambda_i (g_i(x) - c_i)$$

- Write down the FOCs at a candidate solution $(x^{*'}, \lambda^{*'})$,

$$\begin{aligned} \frac{\partial \mathcal{L}(x^*; \lambda^*)}{\partial x_i} &= 0 \qquad \text{for} \qquad i = 1, ..., n \\ \frac{\partial \mathcal{L}(x^*; \lambda^*)}{\partial \lambda_i} &= 0 \qquad \text{for} \qquad i = 1, ..., k \end{aligned}$$

- Solve the system of equations for x^* and λ^*.
- Compute the value of f at all values of x^* that are solutions to the system of equations.
- Note: $\max_{x,\lambda} L(x, \lambda) = f(x^*)$, because $g(x^*) - c = 0$.

5.2.6 When does this work?

Suppose the following conditions hold,

- S is open.
- An interior global optimum, x^*, exists to the optimization problem.
- The constraint qualification is met at x^* and $k < n$.

Then there exists λ^* such that (x^*, λ^*) is a critical point of $\mathcal{L}$.

5.2.7 Digression: The Envelope theorem

We have encountered this result before, when exploring the interpretation of the Lagrange multipliers.

- Suppose we have a generic maximization problem. Write it the following way,

$$\max_{x \in \mathbb{R}^n} \quad f(x; q)$$

  where q are some parameters considered fixed.

- Let $x^*(q)$ be a solution to this problem given some value for the parameters q.
- Define $v(q) = f\left(x^*(q); q\right)$ to be the value function.
- In this simple case without constraints we can write, using the composite function derivative rule (chain rule),

$$\frac{dv(q)}{dq} = \frac{\partial f\left(x^*(q); q\right)}{\partial q} + \frac{\partial f\left(x^*(q); q\right)}{\partial x}\frac{dx^*(q)}{\partial q} = \frac{\partial f\left(x^*(q); q\right)}{\partial q} \tag{13}$$

  since $\partial f / \partial x = 0$ at the optimum x^*.

- This is generalizable to the constraint maximization setting.

The Envelope Theorem

Let $x^*(q)$ of dimension n denote the solution to the maximization problem

$$\begin{aligned} \max_{x \in \mathbb{R}^n} \quad & f(x; q) \\ \text{s.t} \quad & g_1(x; q) = 0 \\ & \vdots \\ & g_k(x; q) = 0 \end{aligned}$$

for any fixed choice of the parameters $q = (q_1, ...q_m)$. Let $v(q) = f\left(x^*(q); q\right)$.

Now consider the following assumptions,

- $f, g_1, ...g_k$ are $\mathcal{C}^1$ in $\mathbb{R}^n$ into $\mathbb{R}$.
- $x^*(q)$ and the set of Lagrange multipliers, $\lambda_1^*(q), ..., \lambda_k^*(q)$ are $\mathcal{C}^1$ functions
- The constraint qualification assumption holds.

Then,

$$\frac{dv(q)}{dq} = \frac{df\left(x^*(q); q\right)}{dq} - \sum_{i=1}^{k} \lambda_i^*(q) \frac{\partial g_i(x^*(q); q)}{\partial q}$$

Note that this is the same a (13), with f replaced by $\mathcal{L} = f - \sum_i \lambda_i g_i$.

5.3 Optimization With Inequality Constraints

Let's now consider the following type of optimization problems,

$$\begin{aligned} \max_{x \in S \subset \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & h_i(x) \leqslant c_i \text{ for } i = 1, ..., l \end{aligned}$$

$$\begin{aligned} \min_{x \in S \subset \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & h_i(x) \leqslant c_i \text{ for } i = 1, ..., l \end{aligned}$$

which is the same as considering the constraint set $\mathcal{D}$ to be,

$$\mathcal{D} = S \cap \{x \in \mathbb{R}^n : h_i(x) \leqslant c_i \text{ for } i = 1, ..., l\}$$

Without any loss of generality, we will consider restrictions with the form, $h_i(x) \leqslant 0$.

Before proceeding, consider the following definition: a constraint $h_i(x) \leqslant 0$ is *effective or binding* at x^* if it holds with equality at x^*.

5.3.1 The Theorem of Kuhn and Tucker

Consider the following assumptions,

1. S is an open set.
2. The functions $f : S \subset \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to \mathbb{R}^l$ are $\mathcal{C}^1$.
3. x^* is an interior local maximum or minimum of f on $\mathcal{D} = S \cap \{x \in \mathbb{R}^n : h_i(x) \leqslant 0 \text{ for } i = 1, ..., l\}$.
4. Let $E \subset 1, ..., l$ denote the set of effective constraints at x^* and $h_E = (h_i)_{i \in E}$. Assume $\text{rank}\left(Dh_E\left(x^*\right)\right) = e$ where e is the number of elements in E.

Then there exists a vector $\lambda^* = \left(\lambda_1^*, ..., \lambda_l^*\right) \in \mathbb{R}^l$ such that the following conditions are met,

KT1 $\lambda_i^* \geqslant 0$, $h_i(x) \leqslant 0$ and $\lambda_i^* h_i(x^*) = 0$ for all $i = 1, ..., l$.

KT2 If dealing with a maximisation problem: $Df(x^*) - \sum_{i=1}^{l} \lambda_i^* Dh_i(x^*) = 0$.

KT2' If dealing with a minimisation problem: $Df(x^*) + \sum_{i=1}^{l} \lambda_i^* Dh_i(x^*) = 0$.

5.3.2 Some considerations about the Kuhn-Tucker Theorem

- This theorem provides *necessary conditions* only.

- The Kuhn-Tucker multipliers can be given a similar interpretation as the Lagrange multipliers: they measure the marginal value (as assessed by the optimising function) of relaxing the restriction they are associated with.
  - If constraint i is not binding/effective then $h_i(x^*) < 0$ and the value of relaxing it is zero, $\lambda_i^* = 0$; i.e. $i \notin E$ implies $\lambda_i = 0$.
  - If constraint i is binding/effective then $h_i(x^*) = 0$ and the value of relaxing it is positive in maximisation problems ($\lambda_i^* > 0$) and negative in minimisation problems ($-\lambda_i^* < 0$). Note that this really depends on how the constraints h_i are defined.
- As in the Theorem of Lagrange, the 4th condition is called the *constraint qualification*. Now, however, only effective (binding) constraints matter in the statement of this assumption since the others do not affect the solution to the problem. It ensures that there exist x such that $h(x) \leqslant 0$.

5.3.3 Procedure to identify the optimum

This is very similar to the procedure used when only equality constraints are considered.

- Given an optimisation problem

$$\begin{aligned} \max_x \quad & f(x) \\ \text{s.t. } & x \in \mathcal{D} = S \cap \{x \in \mathbb{R}^n : h(x) \leqslant 0\} \end{aligned}$$

  write the Lagrangian equation,

$$\mathcal{L}(x;\lambda) = f(x) - \sum_{i=1}^{l} \lambda_i h_i(x)$$

- Write down the FOCs at $(x^{*'}, \lambda^{*'})$,

$$\frac{\partial \mathcal{L}(x^*;\lambda^*)}{\partial x_i} = 0 \quad \text{for} \quad i = 1, ..., n \tag{14}$$

$$\frac{\partial \mathcal{L}(x^*;\lambda^*)}{\partial \lambda_i} \geqslant 0 \quad \text{for} \quad i = 1, ..., l \tag{15}$$

$$\lambda_i^* h_i(x^*) = \lambda_i^* \frac{\partial \mathcal{L}(x^*;\lambda^*)}{\partial \lambda_i} = 0 \quad \text{for} \quad i = 1, ..., l \tag{16}$$

$$\lambda_i^* \geqslant 0 \quad \text{for} \quad i = 1, ..., l \tag{17}$$

- The last two conditions are called the *complementarity or slackness conditions*: at the optimum, either $\lambda_i^* = 0$ or $h_i(x^*) = 0$.

- We then solve the system of equations (14)-(17) for x^* and λ^*.
- And compute the value of f at all values of x^* that are solutions to the system of equations and compare to find the optimum.

5.3.4 Dealing with non-negativity constraints on variables

- Lets assume one of the restrictions in our maximisation problem assumes the simple form,

$$x_1 \geqslant 0$$

- We may deal with such restriction in the same way as we do with any other inequality constraint.
- Or we may notice that there are two possibilities:
  - either it is not binding, and the optimum is an interior solution, meaning that

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0$$

  - or it is binding, and the optimum is at the corner $x_1 = 0$ in which case

$$\frac{\partial \mathcal{L}}{\partial x_1} \text{ may be } \neq 0$$

- We can then construct a *complementarity condition* just as we did for the Lagrange multipliers,

$$x_1 \geqslant 0 \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial x_1} \leqslant 0 \quad \text{and} \quad x_1 \frac{\partial \mathcal{L}}{\partial x_1} = 0$$

5.3.5 When does the Kuhn-Tucker approach work?

- Under the assumptions described in the Kuhn-Tucker theorem, the Kuhn Tucker conditions will identify the optimum (and will possibly have other non-optimal solutions).
- But the maximisation problem might not have a solution, in which case there might or might not be solutions to the Kuhn-Tucker conditions.
- The existence of a solution is guaranteed under some additional conditions:

  Result: If f is a concave function, $h_i, i = 1, \ldots, l$ are convex functions and the feasibility set is non-empty (there exists $x : h(x) \leqslant 0$), then there exists a solution to the maximisation problem and the Kuhn-Tucker conditions are necessary and sufficient for optimality.

Remember that the optimisation process in greatly simplified under some convexity conditions. Some other results are:

- Suppose $\mathcal{D} \in \mathbb{R}^n$ is convex and $f : \mathcal{D} \to \mathbb{R}$ is concave. Then,
  - Any local maximum of f is a global maximum.
  - The set $\operatorname{argmax}\{f(x) : x \in \mathcal{D}\}$ is either empty or convex.
- Suppose $\mathcal{D} \in \mathbb{R}^n$ is convex and $f : \mathcal{D} \to \mathbb{R}$ is strictly concave. Then the set $\operatorname{argmax}\{f(x) : x \in \mathcal{D}\}$ is either empty or contains a single point.

5.3.6 Example

Consider the following utility maximization problem. Let $u(x_1, x_2) = \alpha x_1 + \beta \ln(x_2)$, for $0 < \alpha, \beta < \infty$ and $x' = (x_1, x_2) \in \mathbb{R}^2_+$. Let $p > 0$ denote the relative price of x_1 and let y be income (in terms of units of x_2), so that the budget constraint is $px_1 + x_2 \leq y$. Then, the UMP is

$$\max_{x_1, x_2} \alpha x_1 + \beta \ln(x_2) \text{ s.t. } px_1 + x_2 \leq y,\ x_1 \geq 0,\ x_2 \geq 0.$$

Note that the upper bounds on x_1 and x_2 result from the lower bounds and the budget constraint holding with equality.

