MATHEMATICAL PRELIMINARIES FOR THE CLASS 6.972 [1]

0.1 LINEAR ALGEBRA AND ANALYSIS

In this section, we list some basic definitions, notational conventions, and results from linear algebra and analysis. For related and additional material, we recommend the books by Hoffman and Kunze [71], Lancaster and Tismenetsky [85], and Strang [76] (linear algebra), the books by Ash [72], Ortega and Rheinboldt [70], and Rudin [76] (analysis).

Set Notation

If X is a set and x is an element of X, we write $x \in X$. A set can be specified in the form $X = \{x \mid x \text{ satisfies } P\}$, as the set of all elements satisfying property P. The union of two sets X_1 and X_2 is denoted by $X_1 \cup X_2$ and their intersection by $X_1 \cap X_2$. The symbols $\exists$ and $\forall$ have the meanings "there exists" and "for all," respectively. The empty set is denoted by $\emptyset$.

The set of real numbers (also referred to as scalars) is denoted by $\Re$. The set $\Re$ augmented with $+\infty$ and $-\infty$ is called the *set of extended real numbers*. We denote by $[a, b]$ the set of (possibly extended) real numbers x satisfying $a \leq x \leq b$. A rounded, instead of square, bracket denotes strict inequality in the definition. Thus $(a, b]$, $[a, b)$, and (a, b) denote the set of all x satisfying $a < x \leq b$, $a \leq x < b$, and $a < x < b$, respectively. When working with extended real numbers, we use the natural extensions of the rules of arithmetic: $x \cdot 0 = 0$ for every extended real number x, $x \cdot \infty = \infty$ if $x > 0$, $x \cdot \infty = -\infty$ if $x < 0$, and $x + \infty = \infty$ and $x - \infty = -\infty$ for every scalar x. The expression $\infty - \infty$ is meaningless and is never allowed to occur.

If f is a function, we use the notation $f : X \mapsto Y$ to indicate the fact that f is defined on a set X (its *domain*) and takes values in a set Y (its *range*). If $f : X \mapsto Y$ is a function, and U and V are subsets of X and Y, respectively, the set $\big\{f(x) \mid x \in U\big\}$ is called the *image* or *forward image of* U, and the set $\big\{x \in \Re^n \mid f(x) \in V\big\}$ is called the *inverse image of* V.

0.1.1 Vectors and Matrices

We denote by $\Re^n$ the set of n-dimensional real vectors. For any $x \in \Re^n$, we use x_i to indicate its ith *coordinate*, also called its ith *component*.

[1] Portions of these notes draw from the book Convex Analysis and Optimization, by Bertsekas, Nedic, and Ozdaglar

Vectors in $\Re^n$ will be viewed as column vectors, unless the contrary is explicitly stated. For any $x \in \Re^n$, x' denotes the transpose of x, which is an n-dimensional row vector. The *inner product* of two vectors $x, y \in \Re^n$ is defined by $x'y = \sum_{i=1}^n x_i y_i$. Any two vectors $x, y \in \Re^n$ satisfying $x'y = 0$ are called *orthogonal*.

If x is a vector in $\Re^n$, the notations $x > 0$ and $x \geq 0$ indicate that all coordinates of x are positive and nonnegative, respectively. For any two vectors x and y, the notation $x > y$ means that $x - y > 0$. The notations $x \geq y$, $x < y$, etc., are to be interpreted accordingly.

If X is a set and λ is a scalar we denote by λX the set $\{\lambda x \mid x \in X\}$. If X_1 and X_2 are two subsets of $\Re^n$, we denote by $X_1 + X_2$ the *vector sum*

$$\{x_1 + x_2 \mid x_1 \in X_1,\ x_2 \in X_2\}.$$

We use a similar notation for the sum of any finite number of subsets. In the case where one of the subsets consists of a single vector $\overline{x}$, we simplify this notation as follows:

$$\overline{x} + X = \{\overline{x} + x \mid x \in X\}.$$

Given sets $X_i \subset \Re^{n_i}$, $i = 1, \ldots, m$, the *Cartesian product* of the X_i, denoted by $X_1 \times \cdots \times X_m$, is the subset

$$\big\{(x_1, \ldots, x_m) \mid x_i \in X_i,\ i = 1, \ldots, m\big\}$$

of $\Re^{n_1 + \cdots + n_m}$.

