Monotone, Linear, and Convex Functions

Motivation

We saw in the last lecture that continuous functions preserve the *geometric structure* of the sets they associate. In this lecture, we will see how monotone functions preserve the *order* structure of their domains and how linear functions preserve the *algebraic structure* of the spaces they link. We also introduce convex and concave functions which will be particularly useful when we solve optimization problems.

Monotone Functions

A function between ordered sets X and Y is called **monotone** if it respects the order of X and Y. f is **increasing** if it preserves the ordering so that $x_2 \succsim_X x_1 \implies f(x_2) \succsim_Y f(x_1)$ where $\succsim_i$ is the order on set i. A function is **strictly increasing** if the order relation is strict. Decreasing and strictly decreasing functions are defined analogously. A function f is monotone if it is either increasing or decreasing.

In practice we will often be using functions that map from $\mathbb{R}$ into $\mathbb{R}$ where each set is ordered with $\geq$.

Example: The function $\ln(x)$ is a monotone function that we will often encounter. For any $x_2 \geq x_1$ in $\mathbb{R}_+$, $\ln(x_2) \geq \ln(x_1)$ in $\mathbb{R}$. ■

We also use functions that map from $\mathbb{R}^n$ into $\mathbb{R}$. A function that maps into $\mathbb{R}$ is called a **functional**. The utility functions we use in our theories are examples of functionals.

Example: A Cobb-Douglas utility function, $f(x, y) = x^a y^b$ for $a, b > 0$, is monotone on the domain $(\mathbb{R}^2_+, \geq)$. $(x_2, y_2) \geq (x_1, y_1)$ implies $x_2 \geq x_1$ and $y_2 \geq y_1$. Therefore $f(x_2, y_2) \geq f(x_1, y_1)$. ■

Given any functional f on X and a strictly increasing functional $g : \mathbb{R} \to \mathbb{R}$, their composition $g \circ f : X \to \mathbb{R}$ is called a **monotonic transformation**. A monotonic transformation preserves the ordering implied by f.

Example: Frequently it is easier to work with a log transformation of a function. The log of the Cobb-Douglass function from the previous example is $\ln f(x) = a \ln(x) + b \ln(y)$. This is a monotonic transformation. ■

Note that monotonicity restricts the behavior of a function on comparable elements. It places no restrictions on the action of the function with respect to non-comparable elements.

Linear Functions

A function $f : X \to Y$ between two linear spaces is **linear** if it preserves the linearity of the sets X and Y. That is, for all $x_1, x_2 \in X$ and $\alpha \in \mathbb{R}$, we have

additivity: $f(x_1 + x_2) = f(x_1) + f(x_2)$

homogeneity: $f(\alpha x_1) = \alpha f(x_1)$

Example: The function $f : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $f(x_1, x_2, x_3) = (x_1, x_2, 0)$ is a linear function. Additivity: $f(x_1+y_1, x_2+y_2, x_3+y_3) = (x_1+y_1, x_2+y_2, 0) = (x_1, x_2, 0)+(y_1, y_2, 0)$. Homogeneity: $f(\alpha x_1, \alpha x_2, \alpha x_3) = (\alpha x_1, \alpha x_2, 0) = \alpha(x_1, x_2, 0)$. ■

Example: The function $\ln(x)$ is not a linear mapping. $\ln(1) = 0$, $\ln(2) \approx .7$, $\ln(3) \approx 1.1$. ■

Example: Consider the function $f(x) = 2x + 3$. Most of us grew up referring to this as a linear function. It turns out that f is not in fact linear. $f(x + y) = 2(x + y) + 3 = 2x + 2y + 3 \neq 2x + 2y + 6 = f(x) + f(y)$. $f(\alpha x) = 2\alpha x + 3 \neq 2\alpha x + 3\alpha = \alpha f(x)$. Rather, f is an example of an **affine function** which relate to linear functions in the same way as subspaces relate to affine sets. Affine functions preserve affine sets (lines and planes). ■

Any $m \times n$ matrix $A = (a_{ij})$ defines a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by

$$f(x) = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \cdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix}$$

In the quantitative sequence, the linearity of the expectation operator will become very familiar to you. This is a particular case of a generally desirable property of linear functions. By removing scalars from the arguments of functions and evaluating additive components of the arguments separately, complicated expressions can be made much more tractable.

Properties of Linear Functions

(1) Every linear function $f : X \to Y$ maps the zero vector in X into the zero vector in Y.

(2) Every composition of linear functions is linear.

(3) A linear function (such as a matrix) that has an inverse $f^{-1} : Y \to X$ is said to be nonsingular. The inverse of a nonsingular linear function is also linear.