Define

$$\begin{aligned} g_1(x_1, x_2) &= px_1 + x_2 - y \leq 0, \\ g_2(x_1, x_2) &= -x_1 \leq 0, \\ g_3(x_1, x_2) &= -x_2 \leq 0, \\ g(x_1, x_2)' &= [g_1(x_1, x_2), g_2(x_1, x_2), g_3(x_1, x_2)]. \end{aligned}$$

Then, for $\lambda' = (\lambda_1, \lambda_2, \lambda_3)$,

$$\begin{aligned} \mathcal{L}(x; \lambda) &= u(x_1, x_2) - g(x_1, x_2)'\lambda \\ &= \alpha x_1 + \beta \ln(x_2) - \lambda_1(px_1 + x_2 - y) - \lambda_2(-x_1) - \lambda_3(-x_2). \end{aligned}$$

Maximization with respect to x and λ yields the FOCs at the solution $(x^{\star\prime}, \lambda^{\star\prime})$:

$$\begin{aligned} \alpha - \lambda_1^\star p + \lambda_2^\star &= 0 \\ \beta \frac{1}{x_2^\star} - \lambda_1^\star + \lambda_3^\star &= 0 \\ \lambda_1^\star (px_1^\star + x_2^\star - y) &= 0 \\ \lambda_2^\star x_1^\star &= 0 \\ \lambda_3^\star x_2^\star &= 0 \\ \lambda_i^\star &\geq 0, \quad i = 1, 2, 3. \end{aligned}$$

Note, first, that $\lambda_2^\star \geq 0$ and $\alpha > 0$ implies $\lambda_1^\star > 0$. Similarly, $\lambda_3^\star \geq 0$ and $\beta > 0$ implies $\lambda_1^\star > 0$. So, $\lambda_1^\star$, the shadow value of income, is positive, as a consequence of utility $u(x_1, x_2)$ being strictly monotonic (i.e. $\alpha, \beta > 0$). Therefore, $\lambda_1^\star(px_1^\star + x_2^\star - y) = 0$ implies that the budget constraint holds with equality, i.e. $px_1^\star + x_2^\star = y$.

Next, observe that there are now the following possibilities for the multipliers on the non-negativity constraints:

- $\lambda_2^\star > 0$, which is equivalent to $\alpha - \lambda_1^\star p < 0$, or $\frac{\alpha}{p} < \lambda_1^\star$, in which case $x_1^\star = 0$ and $x_2^\star = y$; in this case, $\lambda_3^\star = 0$ must hold, i.e. $\beta\frac{1}{x_2^\star} = \frac{\beta}{y} = \lambda_1^\star$;
- or $\lambda_2^\star = 0$, i.e. $\frac{\alpha}{p} = \lambda_1^\star$, in which case $x_1^\star > 0$ and $x_2^\star = y - px_1^\star \geq 0$, with $\lambda_3^\star \geq 0$;
  - if $\lambda_3^\star > 0$, i.e. $\beta\frac{1}{x_2^\star} \leq \lambda_1^\star$, then $x_2^\star = 0$ and $x_1^\star = y/p$; but this is impossible, because $x_2^\star = 0$, together with $\beta\frac{1}{x_2^\star} \leq \lambda_1^\star$, would imply that $\lambda_1^\star = +\infty$, but this is incompatible with $\lambda_1^\star = \frac{\alpha}{p} < \infty$;
  - if $\lambda_3^\star = 0$, i.e. $\beta\frac{1}{x_2^\star} = \lambda_1^\star$, then $x_1^\star, x_2^\star > 0$, satisfying the budget constraint with equality.

Conclusion: Either (i) $x_1^\star = 0$ and $x_2^\star = y$, or (ii) interior solution.

(i) $\lambda_2^\star > 0$, or $\frac{\alpha}{p} < \lambda_1^\star$, with $\lambda_3^\star = 0$, i.e. $\frac{\beta}{x_2^\star} = \frac{\beta}{y} = \lambda_1^\star$, implies

$$\frac{\alpha}{p} \leq \lambda_1^\star = \frac{\beta}{y} \quad \Leftrightarrow \frac{\alpha}{\beta/y} < p;$$

(ii) $\lambda_2^\star = 0$, or $\frac{\alpha}{p} = \lambda_1^\star$, with $\lambda_3^\star = 0$, i.e. $\frac{\beta}{x_2^\star} = \lambda_1^\star$, implies

$$\frac{\alpha}{p} = \frac{\beta}{x_2^\star},$$

which, in turn, implies $x_2^\star = \frac{\beta p}{\alpha}$, and hence $x_1^\star = (y - x_2^\star)/p = \frac{y}{p} - \frac{\beta}{\alpha}$, which is positive if, and only if, $\frac{\alpha}{\beta/y} > p$.

So, whether or not $\frac{\alpha}{\beta/y}$ exceeds p determines whether the problem exhibits a corner or interior solution.

5.4 The General Case: Optimization With Mixed Constraints

Lets now consider the following type of optimisation problems,

$$\begin{aligned} \max_{x \in S \subset \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & g_i(x) = 0 \text{ for } i = 1, \ldots, k \\ & h_j(x) \leqslant 0 \text{ for } j = 1, \ldots, l \end{aligned}$$

which is the same as considering the constraint set $\mathcal{D}$ to be,

$$\mathcal{D} = S \cap \left\{ x \in \mathbb{R}^n : g_i(x) = 0, h_j(x) \leqslant 0 \text{ for } i = 1, \ldots, k \text{ and } j = 1, \ldots, l \right\}$$

For ease of notation, define the function $w : \mathbb{R}^n \to \mathbb{R}^{k+l}$,

$$w_i(x) = \begin{cases} g_i(x) & \text{if } i \leqslant k \\ h_i(x) & \text{if } i > k \end{cases}$$

5.4.1 Necessary conditions for optimality with mixed constraints

Consider the following assumptions,

1. S is an open set.
2. The functions $f : S \subset \mathbb{R}^n \to \mathbb{R}$ and $w : \mathbb{R}^n \to \mathbb{R}^{k+l}$ are $\mathcal{C}^1$.
3. x^* is an interior local maximum of f in
   $\mathcal{D} = S \cap \{x \in \mathbb{R}^n : w_i(x) = 0 \text{ for } i = 1, ..., k \text{ and } w_i(x) \leqslant 0 \text{ for } i = k+1, ..., k+l\}$.
4. Let $E \subset 1, ..., k+l$ denote the set of effective constraints at x^* and $w_E = (w_i)_{i \in E}$. Assume rank $(Dw_E(x^*)) = e$ where e is the number of elements in E.

Then there exists a vector $\lambda^* = \left(\lambda_1^*, ..., \lambda_{k+l}^*\right) \in \mathbb{R}^l$ such that the following conditions are met,

1. $w_i(x^*) = 0$ for all $i = 1, ..., k$.
2. $\lambda_i^* \geqslant 0$, $w_i(x) \leqslant 0$ and $\lambda_i^* w_i(x^*) = 0$ for all $i = k+1, ..., k+l$.
3. $Df(x^*) - \sum_{i=1}^{k+l} \lambda_i^* Dw_i(x^*) = 0$.

6 Dynamic Optimization

References: Adda and Cooper, chapt. 2; Bertsekas (1976), Sargent (1987), Stokey and Lucas (1989)

6.1 Discrete Time, Finite Horizon Problems

Setup:

- Given a cake of size W_1.
- At each time period $t = 1, 2, \cdots, T < \infty$, some of the cake can be eaten, resulting in consumption c_t, and the rest is saved for the future.

- Let $u(c_t)$ be flow utility derived form c_t. Assume $u(\cdot)$ is real-valued, differentiable, strictly increasing and strictly concave, satisfying $\lim_{c\to 0} u'(c) = \infty$. Note: Utility $u(\cdot)$ is not indexed by t, i.e. is is *stationary*.
- Finite-lifetime utility is separable across t and given by $\sum_{t=1}^{T} \beta^{t-1} u(c_t)$, where $\beta \in (0,1)$ is a discount factor.

This setup implies that the *state* of the world is captured by the size of the cake at any given time, evolving as

$$W_{t+1} = W_t - c_t \quad t = 1, \cdots, T-1, \tag{18}$$

where $W_t \geq 0$ and $c_t \geq 0$ for $t = 1, \cdots, T$. The state is therefore seen to be a function (or the dual) of the cake eater's decision how much to consume at any given time t, c_t. The consumption path $\{c_t, t = 1, \cdots, T\}$ is what the decision maker controls or optimizes.

Question of interest: What is optimal consumption path $\{c_t^\star, t = 1, \cdots, T\}$? I.e. we seek the solution to

$$\max_{\{c_t, t=1,\cdots,T\}} \sum_{t=1}^{T} \beta^{t-1} u(c_t), \quad \text{given } W_1.$$

Notice first: Since $0 \leq c_t \leq W_1 < \infty$ and $u(\cdot)$ is assumed to be continuous (it must be, since it is assumed to be differentiable), this problem has a solution, by virtue of the Weierstrass Theorem.

6.2 Direct Attack Approach

6.2.1 General Outline: Euler Equations (Necessity) and Sufficiency

Aggregating the constraints imposed by the evolution equation (18) yields

$$W_{T+1} + \sum_{t=1}^{T} c_t = W_1,$$

i.e. the final left-over $W_{T+1} \geq 0$ and the sum of the temporal consumptions c_t, $= 1, \cdots, T$, must add up to what is feasible, W_1.

Setting up the Lagrangian,

$$\mathcal{L}(\mathbf{c}, \lambda) = \sum_{t=1}^{T} \beta^{-1} u(c_t) + \lambda \left[W_1 - W_{T+1} + \sum_{t=1}^{T} c_t \right] + \phi W_{T+1},$$

where $\mathbf{c} = [c_1, \cdots, c_T]$. Note that, as $\lim_{c\to 0} u'(c) = \infty$ by assumption, it will never be optimal to set $c_t = 0$, so the non-negativity constraints on c_t are ignored.

The FOCs are

$$\begin{aligned}\beta^{t-1}u'(c_t^\star) &= \lambda \quad t=1,\cdots,T,\\ \lambda^\star &= \phi^\star.\end{aligned}$$

Combining yields the necessary conditions for optimality, called the *Euler equations*

$$u'(c_t^\star)=\beta u'(c_{t+1}^\star) \quad t=1,\cdots,T-1.$$

These equations relate optimal consumption in adjacent periods, but also across time more generally: For any s,

$$u'(c_t^\star)=\beta^s u'(c_{t+s}^\star), \quad t=1,\cdots,T-s.$$

Satisfying the Euler equations is not sufficient for optimality, because any candidate solution that yields $W_{T+1}=W_T-c_T^\star>0$ is sub-optimal, because the remainder, if consumed, would strictly increase utility (which is strictly increasing by assumption). This is clear from $\phi^\star=\lambda^\star=\beta^{t-1}u'(c_t^\star)>0$, so that the complementary slackness condition $\phi^\star W_{T+1}=0$ implies $W_{T+1}=0$.

In fact, the initial state W_1 and the final state W_{T+1} pin down the end points of the optimal consumption path, and the Euler equations yield the optimal controls $\{c_t^\star, t=1,\cdots,T\}$.

6.2.2 Example: 3-Period Problem

Assume $T=3$ and $u(c)=\ln(c)$. This utility function clearly satisfies $\lim_{c\to 0}u'(c)=\lim_{c\to 0}\frac{1}{c}=\infty$. Then, the Euler equations are

$$\frac{1}{c_1}=\beta\frac{1}{c_2}=\beta^2\frac{1}{c_3},$$

and the resource constraint is $W_1 = c_1+c_2+c_3$, where $W_4=0$ is imposed. These are three equations in three unknowns, with solution

$$c_1^\star=\frac{W_1}{1+\beta+\beta^2},\quad c_2^\star=\frac{\beta W_1}{1+\beta+\beta^2},\quad c_3^\star=\frac{\beta^2 W_1}{1+\beta+\beta^2}.$$

The value function of this problem at time $t=1$ is

$$\begin{aligned}V_1(W_1) &= \ln(c_1^\star)+\beta\ln(c_2^\star)+\beta^2\ln(c_3^\star)\\ &= \ln\left(\frac{W_1}{1+\beta+\beta^2}\right)+\beta\ln\left(\frac{\beta W_1}{1+\beta+\beta^2}\right)+\beta^2\ln\left(\frac{\beta^2 W_1}{1+\beta+\beta^2}\right)\\ &\propto (1+\beta+\beta^2)\ln(W_1).\end{aligned}$$

An infinitesimal increase in the feasible amount W_1 yields an increase in maximum utility of

$$V_1'(W_1)=(1+\beta+\beta^2)\frac{1}{W_1}.$$

Saw before:

$$\begin{aligned} V_1'(W_1) &= \lambda^\star = \beta^{t-1}u'(c_t^\star) \quad t = 1,2,3 \\ &= u'(c_1^\star) = \frac{1}{c_1^\star} = (1+\beta+\beta^2)\frac{1}{W_1}, \end{aligned}$$

so it does not matter in which period the increment on W_1 is consumed, as long as the consumption path retains its optimality via the Euler equations.