Subspaces and Linear Independence

A subset S of $\Re^n$ is called a *subspace* if $ax + by \in S$ for every $x, y \in X$ and every $a, b \in \Re$. An *affine set* in $\Re^n$ is a translated subspace, i.e., a set of the form $\overline{x} + S = \{\overline{x} + x \mid x \in S\}$, where $\overline{x}$ is a vector in $\Re^n$ and S is a subspace of $\Re^n$. The *span* of a finite collection $\{x_1, \ldots, x_m\}$ of elements of $\Re^n$ (also called the *subspace generated* by the collection) is the subspace consisting of all vectors y of the form $y = \sum_{k=1}^m a_k x_k$, where each a_k is a scalar.

The vectors $x_1, \ldots, x_m \in \Re^n$ are called *linearly independent* if there exists no set of scalars $a_1, \ldots, a_m$ such that $\sum_{k=1}^m a_k x_k = 0$, unless $a_k = 0$ for each k. An equivalent definition is that $x_1 \neq 0$, and for every $k > 1$, the vector x_k does not belong to the span of $x_1, \ldots, x_{k-1}$.

If S is a subspace of $\Re^n$ containing at least one nonzero vector, a *basis* for S is a collection of vectors that are linearly independent and whose span is equal to S. Every basis of a given subspace has the same number of vectors. This number is called the *dimension* of S. By convention, the subspace $\{0\}$ is said to have dimension zero. The *dimension of an affine set*

$\overline{x} + S$ is the dimension of the corresponding subspace S. Every subspace of nonzero dimension has an *orthogonal basis*, i.e., a basis consisting of mutually orthogonal vectors.

Given any set X, the set of vectors that are orthogonal to all elements of X is a subspace denoted by $X^\perp$:

$$X^\perp = \{y \mid y'x = 0,\ \forall\ x \in X\}.$$

If S is a subspace, $S^\perp$ is called the *orthogonal complement* of S. It can be shown that $(S^\perp)^\perp = S$. Furthermore, any vector x can be uniquely decomposed as the sum of a vector from S and a vector from $S^\perp$.

Matrices

For any matrix A, we use A_{ij}, $[A]_{ij}$, or a_{ij} to denote its ijth element. The *transpose* of A, denoted by A', is defined by $[A']_{ij} = a_{ji}$. For any two matrices A and B of compatible dimensions, the transpose of the product matrix AB satisfies $(AB)' = B'A'$.

If X is a subset of $\Re^n$ and A is an $m \times n$ matrix, then the image of X under A is denoted by AX (or $A \cdot X$ if this enhances notational clarity):

$$AX = \{Ax \mid x \in X\}.$$

If X is subspace, then AX is also a subspace.

Let A be a square matrix. We say that A is *symmetric* if $A' = A$. We say that A is *diagonal* if $[A]_{ij} = 0$ whenever $i \neq j$. We use I to denote the identity matrix. The *determinant* of A is denoted by $\det(A)$.

Let A be an $m \times n$ matrix. The *range space* of A, denoted by $R(A)$, is the set of all vectors $y \in \Re^m$ such that $y = Ax$ for some $x \in \Re^n$. The *null space* of A, denoted by $N(A)$, is the set of all vectors $x \in \Re^n$ such that $Ax = 0$. It is seen that the range space and the null space of A are subspaces. The *rank* of A is the dimension of the range space of A. The rank of A is equal to the maximal number of linearly independent columns of A, and is also equal to the maximal number of linearly independent rows of A. The matrix A and its transpose A' have the same rank. We say that A has *full rank*, if its rank is equal to $\min\{m, n\}$. This is true if and only if either all the rows of A are linearly independent, or all the columns of A are linearly independent.

The range of an $m \times n$ matrix A and the orthogonal complement of the nullspace of its transpose are equal, i.e.,

$$R(A) = N(A')^\perp.$$

Another way to state this result is that given vectors $a_1, \ldots, a_n \in \Re^m$ (the columns of A) and a vector $x \in \Re^m$, we have $x'y = 0$ for all y such that

$a_i'y = 0$ for all i if and only if $x = \lambda_1 a_1 + \cdots + \lambda_n a_n$ for some scalars $\lambda_1, \ldots, \lambda_n$. This is a special case of Farkas' Lemma, an important result for constrained optimization. A useful application of this result is that if S_1 and S_2 are two subspaces of $\Re^n$, then

$$S_1^\perp + S_2^\perp = (S_1 \cap S_2)^\perp.$$

This follows by introducing matrices B_1 and B_2 such that $S_1 = \{x \mid B_1 x = 0\} = N(B_1)$ and $S_2 = \{x \mid B_2 x = 0\} = N(B_2)$, and writing

$$S_1^\perp + S_2^\perp = R\big([\,B_1' \quad B_2'\,]\big) = N\left(\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}\right)^\perp = \big(N(B_1) \cap N(B_2)\big)^\perp = (S_1 \cap S_2)^\perp$$

A function $f : \Re^n \mapsto \Re$ is said to be *affine* if it has the form $f(x) = a'x + b$ for some $a \in \Re^n$ and $b \in \Re$. Similarly, a function $f : \Re^n \mapsto \Re^m$ is said to be *affine* if it has the form $f(x) = Ax + b$ for some $m \times n$ matrix A and some $b \in \Re^m$. If $b = 0$, f is said to be a *linear function* or *linear transformation*.