Bilinear Functions and Inner Product Spaces

A function $f : X \times Y \to Z$ between linear spaces X, Y, and Z is **bilinear** if it is linear in each factor separately:

$f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y)$

$f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2)$

$f(\alpha x, y) = \alpha f(x, y) = f(x, \alpha y)$ for all $\alpha \in \mathbb{R}$

Bilinear functions are one of the most common types of nonlinear functions and are often used to represent objective functions. They are also encountered in optimization since the second derivative of any smooth function is bilinear. We typically encounter bilinear functionals.

Example: The functional $f(x, y) = xy$ is bilinear since $f(x_1 + x_2, y) = (x_1 + x_2)y = x_1y + x_2y = f(x_1, y) + f(x_2, y)$ and $f(\alpha x, y) = (\alpha x)y = \alpha xy = \alpha f(x, y)$. ■

Example: Any $m \times n$ matrix $A = (a_{ij})$ defines a bilinear functional on $\mathbb{R}^m \times \mathbb{R}^n$ by

$$f(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j$$

In the quantitative sequence you will often have an $n \times n$ weighting matrix and use the functional $f(x, x)$. ■

A bilinear functional f on $X \times X$ is called

symmetric if $f(x, y) = f(y, x)$ for all $x, y \in X$

positive semidefinite if $f(x, x) \geq 0$ for every $x \in X$

positive definite if $f(x, x) > 0$ for every $x \in X$

A symmetric, positive definite bilinear functional on a linear space X is called an **inner product**, denoted $x^T y$. A linear space equipped with an inner product is called an **inner product space**. Every inner product defines a norm given by $||x|| = \sqrt{x^T x}$. Therefore every inner product space is a normed linear space. We spend most of our time in the Euclidean space $\mathbb{R}^n$ with the inner product $x^T y = \sum_{i=1}^{n} x_i y_i$. An inner product space mimics the geometry of Euclidean space and is the most structured of linear spaces. We will revisit inner product spaces in the next lecture and see how useful these spaces are for analyzing data. While all inner product spaces are normed linear spaces, not all normed linear spaces are inner product spaces.

In any inner product space, the **Cauchy-Schwarz inequality** holds:

$$|x^T y| \leq ||x|| \; ||y||$$

Several notations are used to express inner product. Two common alternatives to $x^T y$

are $x \cdot y$ and $\langle x, y \rangle$.

Linear Operators

A **linear operator** is a linear function from a set to itself. Every linear operator on a finite-dimensional space can be represented by a square matrix. We often care about fixed points in political science theory. It turns out that every linear operator $f : X \to X$ has at least one fixed point.

The set of all linear operators on a given space X is denoted $L(X, X)$. If X is finite-dimensional, there is a unique functional det called the **determinant** on $L(X, X)$ with the following properties:

$\det(f \circ g) = \det(f) \det(g)$

$\det(I) = 1$ where I is the identity matrix

$\det(f) = 0$ if and only if f is singular

for all $g, f \in L(X, X)$.

The last property will be particularly useful when we want to check whether an operator is nonsingular. Note that det itself is not a linear functional. You should already be familiar with calculating determinants in low dimensions, typically when inverting matrices.

Another application of linear operators that we will use frequently in optimization are quadratic forms. Let X be a Euclidean space. A functional $Q : X \to \mathbb{R}$ is called a **quadratic form** if there exists a symmetric linear operator $f : X \to X$ such that $Q(x) = x^T f(x)$ for every $x \in X$. Quadratic forms are among the simplest nonlinear functionals we encounter.

Example: The nonlinear function $Q(x, y) = x^2 + 4xy + y^2$ is a quadratic form on $\mathbb{R}^2$. The matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

defines a symmetric linear operator on $\mathbb{R}^2$:

$$f(x,y) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+2y \\ 2x+y \end{bmatrix}$$

$$[x,y] \begin{bmatrix} x+2y \\ 2x+y \end{bmatrix} = x^2 + 4xy + y^2 = Q(x,y) \quad \blacksquare$$

Any $n \times n$ symmetric matrix, A, defines a quadratic form $Q(x) = x^T Ax$. A symmetric matrix A is called

positive definite if $x^T Ax > 0$ for all $x \neq 0 \in X$

positive semidefinite if $x^T Ax \geq 0$ for all $x \neq 0 \in X$

negative definite if $x^T Ax < 0$ for all $x \neq 0 \in X$

negative semidefinite if $x^T Ax \leq 0$ for all $x \neq 0 \in X$

All definite matrices are nonsingular. Recall that if an interior solution to an optimization problem is a local maximum, the Hessian matrix with respect to choice variables is negative definite. For two decision variables, this implies that the second derivative with respect to each choice variable is negative. That is, the main diagonal of H is negative.