Notice that the value function $V_1(W_1)$ essentially captures all there is to know at time $t = 1$ about the T-period problem. This feature, known as the *principle of optimality* due to Richard Bellman, implies that the value function has a recursive structure:

$$V_1(W_1) = \ln(c_1^\star) + \beta V_2(W_2),$$

where $W_2 = W_1 - c_1^\star = \frac{\beta(1+\beta)}{1+\beta+\beta^2}W_1$, and therefore $\frac{\beta W_1}{1+\beta+\beta^2} = \frac{W_2}{1+\beta}$, and

$$\begin{aligned} V_2(W_2) &= \ln(c_2^\star) + \beta\ln(c_3^\star) \\ &= \ln\left(\frac{\beta W_1}{1+\beta+\beta^2}\right) + \beta\ln\left(\frac{\beta^2 W_1}{1+\beta+\beta^2}\right) \\ &= \ln\left(\frac{W_2}{1+\beta}\right) + \beta\ln\left(\frac{\beta W_2}{1+\beta}\right) \\ &\propto (1+\beta)\ln(W_2). \end{aligned}$$

Therefore, the two-period "short-cut"

$$\max_{c_1} \ln(c_1) + \beta V_2(W_2) = \max_{c_1} \ln(c_1) + \beta V_2(W_1 - c_1)$$

yields the FOC

$$\frac{1}{c_1^\star} = \beta V_2'(W_1 - c_1^\star) = \beta(1+\beta)\frac{1}{W_1 - c_1^\star},$$

which, in turn, yields the optimal period 1 consumption $c_1^\star = \frac{W_1}{1+\beta+\beta^2}$.

6.3 Dynamic Programming Approach

6.3.1 General Outline

Suppose there is an additional period, $t = 0$, with initial endowment W_0. Instead of re-iterating the direct attack approach, can previous results be employed?

The dynamic programming approach allows to do this, by essentially converting the $T+1$-period problem into a 2-period problem, using the recursive structure of the value function.

Analogous to the preceding example, in the general setup for the $T+1$ period problem, given W_0, consider the two-period "short-cut"

$$\max_{c_0} u(c_0) + \beta V_1(W_1),$$

where $W_1 = W_0 - c_0$. Assuming differentiability of the value function, the FOC is

$$u'(c_o^\star) = \beta V_1'(W_0 - c_0^\star).$$

Know from the T-period problem that $V_1'(W_1) = \lambda = \beta^{t-1}u'(c_t^\star)$, $t = 1, \cdots, T$, and therefore

$$u'(c_0^\star) = \beta u'(c_1^\star) = \beta^s u'(c_s^\star), \quad s = 0, \cdots, T,$$

i.e. the necessary conditions for optimality (Euler equations) are satisfied.

This approach requires knowledge of the (sequence) of value function(s) $V_1(W_1)$ ($\{V_s(W_s), s = 1, \cdots, T\}$). The value function can typically be derived by backward recursion, starting from a single period problem and successively adding periods.

6.3.2 Example: 3-Period Problem (re-visited)

To build up the value function, first consider the case $T = 2$, starting in period $t = 2$, with $t = 2, 3$, given W_2. The Euler equation yields

$$\frac{1}{c_2} = \frac{\beta}{c_3},$$

and the resource constraint is $W_2 = c_2 + c_3$. Solving these two equations in two unknowns yields

$$c_2\star = \frac{W_2}{1+\beta}, \quad c_3^\star = \frac{\beta W_2}{1+\beta}.$$

This implies the value function for the problem starting at time $t = 2$,

$$\begin{aligned} V_2(W_2) &= \ln(c_2^\star) + \beta \ln(c_3^\star) \\ &= \ln\left(\frac{W_2}{1+\beta}\right) + \beta\left(\frac{\beta W_2}{1+\beta}\right) \\ &= (1+\beta)\ln(W_2) + \ln\left(\frac{1}{1+\beta}\right) + \beta \ln\left(\frac{\beta}{1+\beta}\right) \\ &\propto (1+\beta)\ln(W_2) \end{aligned}$$

The value function $V_2(W_2)$ can then be used in the Bellman style recursion when the first period is added,

$$V_1(W_1) = \max_{c_1} \ln(c_1) + \beta V_2(W_1 - c_1).$$

The FOC is

$$\begin{aligned} \frac{1}{c_1^\star} &= \beta V_2'(W_1 - c_1^\star) \\ &= \beta(1+\beta)\frac{1}{W_1 - c_1^\star}, \end{aligned}$$

and, as was seen before, $c_1^\star = \frac{W_1}{1+\beta+\beta^2}$. The Euler equations then yield $c_2^\star = \frac{\beta W_1}{1+\beta+\beta^2}$ and $c_3^\star = \frac{\beta^2 W_1}{1+\beta+\beta^2}$, as before.

The solution extends to the T-period version of this example in a straightforward manner:

$$c_t = \frac{\beta^{t-1} W_1}{\sum_{s=1}^{T} \beta^{s-1}}, \tag{19}$$

$$V_t(W_t) \propto \left(\sum_{s=1}^{t} \beta^{s-1} \right) \ln(W_t), \quad t = 1, \cdots, T. \tag{20}$$

6.4 Infinite Horizon Problems

6.4.1 Bellman Equation and Policy Function

Suppose the time horizon of the previous cake-eating problem is extended to the infinite future, so that the task is

$$\max_{\{c_t, t \geq 1\}} \sum_{t=1}^{\infty} \beta^{t-1} u(c_t),$$

given $W_1 = W$.

The difference to the finite horizon problem is that the value function V of the infinite horizon problem itself is now stationary, so that now the Bellman equation does not involve time (indices):

$$V(W) = \max_{c \in [0, W]} u(c) + \beta V(W - c) \quad \text{for any } W \geq 0,$$

where the transition equation $W' = W - c$ denotes the remainder after c has been consumed from state W, and as before $\beta \in (0, 1)$.

Here, as before, c denotes the control (or choice) variable, while W denotes the state variable. The two are related via the transition equation. While it is more natural to think of optimizing with respect to the path of the control variable c, one can cast the problem equivalently as an optimization over the path of the state variable. The dual representation, in terms of the state variable, is

$$V(W) = \max_{W' \in [0, W]} u(W - W') + \beta V(W') \quad \text{for any } W \geq 0.$$

Note that, in either representation, the Bellman equation is a *functional equation* because it is used not only to find certain values that form a solution to an optimization problem, but it is used to find the entire value function $V(W)$, for any given $W \geq 0$. The solution is a *fixed point* of the functional equation.

Assuming a solution $V(\cdot)$ exists, the FOC is as before,

$$u'(c^\star) = \beta V'(W - c^\star) = \beta V'(W').$$

As seen before,

$$V'(W) = u'(c^\star)$$

and as this holds for any W (or time),

$$V'(W') = u'(c^{\star'}),$$

so that, combining, this implies the familiar Euler equation:

$$u'(c^\star) = \beta u'(c^{\star'}).$$

The link between the state and control variables is given by the policy function $\phi(\cdot)$,

$$c = \phi(W),$$

so that the transition equation for the state variable can be expressed as

$$W' = \psi(W) = W - \phi(W).$$

By iteration (function composition), this yields $c' = \phi(W') = \phi(W - \phi(W))$. Substituting into the Euler equation casts the problem in terms of the policy function,

$$u'(\phi(W)) = \beta u'(\phi(W - \phi(W)), \quad \text{for any } W \geq 0.$$

In policy analysis, the policy function typically is the object of primary interest.

6.4.2 Example: T-Period Problem (re-visited), in the limit

From the analysis of the various T-period problems, a plausible conjecture of the value function is that it is of the form

$$V(W) = A + B\ln(W),$$

for some A and B which are functions of β.

Using this conjecture, the Bellman equation in the representation using the state variable is

$$A + B\ln(W) = \max_{W'} \ln(W - W') + \beta\left(A + B\ln(W')\right), \quad \text{for any } W \geq 0. \tag{21}$$

The FOC, after differentiation w.r.t. W, is

$$\begin{aligned} \frac{1}{W - W'} &= \beta B \frac{1}{W'} \\ \Rightarrow \quad W' &= \frac{\beta B}{1 + \beta B} W = \psi(W). \end{aligned}$$

Substituting this expression for W' back into the conjecture (21) for the value function yields

$$\begin{aligned} A + B\ln(W) &= \ln\left(\frac{W}{1 + \beta B}\right) + \beta\left(A + B\ln\left(\frac{\beta B W}{1 + \beta B}\right)\right), \quad \text{for any } W \geq 0, \\ &\propto (1 + \beta B)\ln(W) \end{aligned}$$

so that, matching coefficients on $\ln(W)$, it follows that $B = 1 + \beta B$, or $B = \frac{1}{1-\beta}$, and hence

$$V(W) \propto \frac{1}{1-\beta}\ln(W).$$

Using the fact that $V'(W) = u'(c^\star)$, it follows that

$$\begin{aligned} \frac{1}{1-\beta}\frac{1}{W} &= \frac{1}{c^\star} \\ \Rightarrow \quad c^\star &= (1-\beta)W = \phi(W) \\ \Rightarrow \quad W' &= W - \phi(c^\star) = \beta W = \psi(W), \end{aligned}$$

which is the optimal policy function and transition equation.

Note that these solutions are the natural limits of the finite horizon solutions (20) and (19): At $t = 1$, with $W_1 = W$, i.e. (20) and (19) imply

$$\begin{aligned} \lim_{T\to\infty}\left(\sum_{s=1}^{T}\beta^{s-1}\right)\ln(W) &= \frac{\ln(W)}{1-\beta} = V(W) \\ \lim_{T\to\infty}\frac{W}{\sum_{s=1}^{T}\beta^{s-1}} &= (1-\beta)W = \phi(W). \end{aligned}$$

6.5 General Formulation

Consider the infinite horizon problem for an agent with a period t pay-off function $\sigma(s_t, c_t)$, which, in the general setting, is allowed to depend on the state variable s_t, next to the conventional control variable c_t.

The transition equation is $s_{t+1} = \tau(s_t, c_t)$. Note that s_t captures everything that at time t is relevant about the past.

Assume $s \in S$ and $c \in C(s) \subset C$, where the choice possibilities at a given time may be constrained by the state at that time, and C and S are assumed compact.

Assume also that $\sigma(s, c)$ is bounded on $S \times C$.

Again, it is assumed that the problem is stationary, i.e. none of the primitives (pay-off function $\sigma(\cdot,\cdot)$, transition equation $\tau(\cdot,\cdot)$) depend on t, and consequently the value function does not depend on t either.

And the discount factor β satisfies $\beta \in (0, 1)$.

The decision problem is

$$\max_{\{c_t, t\geq 0\}}\sum_{t=0}^{\infty}\beta^{t-1}\sigma(s_t, c_t) \quad \text{s.t.} \quad s_{t+1} = \tau(s_t, c_t), \quad t \geq 0.$$

Using the dynamic programming approach, the Bellman equation in the representation using the control variable is

$$V(s) = \max_{c \in C(s)} \sigma(s,c) + \beta V(s') \quad \text{for all } s \in S,$$

where $s' = \tau(s,c)$.