0.1.2 Topological Properties

Definition 0.1.1: A *norm* $\|\cdot\|$ on $\Re^n$ is a function that assigns a scalar $\|x\|$ to every $x \in \Re^n$ and that has the following properties:

(a) $\|x\| \geq 0$ for all $x \in \Re^n$.

(b) $\|\alpha x\| = |\alpha| \cdot \|x\|$ for every scalar α and every $x \in \Re^n$.

(c) $\|x\| = 0$ if and only if $x = 0$.

(d) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in \Re^n$ (this is referred to as the *triangle inequality*).

The *Euclidean norm* of a vector $x = (x_1, \ldots, x_n)$ is defined by

$$\|x\| = (x'x)^{1/2} = \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}.$$

The space $\Re^n$, equipped with this norm, is called a *Euclidean space*. We will use the Euclidean norm almost exclusively in this book. In particular, *in the absence of a clear indication to the contrary, $\|\cdot\|$ will denote the Euclidean norm*. Two important results for the Euclidean norm are:

Proposition 0.1.1: (Pythagorean Theorem) For any two vectors x and y that are orthogonal, we have

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Proposition 0.1.2: (Schwartz inequality) For any two vectors x and y, we have

$$|x'y| \leq \|x\| \cdot \|y\|,$$

with equality holding if and only if $x = \alpha y$ for some scalar α.

Two other important norms are the *maximum norm* $\|\cdot\|_\infty$ (also called *sup-norm* or *ℓ_∞-norm*), defined by

$$\|x\|_\infty = \max_i |x_i|,$$

and the *ℓ_1-norm* $\|\cdot\|_1$, defined by

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

Sequences

We use both subscripts and superscripts in sequence notation. Generally, we prefer subscripts, but we use superscripts whenever we need to reserve the subscript notation for indexing coordinates or components of vectors and functions. The meaning of the subscripts and superscripts should be clear from the context in which they are used.

A sequence $\{x_k \mid k = 1, 2, \ldots\}$ (or $\{x_k\}$ for short) of scalars is said to *converge* if there exists a scalar x such that for every $\epsilon > 0$ we have $|x_k - x| < \epsilon$ for every k greater than some integer K (depending on ϵ). We call the scalar x the *limit* of $\{x_k\}$, and we also say that $\{x_k\}$ *converges to* x; symbolically, $x_k \to x$ or $\lim_{k\to\infty} x_k = x$. If for every scalar b there exists some K (depending on b) such that $x_k \geq b$ for all $k \geq K$, we write $x_k \to \infty$ and $\lim_{k\to\infty} x_k = \infty$. Similarly, if for every scalar b there exists some K such that $x_k \leq b$ for all $k \geq K$, we write $x_k \to -\infty$ and $\lim_{k\to\infty} x_k = -\infty$.

A sequence $\{x_k\}$ is called a *Cauchy sequence* if for every $\epsilon > 0$, there exists some K (depending on ϵ) such that $|x_k - x_m| < \epsilon$ for all $k \geq K$ and $m \geq K$.

A sequence $\{x_k\}$ is said to be *bounded above* (respectively, *below*) if there exists some scalar b such that $x_k \leq b$ (respectively, $x_k \geq b$) for all k. It is said to be *bounded* if it is bounded above and bounded below. The sequence $\{x_k\}$ is said to be monotonically *nonincreasing* (respectively, *nondecreasing*) if $x_{k+1} \leq x_k$ (respectively, $x_{k+1} \geq x_k$) for all k. If $\{x_k\}$ converges to x and is nonincreasing (nondecreasing), we also use the notation $x_k \downarrow x$ ($x_k \uparrow x$, respectively).

Proposition 0.1.3: Every bounded and monotonically nonincreasing or nondecreasing scalar sequence converges.

Note that a monotonically nondecreasing sequence $\{x_k\}$ is either bounded, in which case it converges to some scalar x by the above proposition, or else it is unbounded, in which case $x_k \to \infty$. Similarly, a monotonically nonincreasing sequence $\{x_k\}$ is either bounded and converges, or it is unbounded, in which case $x_k \to -\infty$.