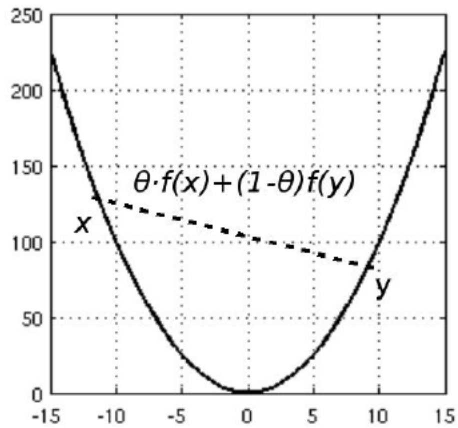
Convex Functions

For many purposes in political science, linearity is too restrictive of an assumption. Convex functions generalize some of the properties of linear functions while providing more suitable functional forms.

A real-valued function f defined on a convex set of a linear space X is **convex** if for every $x_1, x_2 \in S$,

$$f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$$

A function is **strictly convex** if the inequality is strict and **concave** if the inequality is reversed.



The following facts will prove useful:

(1) if f is an invertible function, then f is concave if and only if f^{-1} is convex.

(2) if f and g are convex, $f+g$ is convex

(3) if f is convex, then αf is convex for every $\alpha \geq 0$

(4) logarithmic transformations preserve concavity since ln is concave and increasing

For a twice differentiable function on an open interval $S \subseteq \mathbb{R}$, $f : S \to \mathbb{R}$ is convex if and only if $f''(x) \geq 0$ for every $x \in S$ and concave if and only if $f''(x) \leq 0$ for every $x \in S$. Strict inequality implies strict convexity or concavity.

In the next lecture, we will see how to differentiate functions defined on higher dimensional domains. To preview this, the generalization of the second derivative on a one dimensional domain is given by a symmetric square matrix that is referred to as a **Hessian**, H_f. In higher dimensions, we would like to exploit the sign of the derivative to identify whether a function is convex or concave at some point. But how should we think of the "sign" of a matrix? The concept of definiteness that we saw above will generalize the idea of positive and negative into higher dimensions. $f : \mathbb{R}^n \to \mathbb{R}^m$ is convex at x iff and only if $H_f(x)$ is positive semidefinite and concave iff $H_f(x)$ is negative semidefinite.

A functional f on a convex set S of a linear space is **quasiconvex** if

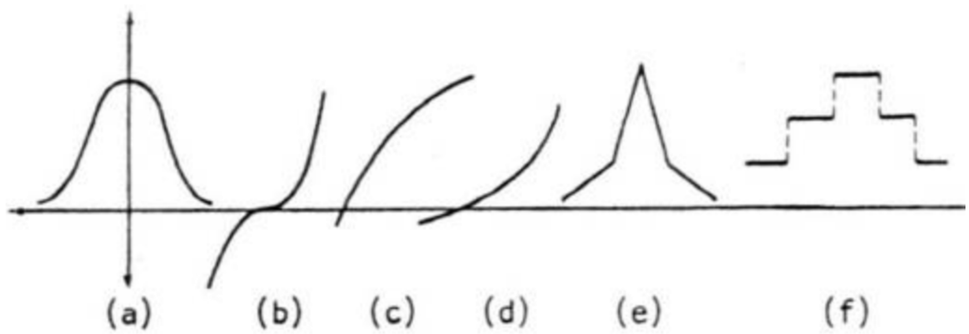
$$f(\alpha x_1 + (1-\alpha)x_2) \leq \max\{f(x_1), f(x_2)\}$$

for every $x_1, x_2 \in S$ and $\alpha \in [0,1]$. A function is **strictly quasiconvex** if the inequality is strict. A function is **quasiconcave** if

$$f(\alpha x_1 + (1-\alpha)x_2) \geq \min\{f(x_1), f(x_2)\}$$

for every $x_1, x_2 \in S$ and $\alpha \in [0,1]$. We often encounter quasiconcave functions rather than quasiconvex functions.

Geometrically a function is quasiconcave if the function along a line joining any two points in the domain lies above at least one of the endpoints. Any monotone function is both quasiconcave and quasiconvex. Each of the following functions is quasiconcave.



Exercises

1) Let $f_1, f_2, ..., f_n$ be convex functions and $\alpha_1, \alpha_2, ...\alpha_n \geq 0$. Prove that $f(x) = \alpha_1 f_1(x) + ... + \alpha_n f_n(x)$ is convex. Is $\alpha_1 f_1 - \alpha_2 f_2$ convex? Prove your answer.

2) Prove the Cauchy-Schwarz inequality for $\mathbb{R}^n$.

3) Prove the following: $L : \mathbb{R}^l \to \mathbb{R}$ is a continuous, linear functional if and only if there exists a $y \in \mathbb{R}^l$ such that for all $x \in \mathbb{R}^l$, $L(x) = y^T x$.

Source Material

Carter, M. (2001). *Foundations of Mathematical Economics*. Cambridge, MA: The MIT Press.

Corbae, D., Stinchcombe, M.B., & Zeman, J. (2009). *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*. Princeton, NJ: Princeton University Press.