Equivalently, the Bellman equation can be expressed in the representation using the state variable. For this, the transition equation is inverted to back out $c = \tau_c^{-1}(s,s')$, where $\tau_c^{-1}(s,\cdot)$ denotes the inverse of $\tau(s,c)$ with respect to c. Then, define

$$\begin{aligned} \gamma(s,s') &= \sigma\left(s, \tau_c^{-1}(s,s,')\right) \\ \Gamma(s) &= \{s' : c = \tau^{-1}(s,s') \in C(s)\}, \end{aligned}$$

and use these to re-express the Bellman equation as

$$V(s) = \max_{s' \in \Gamma(s)} \gamma(s,s') + \beta V(s'). \tag{22}$$

Denote the optimal policy function in the state variable representation by $s' = \psi(s)$. Combining with the transition equation yields the optimal policy function linking state and control variables, $c = \tau^{-1}(s, \psi(s)) = \phi(s)$.

Two results of interest:

> Theorem: (Stokey and Lucas, 1989) Assume that $\gamma(s,s')$ is real-valued, continuous, and bounded, $\beta \in (0,1)$ and $\Gamma(s)$ non-empty for all $s \in S$, compact and continuous in s. Then there exists a unique value function $V(s)$ that solves (22).
>
> This result follows from the fact that, with $\beta \in (0,1)$, the search can be cast as a contraction mapping $\mathcal{T}$, defined as
>
> $$\mathcal{T}(W)(s) = \max_{s' \in \Gamma(s)} \gamma(s,s') + \beta W(s') \quad \text{for any } s \in S.$$
>
> The mapping $\mathcal{T}$ starts with a guess for the value function, $W(\cdot)$, and upon maximization w.r.t. s' over $\Gamma(s)$ produces a function $\mathcal{T}(W)(s)$, $s \in S$, which may or may not agree with $W(s)$. If it does, i.e. when $W(s) = \mathcal{T}(W)(s)$ for all $s \in S$, then this is a fixed point in the space of candidate value functions and hence a solution. The Theorem states that, under the stated assumptions, there exists such a solution, and this solution is unique.
>
> The argument demonstrating the existence of a fixed point in the function space is based on a result by Blackwell (1965) which relies on appropriate discounting, implied by $\beta \in (0,1)$, and monotonicity of the operator $\mathcal{T}$.

- Theorem: (Stokey and Lucas, 1989) Assume that $\gamma(s, s')$ is real-valued, continuous, and bounded, $\beta \in (0, 1)$ and S is a convex subset of $\mathbb{R}^k$, and that $\Gamma(s)$ is non-empty for all $s \in S$, as well as compact, convex, and continuous. Then the unique solution to (22) is strictly concave. Furthermore, $\psi(s)$ is a continuous and single-valued function.

6.6 Applied Example: Optimal Growth

References: Brock and Mirman (1972), Koopmans (1963), Cass (1965)

Suppose capital at time t, k_t, can be used to produce output using a technology characterized by the production function $f(cdot)$ that can be consumed at time t, denoted by c_t, or turned into period $t+1$'s capital k_{t+1}; i.e. $f(k_t) = c_t + k_{t+1}$. A representative consumer's objective is to maximize discounted future flow utilities of consumption $u(c_t), t = 1, 2, \cdots$, subject to these temporal resource constraints, i.e.

$$\max_{\{c_t, t=0,1,\cdots\}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$
$$\text{s.t.} c_t + k_{t+1} = f(k_t), k_0 \text{ given}, c_t \geq 0.$$

Here, $0 < \beta < 1$, $u(\cdot)$ satisfies $\lim_{c\to 0} u'(c) = +\infty$, $u'(\cdot) > 0$ and $u''(\cdot) < 0$, and $f(\cdot)$ satisfies $\lim_{k\to 0} f'(k) = +\infty$, $\lim_{k\to\infty} f'(k) = 0$, $f'(\cdot) > 0$ and $f''(\cdot) < 0$.

The Bellman equation in state variable representation is

$$V(k_t) = \max_{k_{t+1}} u\left(f(k_t) - k_{t+1}\right) + \beta V(k_{t+1}).$$

This yields the FOC:

$$u'\left(f(k_{t+1}) - k_{t+1}\right) = \beta V'(k_{t+1}) = u'\left(f(k_{t+1}) - \psi(k_t)\right), \tag{23}$$

where $k_{t+1} = \psi(k_t)$ is the transition equation characterizing the optimal path of the state variable.

Moreover, the shadow value of capital is

$$V'(k_t) = u'\left(f(k_t) - k_{t+1}\right) f'(k_t) > 0. \tag{24}$$

Is there a maximal capital stock $\bar{k}$ which can be sustained as a possible stationary long-run solution? If so, it must be the case that, eventually, $\bar{k} = f(\bar{k})$ and $c = 0$. Since $f(\cdot)$ is strictly concave, with $\lim_{k\to 0} f'(k) = +\infty$, $0 < \bar{k} < \infty$ exists. And it is easy to see that $k_{t+1} = f(k_t) \to \bar{k}$ as $t \to \infty$. So, if $k_0 \in (0, \bar{k}]$, then $k_t \in (0, \bar{k}]$ for all t. Assume henceforth that $k_0 \in (0, \bar{k}]$.

Observe that $\gamma(k_t, k_{t+1}) = \ln\left(f(k_t) - k_{t+1}\right)$ is real-valued, continuous and bounded for $k_t \in (0, \bar{k}] = S$, a convex set. Hence, by the aforementioned Stokey-Lucas Theorem, $V(k)$ is concave and $\psi(k)$ is continuous and single-valued.

Note that (24), together with continuity of $f(\cdot)$, $f'(\cdot)$ and $u'(\cdot)$, implies continuity of $V'(\cdot)$.

Claim: $\psi(k)$ is non-decreasing in k.

Proof: Consider k_i, $i = 1, 2$. By (23),

$$u'\left(f(k_i) - \psi(k_i)\right) = \beta V'(\psi(k_i)), \quad i = 1, 2.$$

Suppose $k_1 > k_2$ and, to the contrary of the claim, $\psi(k_1) < \psi(k_2)$. Concavity of $V(\cdot)$ and $u(\cdot)$ implies

$$\begin{aligned} & V'(\psi(k_1)) > V'(\psi(k_2)) \\ \Rightarrow\ & u'\left(f(k_1) - \psi(k_1)\right) > u'\left(f(k_2) - \psi(k_2)\right) \\ \Rightarrow\ & f(k_1) - \psi(k_1) < f(k_2) - \psi(k_2) \\ \Rightarrow\ & \psi(k_2) - \psi(k_1) < f(k_2) - f(k_1) < 0, \end{aligned}$$

which is a contradiction to the supposition and thereby proves the claim.

Since $k_t \in (0, \bar{k}]$ for all t, $\psi(k_t)$ is bounded. Since $\psi(k)$ is non-decreasing, the sequence $\{\psi(k_t), t = 0, 1, 2, \cdots\}$ is monotone. Therefore, this sequence converges to a limit k_∞.

Does the limit k_∞ depend on the initial condition k_0? Eventually, $k_{t+1} = k_t = k_\infty$, and (23) implies $u'\left(f(k_\infty) - k_\infty\right) = \beta V(k_\infty)$, while (24) implies $u'\left(f(k_\infty) - k_\infty\right) f'(k_\infty) = V'(k_\infty)$. Hence,

$$\beta f'(k_\infty) = 1 \tag{25}$$

determined the unique stationary state k_∞, independent of k_0 (and, incidentally, independent of $u(\cdot)$).

Proceeding with the functional form assumptions $u(c) = ln(c)$ and $f(k) = Ak^\alpha$, for $A > 0$ and $0 < \alpha < 1$, consider the guess for the value function $V(k) = E + F \ln(k)$, for some $E, F > 0$. By (23), this guess implies

$$\frac{1}{Ak_t^\alpha - k_{t+1}} = \beta F \frac{1}{k_{t+1}},$$

and so $k_{t+1} = \frac{\beta F}{1+\beta F} A k_t^\alpha$. Therefore,

$$\begin{aligned} E + F \ln(k_t) &= \ln\left(Ak_t^\alpha \frac{1}{1+\beta F}\right) + \beta\left[E + F \ln\left(\frac{\beta F}{1+\beta F} A k_t^\alpha\right)\right] \\ &\propto (\alpha + \alpha\beta F) \ln(k_t), \end{aligned}$$

and so $F = \alpha + \alpha\beta F = \frac{\alpha}{1-\alpha\beta}$. Then,

$$V(k) \propto \frac{\alpha}{1-\alpha\beta} \ln(k),$$

and the transition equation for the optimal path of the state variable is

$$k_{t+1} = \psi(k_t) = \alpha\beta A k_t^\alpha,$$

and the policy function characterizing the optimal path for the control variable is

$$c_t = \phi(k_t) = (1 - \alpha\beta)Ak_t^{\alpha}.$$

Note that the transition equation for the state variable implies

$$k_\infty = \alpha\beta A k_\infty^{\alpha} = (\alpha\beta A)^{\frac{1}{1-\alpha}}.$$

This is consistent with (25):

$$\beta f'(k_\infty) = \beta\alpha A k_\infty^{\alpha-1} = \alpha\beta A\left(\alpha\beta A\right)^{-1} = 1.$$

Exercise: Consider the finite horizon variant of this problem, i.e. the T-period problem

$$\max_{\{c_t, t=0,1,\cdots,T_1\}} \sum_{t=0}^{T-1} \beta^t u(c_t)$$
$$\text{s.t.} c_t + k_{t+1} = f(k_t),\ t = 0, \cdots, T_1;\ k_0 \text{ given},\ c_t \geq 0.$$

Show that the value function of this problem, $V_T(k_0)$, satisfies

$$V_T(k_0) \propto \alpha \ln(k_0)\left[\sum_{s=0}^{T-1}(\alpha\beta)^s\right].$$

7 Differential Equations

References: Simon and Blume, chapters 24 and 25; Chiang, chapters 14 and 15.

- Dynamic problems are the rule rather than the exception in modern economics. Common examples are:
  - the problem of the consumer deciding about how much to consume today given her forecasts of future income;
  - precautionary savings;
  - the problem of the firm managing its stocks when holding stocks is costly, re-stocking is costly and/or each period demand is uncertain.
  - the problem of investing in education, both from the agent and the government perspectives.
- We are frequently interested in knowing how a variable evolves over time and how its past path influences its future behaviour.
- We first introduce difference equations and will then discuss differential equations.
- The difference between the two lies in how time is considered: difference equations use a discrete version of time while differential equations apply when time is continuous.

7.1 Difference equations

- Consider a world in which time is discrete, $t = 1, 2, 3... \in \mathbb{N}$.
- Let y be a variable that evolves with time and y_t be the specific value this variable assumes at time t.
- The evolution of this variable from $t-1$ to t is described by the difference,

$$\Delta y_t = y_t - y_{t-1}$$

- We can now describe the pattern of evolution of y using this notation. This could be,

$$y_t = g(y_{t-1}, t)$$

  This is an example of a first order *difference equation*.
- Typically, only linear difference equations can be solved analytically. More complicated equations require numerical techniques to be solved.