The *supremum* of a nonempty set X of scalars, denoted by $\sup X$, is defined as the smallest scalar x such that $x \geq y$ for all $y \in X$. If no such scalar exists, we say that the supremum of X is ∞. Similarly, the *infimum* of X, denoted by $\inf X$, is defined as the largest scalar x such that $x \leq y$ for all $y \in X$, and is equal to $-\infty$ if no such scalar exists. For the empty set, we use the convention

$$\sup(\emptyset) = -\infty, \qquad \inf(\emptyset) = \infty.$$

(This is somewhat paradoxical, since we have that the sup of a set is less than its inf, but works well for our analysis.) If $\sup X$ is equal to a scalar $\overline{x}$ that belongs to the set X, we say that $\overline{x}$ is the *maximum point* of X and we often write

$$\overline{x} = \sup X = \max X.$$

Similarly, if $\inf X$ is equal to a scalar $\overline{x}$ that belongs to the set X, we often write

$$\overline{x} = \inf X = \min X.$$

Thus, when we write $\max X$ (or $\min X$) in place of $\sup X$ (or $\inf X$, respectively) we do so just for emphasis: we indicate that it is either evident, or it is known through earlier analysis, or it is about to be shown that the maximum (or minimum, respectively) of the set X is attained at one of its points.

Given a scalar sequence $\{x_k\}$, the supremum of the sequence, denoted by $\sup_k x_k$, is defined as $\sup\{x_k \mid k = 1, 2, \ldots\}$. The infimum of a sequence is similarly defined. Given a sequence $\{x_k\}$, let $y_m = \sup\{x_k \mid k \geq m\}$, $z_m = \inf\{x_k \mid k \geq m\}$. The sequences $\{y_m\}$ and $\{z_m\}$ are nonincreasing and nondecreasing, respectively, and therefore have a limit whenever $\{x_k\}$ is bounded above or is bounded below, respectively (Prop. 0.1.3). The limit of y_m is denoted by $\limsup_{k\to\infty} x_k$, and is referred to as the *limit superior* of $\{x_k\}$. The limit of z_m is denoted by $\liminf_{k\to\infty} x_k$, and is referred to as the *limit inferior* of $\{x_k\}$. If $\{x_k\}$ is unbounded above, we write $\limsup_{k\to\infty} x_k = \infty$, and if it is unbounded below, we write $\liminf_{k\to\infty} x_k = -\infty$.

Proposition 0.1.4: Let $\{x_k\}$ and $\{y_k\}$ be scalar sequences.

(a) There holds

$$\inf_k x_k \leq \liminf_{k\to\infty} x_k \leq \limsup_{k\to\infty} x_k \leq \sup_k x_k.$$

(b) $\{x_k\}$ converges if and only if $\liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k$ and, in that case, both of these quantities are equal to the limit of x_k.

(c) If $x_k \leq y_k$ for all k, then

$$\liminf_{k\to\infty} x_k \leq \liminf_{k\to\infty} y_k, \qquad \limsup_{k\to\infty} x_k \leq \limsup_{k\to\infty} y_k.$$

(d) We have

$$\liminf_{k\to\infty} x_k + \liminf_{k\to\infty} y_k \leq \liminf_{k\to\infty}(x_k + y_k),$$

$$\limsup_{k\to\infty} x_k + \limsup_{k\to\infty} y_k \geq \limsup_{k\to\infty}(x_k + y_k).$$

A sequence $\{x_k\}$ of vectors in $\Re^n$ is said to converge to some $x \in \Re^n$ if the ith coordinate of x_k converges to the ith coordinate of x for every i. We use the notations $x_k \to x$ and $\lim_{k\to\infty} x_k = x$ to indicate convergence for vector sequences as well. The sequence $\{x_k\}$ is called bounded (respectively, a Cauchy sequence) if each of its corresponding coordinate sequences is bounded (respectively, a Cauchy sequence). It can be seen that $\{x_k\}$ is bounded if and only if there exists a scalar c such that $\|x_k\| \leq c$ for all k.

Definition 0.1.2: We say that a vector $x \in \Re^n$ is a *limit point* of a sequence $\{x_k\}$ in $\Re^n$ if there exists a subsequence of $\{x_k\}$ that converges to x.

Proposition 0.1.5: Let $\{x_k\}$ be a sequence in $\Re^n$.

(a) If $\{x_k\}$ is bounded, it has at least one limit point.

(b) $\{x_k\}$ converges if and only if it is bounded and it has a unique limit point.

(c) $\{x_k\}$ converges if and only if it is a Cauchy sequence.