7.1.1 Solving a first order linear difference equation with constant coefficients

- Lets start with the following simple equation,

$$y_{t+1} = cy_t + d$$

  where c and d are constants.
- We can start from the beginning of times to write,

$$\begin{aligned}
y_1 &= cy_0 + d \\
y_2 &= cy_1 + d = c\left(cy_0 + d\right) + d = c^2 y_0 + cd + d \\
y_3 &= c\left(c^2 y_0 + cd + d\right) + d = c^3 y_0 + c^2 d + cd + d \\
&... \\
y_t &= c^t y_0 + d\sum_{i=0}^{t-1} c^i \\
&= c^t y_0 + d\left[\sum_{i=0}^{\infty} c^i - \sum_{i=t}^{\infty} c^i\right] \\
&= c^t y_0 + d\left[\sum_{i=0}^{\infty} c^i - c^t \sum_{i=0}^{\infty} c^i\right] \\
&= c^t y_0 + d\frac{1-c^t}{1-c}.
\end{aligned}$$

- Hence, assuming $c \neq 0$ the solution to this difference equation is,

$$y_t = c^t y_0 + d\frac{1-c^t}{1-c}$$

while if $c = 0$,

$$y_t = d \text{ for all } t$$

7.2 First order differential equations

However, time is continuous and we are often interested in treating it as such.

When treating time as continuous, we model dynamics using differential equations.

- In this case, we can use derivatives to express the evolution of a variable over time.
- Say we start with the simplest difference equation,

$$y_{t+1} - y_t = cy_t$$

- But since t is continuous, we can apply this rule to a smaller increment in t,

$$y_{t+\Delta t} - y_t = cy_t\Delta t$$

or rearranging

$$\frac{y_{t+\Delta t} - y_t}{\Delta t} = cy_t$$

- Taking the limits as $\Delta t \to 0$,

$$\frac{dy_t}{dt} = cy_t$$

- The above is an example of a *first order homogeneous differential equation*.

7.2.1 Some considerations about differential equations

- Instead of using the time index, the typical notation in differential equations is to explicitly describe t as the argument of the function, $y(t)$.
- Moreover, the argument of the function is frequently omitted. The first derivative with respect to time is frequently written as $\dot{y}$ or y', the second as $\ddot{y}$ or y'', and so on...
- When dealing with differential equations, the argument of the function y does not need to be time. The following is a differential equation, where x is some real variable

$$y'(x) = ay(x) + bx$$

- The above equation is an example of a first-order differential equation because only the first derivative is included in the relation.
- There are many types of differential equations. The general form for a first order differential equation is:

$$y' = g(y,t)$$

and the following are a few special cases:

  - Separable differential equation: $g(y,t) = u(y)v(t)$
  - Linear non-homogeneous differential equation: $g(y,t) = u(t)y + v(t)$
  - Linear homogeneous differential equation: $g(y,t) = u(t)y$
  - Linear differential equation with constant coefficients: $g(y,t) = cy + k$

- If the function $g(y,t)$ does not depend on t, i.e. when it reduces to $g(y,t) = \tilde{g}(y)$ for all t, the differential equation $y' = \tilde{g}(y)$ is called *autonomous* or *time-independent*. If, on the other hand, $g(y,t)$ does explicitly depend on t, the differential equation $y' = g(y,t)$ is called *non-autonomous* or time-dependent.

7.2.2 Solution methods for some first order differential equations

Linear homogeneous first-order differential equations with constant coefficients

- The first order equations of this type can be written as

$$\frac{dy}{dt} = cy.$$

This differential equation is autonomous.

- This is equivalent to write,

$$\frac{1}{y}\frac{dy}{dt} = c$$

- We can now form the indefinite integral on both sides to obtain,

$$\int \frac{1}{y}\frac{dy}{dt}dt = \int c\,dt$$

which yields,

$$\ln y(t) = ct + k \qquad \text{or} \qquad y(t) = e^k e^{ct} = ae^{ct}$$

This is the *general solution*.

- The *definite solution* can be computed if we have some *boundary condition*: $y(0) = y_0$,

$$y_0 = ae^0 = a \qquad \text{and thus} \qquad y(t) = y_0 e^{ct}$$

Linear homogeneous first-order differential equations with variable coefficient

- These sort of differential equations take the form,

$$\frac{dy}{dt} = u(t)y.$$

  They are non-autonomous.

- Proceeding as before, we have,

$$\frac{1}{y}\frac{dy}{dt} = u(t)$$

  which we can integrate to yield,

$$\begin{aligned} \ln y(t) &= \int u(t)dt \\ &= k + U(t) \end{aligned}$$

  where $U(t)$ is the primitive of u, i.e. $u(t) = U'(t)$. But then the solution is:

$$y = e^k e^{U(t)} = Ke^{U(t)}$$

  where K is the integrating constant to be determined using an initial condition.

Separable (not necessarily linear) first-order differential equations

- More generally, we say that a first-order differential equation is *separable* if it can be written as follows,

$$\frac{dy}{dt} = F(t)G(y)$$

- In such case we follow the path described above to say,

$$\frac{1}{G(y)}\frac{dy}{dt} = F(t)$$

  thus leading to the solution,

$$\int \frac{1}{G(y)}dy = \int F(t)dt$$

- Notice that whether this can be explicitly solved for $y(t)$ depends on the functions G and F.

- **Example**: Let $u(x)$ be a utility function, defined over wealth x. Assume $u \in \mathcal{C}^2$. The Arrow-Pratt measure of relative risk-aversion at wealth level x is defined as the elasticity of marginal utility $u'(x)$ w.r.t. wealth,

$$\nu(x) = -\frac{u''(x)x}{u'(x)}.$$

Suppose $\nu(x) = \eta$ for all x, i.e. relative risk aversion is constant. What type of utility function induces

$$\eta = -\frac{u''(x)x}{u'(x)} \quad \text{for all } x?$$

The approach is to treat this as a differential equation for marginal utility $\mu(x) = u'(x)$, solve for $\mu(x)$ and integrate to get $u(x)$. Note that, here, x plays the role of t.

The expression for η implies

$$\eta = -\frac{\mu'(x)x}{\mu(x)} \quad \Rightarrow \quad \mu'(x) = \frac{d\mu(x)}{dx} = -\frac{\eta\mu(x)}{x} = F(x)G(\mu(x)).$$

This is a separable differential equation in $\mu(x)$ and x, with $F(x) \propto \frac{1}{x}$ and $G(\mu(x)) \propto \mu(x)$.

Separating μ from x yields

$$\frac{d\mu(x)}{\mu(x)} = -\eta\frac{dx}{x},$$

and forming the indefinite integral leads to

$$\int \frac{d\mu(x)}{\mu(x)} = -\eta \int \frac{dx}{x}.$$

The solution to these expressions is

$$\ln(\mu(x)) = -\eta \ln(x) + c,$$

for some constant c. Therefore,

$$\mu(x) = bx^{-\eta},$$

where b is another constant.

Since $u'(x) = \mu(x)$, $u(x) = \int \mu(x) = \int bx^{-\eta}dx$, or

$$\begin{aligned} u(x) &= a + b\ln(x) && \text{if } \eta = 1; \\ &= a + \frac{b}{1-\eta}x^{1-\eta} && \text{if } \eta \neq 1. \end{aligned}$$

Linear non-homogeneous first-order differential equations with constant coefficients

- These differential equations cane be written as follows

$$\frac{dy}{dt} = c_0 + c_1 y.$$

They are autonomous.

- To solve them we follow in two steps.

- *Step 1*: start with a guess for a particular solution: suppose $y(t) = k$, constant. We can write,

$$0 = c_0 + c_1 k \qquad \Rightarrow \qquad k = -\frac{c_0}{c_1}, \text{ provided } c_1 \neq 0,$$

confirming that for $k = -c_0/c_1$, our guess is indeed a solution.

- *Step 2*: now transform the variable y to $z = y - k$ and notice that $dz/dt = dy/dt$, because k is constant across t by hypothesis. We can then write,

$$\begin{aligned} \frac{dz}{dt} &= c_0 + c_1 y \\ &= c_0 + c_1 k + c_1 z \\ &= c_1 z \end{aligned}$$

But we know the solution to this homogeneous equation: it is $z(t) = ae^{c_1 t}$. But since $y = z + k = z - c_0/c_1$, the solution in terms of y is,

$$y(t) = ae^{c_1 t} - \frac{c_0}{c_1}$$

where a can be computed using some initial condition.

- This means that the solution to the linear first order differential equation with constant coefficients is the sum of a *particular solution,* $-c_0/c_1$, and the *general solution,* $ae^{c_1 t}$.

- If c_1 is negative, y converges to the particular solution, $-c_0/c_1$. This is the *steady state* value when the argument of function y is time.

- To verify the solution, compare the derivative with respect to time of the solution,

$$y' = ac_1 \exp(c_1 t),$$

with the differential equation, evaluated at the solution,

$$y' = c_0 + c_1 \left(a \exp(c_1 t) - \frac{c_0}{c_1} \right) = ac_1 \exp(c_1 t).$$

As they agree, this is indeed a solution.

General linear non-homogeneous first-order differential equations

- These equations are of the following type:

$$y' + u(t)y = v(t) \tag{26}$$

These are non-autonomous.

- To solve them we first try to simplify them. A method of doing so is to multiply the whole equation by an integrating factor, $I(t)$. This will not change the solution but make the equation simpler to solve.
- Suppose we multiply the differential equation (26) by the an integrating factor $I(t)$ that satisfies the following condition:

$$I(t)u(t) = I'(t) \tag{27}$$

to obtain

$$I(t)y' + I(t)u(t)y = I(t)v(t) \tag{28}$$

- Now notice that, by the product differentiation rule;

$$I(t)y' + I(t)u(t)y = \frac{d}{dt}(I(t)y)$$

and thus the transformed differential equation (28) can be written as:

$$\frac{d}{dt}(I(t)y) = I(t)v(t)$$

- we can now integrate both sides with respect to t to obtain:

$$I(t)y = \int I(t)v(t)dt$$

which yields

$$y = \frac{1}{I(t)}\int I(t)v(t)dt$$

- The main problem, of course, is how to find the appropriate integrating factor.
  - We know that $I(t)$ should fulfil condition (27).
  - But then:

$$\frac{I'(t)}{I(t)} = u(t)$$

which can be re-written as

$$\frac{d}{dt}\ln I(t) = u(t)$$

  - Integrating yields:

$$\ln I(t) = U(t)$$

or

$$I(t) = \exp\{U(t)\} \tag{29}$$

where U is the primitive of u.

  - Notice that we use the primitive, excluding the constant from the integration. This does not change the validity of the integrating factor and we, therefore, choose its simplest version.

- Using the integrating factor in (29) the solution can now be written as:

$$y = \exp\{-U(t)\} \int \exp\{U(t)\} v(t) dt$$

- Notice that the integral above is indefinite, not the primitive, and so there will be an integrating constant when applied to concrete cases.

7.3 Higher order differential equations and systems of differential equations

- Consider the following nth-order differential equation,

$$x^{(n)} + a_1 x^{(n-1)} + ... + a_{n-1} x' + a_n x = c$$

  where $a_1, ..., a_n, c$ are constants.

- This equation can be represented by a system of first order linear equations by defining,

$$\begin{aligned} y_0 &= x \\ y_i &= x^{(i)} \quad \text{for } i = 1, ...n-1 \end{aligned}$$

  and thus, $y_i' = x^{(i+1)} = y_{i+1}$.

- We can now write the equivalent system of n differential equations in n unknowns as,

$$\begin{cases} y_i' = y_{i+1} & \text{for } 0 \leqslant i \leqslant n-2 \\ y_{n-1}' = -a_1 y_{n-1} - a_2 y_{n-2} - ... - a_n y_0 + c \end{cases}$$

- This is just a particular case of a general system of first order differential equations,

$$y_i' = \sum_{j=1}^{n} a_{ij} y_j + c_i \quad \text{for } i = 1, ..., n$$

  or, in vector notation,

$$y' = Ay + c$$

  where y, y' and c are $n \times 1$ and A is $n \times n$.

7.3.1 Preliminaries

- Consider the homogenous system $y' = Ay$ of n equations.
- From the study of univariate differential equations, we might guess that $y = \alpha \exp(\lambda t)$ is a solution, for some $(n \times 1)$ vector α and a scalar λ.
- To verify the guess, note that $y' = \alpha\lambda \exp(\lambda t)$, and evaluating the system of differential equations at the guess yields therefore

$$y' = \alpha\lambda \exp(\lambda t) = A\alpha \exp(\lambda t).$$

  This, in turn, implies

$$(A - \lambda I_n)\alpha \exp(\lambda t) = 0.$$

  Since $\exp(\lambda t) > 0$ for any λ and t, it follows that validity of the guess requires that

$$(A - \lambda I_n)\alpha = 0,$$

  i.e. that λ be an eigenvalue of A, and α be the associated eigenvector. So, in order to find solutions to the system $y' = Ay$, one needs to find the eigenvectors and eigenvalues of A.

- Two facts:

  (i) If y_1 and y_2 solve $y' = Ay$, then any convex combination of them also forms a solution.

  (ii) Suppose $Y = [y_1, \cdots, y_n]$ are solutions to the system of n differential equations $y' = Ay$. Then, if $\det(Y) \neq 0$, the collection $\{y_1, \cdots, y_n\}$ is said to form a fundamental set of solutions, and the general solution is

$$y = \beta_1 y_1 + \cdots + \beta_n y_n,$$

  for some coefficients $\beta_1, \cdots, \beta_n$.

7.3.2 Solving systems of linear differential equations with constant coefficients when the eigenvalues of A are all distinct

- Consider the system of differential equations written in matrix format,

$$y' = Ay + c$$

- Assume A is full-rank (invertible) and has n distinct eigenvalues.
- Let $\lambda_1, \ldots, \lambda_n$ be the n distinct eigenvalues of A.

- Then we know there exists an orthogonal matrix P such that $P^{-1}AP = D$ where $D = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$. (Recall: Schur decomposition of A: $A = PDP^{-1}$, leading to the diagonalization of A as $D = P^{-1}AP$; here, P is the matrix of eigenvectors of A.)

- Now define $w = P^{-1}y'$. Then $Pw = y'$ and,

$$\begin{aligned} w' &= P^{-1}y'' \\ &= P^{-1}(Ay + c)' \\ &= P^{-1}Ay' \\ &= P^{-1}APw = Dw \end{aligned}$$

- But this is a system of independent homogeneous first order linear equations, each equation being $w_i' = \lambda_i w_i$ for $i = 1, \ldots, n$.

- The solution to each of these is

$$w_i = k_i e^{\lambda_i t}$$

  or, in matrix notation,

$$w = Ke$$

  where $K = \text{diag}\{k_1, \ldots, k_n\}$ and $e = \left[e^{\lambda_1 t} \ldots e^{\lambda_n t}\right]'$.

- Thus, the solution for y can now be derived,

$$\begin{aligned} y &= A^{-1}(y' - c) \\ &= A^{-1}(Pw - c) \\ &= A^{-1}PKe - A^{-1}c \end{aligned}$$

- Note that $\{\lambda_1, \cdots, \lambda_n\}$ distinct (and non-zero) implies that, for any $t > 0$, the system $w' = Dw$ has a fundamental set of s solutions (and hence a general solution), which in turn translates into a fundamental set of solutions (and a general solution) of the system $y' = Ay + c$.

7.3.3 Solving systems of linear differential equations with constant coefficients in the general case with possibly non-distinct eigenvalues of A

- We will consider the system of two differential equations,

$$\begin{cases} y_1' = a_{11}y_1 + a_{12}y_2 + c_1 \\ y_2' = a_{21}y_1 + a_{22}y_2 + c_2 \end{cases}$$

- We will follow in steps:
  - *Step 1*: Guess some particular solution and check it is indeed a solution.
    - We choose $y_1 = k_1$ and $y_2 = k_2$.
    - Let $k = [k_1, k_2]'$, and let A be the coefficient matrix.
    - With this solution, the system becomes $Ak = -c$ which implies $K = -A^{-1}c$, provided A has full rank, i.e. $a_{11}a_{22} - a_{12}a_{21} \neq 0$; this is a particular solution to the system of equations.
  - *Step 2*: Transform the variable to obtain a system of homogeneous equations and solve it.
    - Use the change of variable $z = y - k$ which then implies,

$$\begin{aligned} y &= z - A^{-1}c \\ \text{and} \quad z' &= y' \end{aligned}$$

    - Replacing in the system yields,

$$\begin{aligned} z' &= Ay + c \\ &= A(z - A^{-1}c) + c \\ &= Az \end{aligned}$$

    - We guess a solution,

$$\begin{aligned} z_1 &= \alpha_1 e^{\lambda t} \\ z_2 &= \alpha_2 e^{\lambda t} \end{aligned}$$

      which yields,

$$\begin{aligned} z_1' &= \alpha_1 \lambda e^{\lambda t} \\ z_2' &= \alpha_2 \lambda e^{\lambda t} \end{aligned}$$

    - Replacing in the system $z' = Az$ yields,

$$\alpha \lambda e^{\lambda t} = A\alpha e^{\lambda t}$$

      where $\alpha = (\alpha_1, \alpha_2)'$ and $e^{\lambda t}$ is a positive scalar.
    - This results in the system,

$$(A - \lambda I)\alpha = 0$$

      which has a non-zero solution only if $|A - \lambda I| \neq 0$ and thus λ_1, λ_2 are the eigenvalues of A (possibly non-different) and α are the eigenvectors.

∗ Solving for the eigenvalues and computing the eigenvectors we obtain,

$$\begin{cases} z_1 = \alpha_{11} e^{\lambda_1 t} + \alpha_{12} e^{\lambda_2 t} \\ z_2 = \alpha_{21} e^{\lambda_1 t} + \alpha_{22} e^{\lambda_2 t} \end{cases}$$

– *Step 3*: Since $y = z + k$, the solution can now be constructed,

$$y = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} \\ e^{\lambda_2 t} \end{bmatrix} - A^{-1} c$$

– The solution can be verified by comparing (i) y' derived from it with (ii) the result from substituting the solution into the initial system of differential equations.

$$\begin{aligned} (i)\ y' &= \begin{bmatrix} \alpha_{11}\lambda_1 e^{\lambda_1 t} & \alpha_{12}\lambda_2 e^{\lambda_2 t} \\ \alpha_{21}\lambda_2 e^{\lambda_1 t} & \alpha_{22}\lambda_2 e^{\lambda_2 t} \end{bmatrix} \\ &= [\lambda_1\alpha_1, \lambda_2\alpha_2] \begin{bmatrix} e^{\lambda_1 t} \\ e^{\lambda_2 t} \end{bmatrix} \\ (ii)\ y' &= Ay + c \\ &= A \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} \\ e^{\lambda_2 t} \end{bmatrix} \\ &= A [\alpha_1, \alpha_2] \begin{bmatrix} e^{\lambda_1 t} \\ e^{\lambda_2 t} \end{bmatrix}, \end{aligned}$$

and the two expressions agree because $A\alpha_i = \lambda_i \alpha_i$, $i = 1, 2$.

7.4 Applied Example: Optimal Auctions

References: Riley and Samuelson (1981)

A seller faces n potential buyers.

Buyer i has reservation value v_i, $i = 1, \cdots, n$.

Suppose the reservation values v_i are independently and identically distributed (so-called independent private values). Specifically, suppose that $v_i \overset{i.i.d.}{\sim} u[0, \theta]$, $\theta > 0$.

There are various auctions formats:

- English (ascending) auction;
- Dutch (descending) auction;
- first prize sealed bid auction;
- second prize sealed bid auction.

In equilibrium, buyer i bids $b_i = b(v_i)$, $i = 1, \cdots, n$, increasing in v_i; i.e. the equilibrium bidding function $b(\cdot)$ monotonically maps reservation values into bids.

A bidder's expected gain is his/her reservation value, multiplied by the probability of winning the auction, minus the expected auction payment.

Let $\psi(b_1, \cdots, b_n)$ be the function that maps bids into auction payments.

W.l.o.g., consider buyer 1. If buyer 1 submits bid b, then, conditional on winning, he/she faces expected payment

$$P(b) = \mathbb{E}[\psi(b, b_2, \cdots, b_n) | b > b_i, \forall i > 1],$$

where the expectation is taken with respect to $b_2, \cdots, b_n$. This is complicated...

The range of the equilibrium bidding function $b(\cdot)$ is $\mathcal{B} = [0, b(\theta)]$, and hence buyer 1 will submit a bid $b_1 \in \mathcal{B}$. Note that $b_1 = b(x)$ for some $x \in [0, v_1]$, and since $b(\cdot)$ is monotonic, $x = b^{-1}(b_1)$. So determining the optimal bid is equivalent to determining an optimal value x. Therefore, $b(\cdot)$ represents equilibrium bidding strategies iff buyer 1's optimal bid $b_1^\star = b(v_1) = b(x^\star)$, i.e. $x^\star = v_1$ (truthful bidding; auctions are so-called revelation mechanisms).

Then, buyer 1's expected gain from the auction, given his/her reservation value and bid (in terms of x) and assuming all other bids follow equilibrium strategies, is

$$\pi(x, v_1) = v_1 \left(\frac{x}{\theta}\right)^{n-1} - P(x),$$

which buyer 1 maximizes w.r.t. x.

Participation constraint: Buyer 1 will only submit a bid if $v_1 > v^\star$, where $v^\star$ satisfies $\pi(v^\star, v^\star) = 0$, i.e. $P(v^\star) = v^\star \left(\frac{v^\star}{\theta}\right)^{n-1}$.

Maximization w.r.t. x of $\pi(x, v_1)$ yields

$$v_1(n-1)\frac{(x^\star)^{n-2}}{\theta^{n-1}} - P'(x^\star) = 0.$$

This must hold in equilibrium, i.e. when $x^\star = v_1$, so that

$$P'(v_1) = (n-1)\left(\frac{v_1}{\theta}\right)^{n-1}, \text{ for any } v_1 \in [v^\star, \theta].$$

This is a differential equation characterizing the expected payment function.

Integrating,

$$\int_{v^\star}^{v_1} (n-1)\left(\frac{x}{\theta}\right)^{n-1} dx = P(v_1) - P(v^\star),$$

which, combined with the boundary condition (participation constraint) implies

$$\begin{aligned} P(v_1) &= \frac{(v^\star)^n}{\theta^{n-1}} + \frac{n-1}{n}\frac{1}{\theta^{n-1}}\left[v_1^n - v^{\star n}\right] \\ &= v_1\left(\frac{v_1}{\theta}\right)^{n-1} - \frac{1}{n}\frac{1}{\theta^{n-1}}\left[v_1^n - v^{\star n}\right] \text{ for all } v_1 > v^\star. \end{aligned}$$

The first part if the buyer 1's expected gross surplus from winning. The more competitive the auction, i.e. the more bidders, the more of this gross surplus the winning bidder will have to hand over to the seller.

Note that the preceding result for the expected payment function $P(v)$ could be derived in ignorance about the specific auction format, i.e. it holds for any auction format (so-called Revenue Equivalence).

The expected payment $P(v)$ equals the bid $b(v)$ multiplied by the probability of winning, q. Note that monotonicity of $b(\cdot)$ implies

$$q = \Pr(b(v_1) > b(v_i), \forall i > 1) = \Pr(v_1 > v_i, \forall i > 1) = \left(\frac{v_1}{\theta}\right)^{n-1}.$$

Hence,

$$b(v) = P(v)/q = v - \frac{1}{n}\frac{v_1^n - v^{\star n}}{v^{n-1}}.$$

Note that the bidder's reservation value v exceeds his/her optimal bid $b(v)$ (so-called bid shading).

8 Probability Theory

References: Bickel and Doksum (1977), Appendix A; Billingsley (1995), Durrett (1996).

8.1 Probability Space

Recall from the Introduction to Measure Theory: A probability space is a measurable space $(\Omega, \mathcal{A})$ with probability measure P, where Ω is the *sample space*, i.e. the set of all possible outcomes of a random experiment, whose subsets are called *events*, and $\mathcal{A}$ is a σ-algebra on Ω.

Next to the property that probability measure is normalized $P(\Omega) = 1$, it satisfies all the other properties of measures, as defined before.

8.2 Discrete Probability Models

A probability model is *discrete* if Ω is finite or countably infinite, so $\Omega = \{\omega_1, \omega_2, \cdots\}$, and every element is assigned a probability. Then, for any event $A \in \mathcal{A}$,

$$P(A) = \sum_i 1_{\{\omega_i \in A\}} P(\{\omega_i\}).$$

8.3 Conditional Probability and Independence

- Given an event $B \in \mathcal{A}$ with $P(B) > 0$ and any other event $A \in \mathcal{A}$, the conditional probability of A, given B, is defined as
$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$
It characterizes the chance assigned to the occurrence of A, given that states of the world occurred that led to the occurrence of B.

- If $A_1, A_2, \cdots$ ... are pairwise disjoint events, and $P(B) > 0$, then
$$P\left(\cup_{i=1}^{\infty} A_i |\, B\right) = \sum_{i=1}^{\infty} P(A_i|B).$$
In fact, for given B, $P(\cdot|B)$ itself is a probability measure on $(\Omega, \mathcal{A})$, called the *conditional probability measure* given B.

- It is important to recognize, however, that for a given event A the conditional probability $P(A|\cdot)$ is random because the conditioning event itself is random.

- If $B_1, \cdots, B_n$ are pairwise disjoint events with $P(B_i) > 0$ for $i = 1, \cdots, n$ and $\Omega = \cup_{i=1}^{n} B_i$, then
$$A = \cup_{i=1}^{n} (A \cap B_i)$$
implies
$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A|B_i) P(B_i),$$

because $P(A \cap B_i) = P(A|B_i)P(B_i)$ by the definition of conditional probability, $i = 1, \cdots, n$.

- If $P(A) > 0$, combining the above expressions yields *Bayes rule*,

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}.$$

- The events A and B are said to be *independent* if

$$P(A \cap B) = P(A)P(B).$$

If $P(B) > 0$ and A and B are independent, then

$$P(A|B) = P(A),$$

i.e. knowledge of B does not restrict the state of the world that induce A and hence does not affect the probability of A.

8.4 Probabilities on Euclidean Space

- In applied work, random experiments typically yield outcomes that are real numbers. Their sample space can be thought of as Euclidean space, say of dimension k, $\mathbb{R}^k$, and the associated (Borel) σ-algebra $\mathcal{B}^k$ is defined on all open k-dimensional rectangles, i.e. sets of the form $(a_1, b_1) \times \cdots \times (a_k, b_k) = \{(x_1, \cdots, x_k) : a_i < x_i < b_i, i = 1, \cdots, k\}$.

- A discrete probability distribution on $\mathbb{R}^k$ is a probability measure P such that $\sum_{i=1}^{m} P(\{x_i\}) = 1$ for some points $x_i \in \mathbb{R}^k$ and some $m \in \mathbb{N}$. The elements in the sum are called *probability mass function* (pmf) or *frequency function*.

- An integrable, non-negative function p on $\mathbb{R}^k$ with the property

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(t_1, \cdots, t_k) dt_1 \cdots dt_k = 1$$

is called *probability density function* (pdf).

- An (absolutely) *continuous probability distribution* on $\mathbb{R}^k$ is a probability P defined by

$$P(A) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} 1_{\{(t_1, \cdots, t_k) \in A\}} p(t_1, \cdots, t_k) dt_1 \cdots dt_k.$$

A special case is the probability of a k-dimensional rectangle,

$$P\left((a_1, b_1) \times \cdots \times (a_k, b_k)\right) = \int_{a_1}^{b_1} \cdots \int_{a_k}^{b_k} p(t_1, \cdots, t_k) dt_1 \cdots dt_k.$$

- A *cumulative distribution function* (CDF) is defined by

$$F(x_1, \cdots, x_k) = P\left((-\infty, x_1] \times \cdots \times (-\infty, x_k]\right), \quad (x_1, \cdots, x_k) \in \mathbb{R}^k.$$

When $k = 1$, F is a function defined on $\mathbb{R}$ with the following properties:

(i) $0 \leq F \leq 1$;

(ii) $x \leq y$ implies $F(x) \leq F(y)$ (monotonicity);

(iii) $x_n \searrow x$ implies $F(x_n) \searrow F(x)$ (continuity from the right);

(iv) $\lim_{x\to\infty} F(x) = 1$, and $\lim_{x\to-\infty} F(x) = 0$.

8.5 Random Variables and Vectors: Transformations

- A real-valued random vector X is a function from Ω to $\mathbb{R}^k$ such that $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}$ for every $B \in \mathcal{B}^k$.

- The probability distribution of X, denoted by P_X, therefore derives from the probability measure P,

$$P_X(B) = P\left(\{\omega : X(\omega) \in B\}\right).$$

A random vector is said to have a continuous or discrete distribution according to whether its probability distribution is continuous or discrete; similarly for its frequency function, density function etc.

- Let g be any measurable function from $\mathbb{R}^k$ to $\mathbb{R}^m$, i.e. $g^{-1}(B) = \{y \in \mathbb{R}^k : g(y) \in B\} \in \mathcal{B}^k$ for every $B \in \mathcal{B}^m$. Then, the random variable $g(X)$ is defined by $g(X)(\omega) = g(X(\omega))$.

- The probability distribution of $g(X)$, $P_{g(X)}$, is determined by P_X (which, in turn, is determined by P):

$$P_{g(X)}(B) = P_X\left(X \in g^{-1}(B)\right) = P\left(\{\omega : X(\omega) \in g^{-1}(B)\}\right).$$

- Suppose that X is a continuously distributed scalar random variable with density p_X and support S, and g is real-valued, one-to-one, and g' exists and is non-zero on its support. Then, $g(X)$ has density

$$p_{g(X)}(t) = \frac{p_X(g^{-1}(t))}{|g'(g^{-1}(t))|}, \quad t \in g(S).$$

This is the *change of variables formula*.

- It generalizes to vector-valued, continuously distributed random variables. If X is a continuously distributed random vector in $\mathbb{R}^k$ with density p_X and support $S \subseteq \mathbb{R}^k$, and $g : S \to g(S) \subseteq \mathbb{R}^k$ is real-valued, one-to-one, and ∇g exists and is non-singular on its support. Then, $g(X)$ has density

$$p_{g(X)}(t) = p_X(g^{-1}(t))|\nabla g(g^{-1}(t))|^{-1}, \quad t \in g(S).$$

- Example: Let $g(X) = \mu + \sigma X$, for μ and $\sigma \neq 0$ two non-stochastic parameters. Then,

$$\begin{aligned} g^{-1}(t) &= \frac{t-\mu}{\sigma} \\ g'(X) &= \sigma \quad \text{for any } X \\ |g'(g^{-1}(t))|^{-1} &= \sigma^{-1} \\ \Rightarrow \quad p_{g(X)}(t) &= \frac{1}{\sigma} p_X\left(\frac{t-\mu}{\sigma}\right). \end{aligned}$$

- If (X, Y) is continuously distributed with (joint) density $p_{X,Y}$, then X has *marginal density*

$$p_X(x) = \int_{-\infty}^{\infty} p_{(X,Y)}(x, y)dy,$$

and, if $p_X(x) > 0$, then Y has conditional density

$$p_{Y|X}(y|X = x) = \frac{p_{(X,Y)}(x, y)}{p_X(x)}.$$

8.6 Expectations

- If X is a discrete random variable with values $\{x_1, x_2, \cdots\}$, the *expectation* (or *mean*) of X is defined by

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i p_X(x_i),$$

provided this sum is finite (otherwise, the expectation of X does not exist). Finiteness is ensured if $\mathbb{E}(|X|) = \sum_{i=1}^{\infty} |x_i| p_X(x_i) < \infty$.

- More generally, if X is an n-dimensional discrete random vector and g is a real-valued function on $\mathbb{R}^n$,

$$\mathbb{E}(g(X)) = \sum_{i=1}^{\infty} g(x_i) p_X(x_i),$$

provided this sum is finite. Finiteness is ensured if $\mathbb{E}(|g(X)|) = \sum_{i=1}^{\infty} |g(x_i)| p_X(x_i) < \infty$.

- This generalizes further to functions of continuously distributed random vectors, where sums are replaced by integrals:

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n) p_X(x_1, \cdots, x_n) dx_1 \cdots x_n,$$

provided the integral is finite. Finiteness is ensured if

$$\mathbb{E}(|g(X)|) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |g(x_1, \cdots, x_n)| p_X(x_1, \cdots, x_n) dx_1 \cdots x_n < \infty.$$

- The expectation operator is a linear operator, so it commutes with other linear operator; e.g. for non-stochastic parameters μ and α,

$$\mathbb{E}(\mu + \alpha X) = \mu + \alpha \mathbb{E}(X).$$

8.7 Moments

- If $k \in \mathbb{N}$ and X is a scalar random variable, the kth *moment* of X is defined as the expectation of X^k,

$$\begin{aligned}\mathbb{E}(X^k) &= \sum_{i=1}^{\infty} x_i^k p_X(x_i) \quad \text{if } X \text{ is discrete} \\ &= \int_{-\infty}^{\infty} x^k p_X(x) dx \quad \text{if } X \text{ is continuous.}\end{aligned}$$

- The kth *central moment* of X is $\mathbb{E}\left[(X - \mathbb{E}(X))^k\right]$. The second central moment is the variance of X, $\text{var}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]$, and its non-negative square root is called the standard deviation of X.

- For $X \in \mathbb{R}^k$, $\mathbb{E}(X) \in \mathbb{R}^k$, and for $Y \in \mathbb{R}^m$,

$$\text{cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))'\right],$$

  a $k \times m$ matrix. Hence,

$$\text{var}(X) = \text{cov}(X, X) = \mathbb{E}\left[(X - \mathbb{E}(X))(X - \mathbb{E}(X))'\right],$$

  a positive, semi-definite and symmetric $k \times k$ matrix.

- If X and Y are independent,

$$\begin{aligned}\text{cov}(X, Y) &= \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))'\right] \\ &= \mathbb{E}\left[X - \mathbb{E}(X)\right]\mathbb{E}\left[Y - \mathbb{E}(Y)'\right] \\ &= 0\end{aligned}$$

  so independence implies zero covariance. The converse is not generally true.

- If A and B are non-stochastic $l \times k$ and $n \times m$ matrices, and a and b are non-stochastic $l \times 1$ and $n \times 1$ vectors, then

$$\text{cov}(AX + a, BY + b) = A\text{cov}(X, Y)B',$$

  so that $AX + a$ is seen to have variance $A\text{var}(X)A'$.

- When the expectation of a random variable is taken with respect to a conditional distribution, it is called a *conditional expectation*. If $F(x|Y = y)$ is the conditional distribution of X, given $Y = y$, then for any real function $g(X)$,

$$\mathbb{E}[g(X)|Y = y] = \int_{-\infty}^{\infty} g(x)F(dx|Y = y),$$

assuming the integral is finite. For any real function $h(x, y)$, by Fubini's Theorem,

$$\begin{aligned}\mathbb{E}[h(X,Y)] &= \int h(x,y)F(dx,dy) \\ &= \int_y \left[\int_x h(x,y)F(dx|Y=y)\right] F(dy) \\ &= \mathbb{E}_Y\mathbb{E}_{X|Y}h(X,Y) \\ &= \mathbb{E}_X\mathbb{E}_{Y|X}h(X,Y),\end{aligned}$$

the *Law of Iterated Expectations*. This re-enforces the fact that conditional distributions and conditional moments are random.

- The conditional variance of scalar Y, given scalar X, is $\text{var}(Y|X) = \mathbb{E}_{Y|X}\left[(Y - \mathbb{E}[Y|X])^2\right]$; analogous formulae apply in the case of vectors. It is related to the unconditional variance of Y by the formula

$$\begin{aligned}\text{var}(Y) &= \mathbb{E}_X\left[\mathbb{E}_{Y|X}\left[\left(Y - \mathbb{E}_{Y|X}[Y|X] + \mathbb{E}_{Y|X}[Y|X] - \mathbb{E}(Y)\right)^2\right]\right] \\ &= \mathbb{E}_X\left[\mathbb{E}_{Y|X}\left[\left(Y - \mathbb{E}_{Y|X}[Y|X]\right)^2 \middle| X\right]\right] + \mathbb{E}_X\left[\left(\mathbb{E}_{Y|X}[Y|X] - \mathbb{E}_Y(Y)\right)^2\right] \\ &\quad +2\mathbb{E}_X\left[\mathbb{E}_{Y|X}\left[\left(Y - \mathbb{E}_{Y|X}[Y|X]\right)\left(\mathbb{E}_{Y|X}[Y|X] - \mathbb{E}_Y(Y)\right)\middle| X\right]\right] \\ &= \mathbb{E}_X\left[\text{var}(Y|X)\right] + \mathbb{E}_X\left[\left(\mathbb{E}_{Y|X}[Y|X] - \mathbb{E}(Y)\right)^2\right] \\ &\quad +2\mathbb{E}_X\left[\mathbb{E}_{Y|X}\left[\left(Y - \mathbb{E}_{Y|X}[Y|X]\right)\middle| X\right]\left(\mathbb{E}_{Y|X}[Y|X] - \mathbb{E}_Y(Y)\right)\right] \\ &= \mathbb{E}_X\left[\text{var}(Y|X)\right] + \mathbb{E}_X\left[\left(\mathbb{E}_{Y|X}[Y|X] - \mathbb{E}(Y)\right)^2\right],\end{aligned}$$

i.e. the unconditional variance is the expectation of the conditional variance plus the variance of the conditional mean.

8.8 Limit Theorems in Statistics

8.8.1 Modes of Convergence

- Consider a sequence of random variables $Y_1, Y_2, \cdots, Y_n, \cdots$. These random variables are all functions of the same state of the world, $\omega \in \Omega$. Denote the joint CDF of the finite subsequence $(Y_1, \cdots, Y_n)$ by $F_{Y_1,\cdots,Y_n}(y_1, \cdots, y_n)$, as defined above.
- Preliminary remark: There are several concepts for the limit of a sequence of random variables, $\lim_{n\to\infty} Y_n = Y_0$. Since the Y_n are functions of states of the world, these limit concepts correspond to different ways of defining limits of functions.

1. Y_n *converges in probability* to Y_0, denoted $Y_n \xrightarrow{p} Y_0$ or $\text{plim}_n Y_n = Y_0$, if, for each $\epsilon > 0$, $\lim_{n\to\infty} P\left(|Y_n - Y_0| > \epsilon\right) = 0$.

2. Y_n *converges almost surely* (or *with probability 1*) to Y_0, denoted $Y_n \xrightarrow{a.s.} Y_0$, if, for each $\epsilon > 0$, $\lim_{n\to\infty} P(\sup_{n'\geq n} |Y_{n'} - Y_o| > \epsilon) = 0$.

- To compare convergence in probability and almost surely, let $W_n = \{\omega : |Y_n(\omega) - Y_0| > \epsilon)$ and note that $W_n \subset \bigcup_{n'\geq n} W_{n'}$. Convergence in probability is equivalent to $\lim_n P(W_n) = 0$, while almost sure convergence is equivalent to $\lim_n P\left(\bigcup_{n'\geq n} W_{n'}\right) = 0$. Hence, the latter implies the former and therefore constitutes a stronger (more demanding) notion of convergence.

3. Y_n *converges in* $\mathcal{L}^p$ to Y_0 if $\lim_n \mathbb{E}\left[|Y_n - Y_o|^p\right] = 0$, $p \in \mathbb{N}$. For $p = 2$, this is called convergence in quadratic mean.

4. Y_n *converges in distribution* to Y_0, denoted by $Y_n \xrightarrow{d} Y_0$, if the CDF of Y_n converges to the CDF of Y_0 at each continuity point of Y_0; i.e. if F_{Y_0} is continuous at y, then $\lim_n F_{Y_n}(y) = f_{Y_0}(y)$.

- Note that, for convergence in distribution (4), the random variables Y_n and Y_0 can be defined on different probability spaces; this is not the case for the other modes of convergence (1-3). Almost sure convergence and convergence in probability imply convergence in distribution, but the converse is not true, unless the limiting distribution has all all probability mass at a single point (i.e. the limiting distribution is degenerate).

- If two random variables Y_n and Z_n satisfy $|Y_n - Z_n| \xrightarrow{p} 0$ and $Y_n \xrightarrow{d} Y_0$, then $Z_n \xrightarrow{d} Y_0$ as well. In this case, Y_n and Z_n are said to be *asymptotically equivalent*.

- $Y_n \xrightarrow{p} Y_0$ implies $g(Y_n) \xrightarrow{p} g(Y_0)$ for any continuous function $g(\cdot)$.

- The preceding result has an analogue for convergence in distribution: $Y_n \xrightarrow{d} Y_0$ implies $g(Y_n) \xrightarrow{d} g(Y_0)$ for any continuous function $g(\cdot)$.

- Like the Landau symbols, there exist a o_p and O_p notation for convergence properties (Mann-Wald symbols).
  - The definition of o_p is $Y_n \xrightarrow{p} Y_0$ is equivalent to $Y_n = Y_0 + o_p(1)$, and $n^{-\alpha}(Y_n - Y_0) \xrightarrow{p} 0$ is equivalent to $Y_n - Y_o = o_p(n^\alpha)$.
  - The notation $Y_n = O_p(1)$ is defined to mean that, given $\epsilon > 0$, there exists a large M (not depending on n) such that $P(|Y_n| > M) < \epsilon$ for all n. A sequence Y_n with this property is called *stochastically bounded*.

8.8.2 Laws of Large Numbers and Central Limit Theorems

- Consider a sequence of random variables $Y_1, Y_2, \cdots, Y_n, \cdots$ and a corresponding sequence $X_n = \frac{1}{n}\sum_{i=1}^n Y_i$. Laws of large numbers (LLNs) are concerned with the conditions under

which X_n converges to a constant, either in probability (weak laws) or almost surely (strong laws).

- There exists a variety of such LLNs. They essentially differ in terms of the assumptions they make on the statistical dependence across the Y_i and the existence of their moments.
- The *Khinchine Weak Law of Large Numbers*: If the Y_n are i.i.d. and $\mathbb{E}(Y_n) = \mu$, then $X_n \xrightarrow{p} \mu$.
- The *Kolmogorov Strong Law of Large Numbers*: If the Y_n are i.i.d. and $\mathbb{E}(Y_n) = \mu$, then $X_n \xrightarrow{a.s.} \mu$. Obviously, this result encompasses the preceding WLLN.
- Assume now that $\mathbb{E}(Y_i) = 0$ for all i. Consider the sequence $Z_n = \sqrt{n}X_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i$. Central Limit Theorems (CLTs) are concerned with conditions under which Z_n converges in distribution to a normal random variable Z_0.

  Aside: A standard normal random V variable has density $\phi(x) = \frac{1}{sqrt2\pi}\exp\left(-\frac{1}{2}x^2\right)$, $x \in \mathbb{R}$. Using the change of variables formula, the random variable $W = \mu + \sigma V$ has density $\frac{1}{\sigma}\phi\left(\frac{W-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(w-\mu)^2\right)$, $w \in \mathbb{R}$. The normal distribution is completely characterized my its mean μ and variance parameters σ^2, and denoted by $N(\mu, \sigma^2)$.

  A multivariate normal random vector $x \in \mathbb{R}^k$ with mean vector $\mu \in \mathbb{R}^k$ and $k \times k$ variance-covariance matrix Σ, positive definite symmetric, has density

$$\frac{1}{\sqrt{(2\pi)^k|\Sigma|}}\exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right).$$

  If the elements of x are have zero covariance, i.e. $\Sigma = \text{diag}\{\sigma_i^2, i = 1, \cdots, k\}$ and therefore $|\Sigma| = \prod_{i=1}^{k}\sigma_i^2$, then this density simplifies to

$$\frac{1}{\sqrt{\prod_{i=1}^{k}[2\pi\sigma_i^2]}}\exp\left(-\frac{1}{2}\sum_{i=1}^{k}\frac{x_i-\mu}{\sigma_i^2}\right) = \prod_{i=1}^{k}\frac{1}{\sigma_i}\phi\left(\frac{x_i-\mu}{\sigma_i}\right),$$

  i.e. the joint density factorizes into the product of the component densities, and therefore, by definition of statistical independence, in the case of the normal distribution zero covariance implies independence.

- There exists a variety of CLTs as well. They also essetially differ in terms of the assumptions they make on the statistical dependence across the Y_i and the existence of their moments.
- The *Lindeberg-Lévy Central Limit Theorem*: If the Y_n are i.i.d. with $\mathbb{E}(Y_n) = 0$ and $\mathbb{E}(Y_n^2) = \sigma^2 < \infty$, then $Z_n \xrightarrow{d} Z_0 \sim N(0, \sigma^2)$.

8.8.3 An Example

- Let Y_i be the outcome of the ith fair coin toss in a sequence of n independent tosses, $i = 1, \cdots, n$; i.e. $Y_i = 1$ if the coin comes up heads, and $Y_i = 0$ if it comes up tails, and as the coin is fair, $P(Y_i = 1) = P(Y_i = 0) = \mathbb{E}[Y_i] = \frac{1}{2}$, for all $i = 1, \cdots, n$. So the probability space associated with a coin toss has sample space $\Omega = \{0, 1\}$, the σ-algebra consists of $\{0, 1, \emptyset, \{0, 1\}\}$, and is endowed with the binomial probabilities $P(Y_i = 1) = P(Y_i = 0) = \frac{1}{2}$.
- Define the average outcome of the n coin tosses, $X_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$, and note that $X_n \in [0, 1]$. The random variable X_n has the unit interval as sample space, with Borel σ-field of open sets in the unit interval.
- Since the Y_i are assumed i.i.d. with mean $\mathbb{E}(Y_i) = \frac{1}{2}$ for all $i = 1, \cdots, n$, the Kolmogorov SLLN implies that $X_n \overset{a.s.}{\to} \frac{1}{2}$.
- Furthermore, since $\text{var}(Y_i) = \mathbb{E}\left[(Y_i - \mathbb{E}(Y_i))^2\right] = \mathbb{E}(Y_i^2) - \left(\mathbb{E}(Y_i)^2\right) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} < \infty$, the Lindeberg-Lévy CLT implies that $Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(Y_i - \frac{1}{2}\right) \overset{d}{\to} N\left(0, \frac{1}{4}\right)$.
- Notice that $Z_n = \sqrt{n}(X_n - \mu)$ multiplies $X_n - \mu$ by $\sqrt{n}$ which increases in n just at the right rate to balance the variance of $X_n - \mu$ which shrinks to zero as n tends to infinity: $\text{var}(X_n - \mu) = \frac{1}{n^2}\sum_{i=1}^{n}\text{var}(Y_i) = \frac{1}{n}\text{var}(Y_1) = \frac{1}{4n}$.