Closed and Open Sets

We say that x is a *closure point* or *limit point* of a set $X \subset \Re^n$ if there exists a sequence $\{x_k\}$, consisting of elements of X, that converges to x. The *closure* of X, denoted $\mathrm{cl}(X)$, is the set of all limit points of X.

Definition 0.1.3: A set $X \subset \Re^n$ is called *closed* if it is equal to its closure. It is called *open* if its complement (the set $\{x \mid x \notin X\}$) is closed. It is called *bounded* if there exists a scalar c such that the magnitude of any coordinate of any element of X is less than c. It is called *compact* if it is closed and bounded.

Definition 0.1.4: A *neighborhood* of a vector x is an open set containing x. We say that x is an *interior* point of a set $X \subset \Re^n$ if there exists a neighborhood of x that is contained in X. A vector $x \in \mathrm{cl}(X)$ which is not an interior point of X is said to be a *boundary* point of X.

Let $\|\cdot\|$ be a given norm in $\Re^n$. For any $\epsilon > 0$ and $x^* \in \Re^n$, consider the sets

$$\big\{x \mid \|x - x^*\| < \epsilon\big\}, \qquad \big\{x \mid \|x - x^*\| \leq \epsilon\big\}.$$

The first set is open and is called an *open sphere* centered at x^*, while the second set is closed and is called a *closed sphere* centered at x^*. Sometimes the terms *open ball* and *closed ball* are used, respectively.

Proposition 0.1.6:

(a) The union of finitely many closed sets is closed.

(b) The intersection of closed sets is closed.

(c) The union of open sets is open.

(d) The intersection of finitely many open sets is open.

(e) A set is open if and only if all of its elements are interior points.

(f) Every subspace of $\Re^n$ is closed.

(g) A set X is compact if and only if every sequence of elements of X has a subsequence that converges to an element of X.

(h) If $\{X_k\}$ is a sequence of nonempty and compact sets such that $X_k \supset X_{k+1}$ for all k, then the intersection $\cap_{k=0}^{\infty} X_k$ is nonempty and compact.

The topological properties of subsets of $\Re^n$, such as being open, closed, or compact, do not depend on the norm being used. This is a consequence of the following proposition, referred to as the *norm equivalence property in* $\Re^n$, which shows that if a sequence converges with respect to one norm, it converges with respect to all other norms.

Proposition 0.1.7: For any two norms $\|\cdot\|$ and $\|\cdot\|'$ on $\Re^n$, there exists a scalar c such that $\|x\| \leq c\|x\|'$ for all $x \in \Re^n$.

Using the preceding proposition, we obtain the following.

Proposition 0.1.8: If a subset of $\Re^n$ is open (respectively, closed, bounded, or compact) with respect to some norm, it is open (respectively, closed, bounded, orcompact) with respect to all other norms.

Continuity

Let $f : X \mapsto \Re^m$ be a function, where X is a subset of $\Re^n$, and let x be a point in X. If there exists a vector $y \in \Re^m$ such that the sequence $\big\{f(x_k)\big\}$

converges to y for every sequence $\{x_k\} \subset X$ such that $\lim_{k\to\infty} x_k = x$, we write $\lim_{z\to x} f(z) = y$. If there exists a vector $y \in \Re^m$ such that the sequence $\big\{f(x_k)\big\}$ converges to y for every sequence $\{x_k\} \subset X$ such that $\lim_{k\to\infty} x_k = x$ and $x_k \leq x$ (respectively, $x_k \geq x$) for all k, we write $\lim_{z\uparrow x} f(z) = y$ [respectively, $\lim_{z\downarrow x} f(z)$].

Definition 0.1.5: Let X be a subset of $\Re^n$.

(a) A function $f : X \mapsto \Re^m$ is called *continuous* at a point $x \in X$ if $\lim_{z\to x} f(z) = f(x)$.

(b) A function $f : X \mapsto \Re^m$ is called *right-continuous* (respectively, *left-continuous*) at a point $x \in X$ if $\lim_{z\downarrow x} f(z) = f(x)$ [respectively, $\lim_{z\uparrow x} f(z) = f(x)$].

(c) A real-valued function $f : X \mapsto \Re$ is called *upper semicontinuous* (respectively, *lower semicontinuous*) at a point $x \in X$ if $f(x) \geq \limsup_{k\to\infty} f(x_k)$ [respectively, $f(x) \leq \liminf_{k\to\infty} f(x_k)$] for every sequence $\{x_k\}$ of elements of X converging to x.

If $f : X \mapsto \Re^m$ is continuous at every point of a subset of its domain X, we say that f *is continuous over that subset*. If $f : X \mapsto \Re^m$ is continuous at every point of its domain X, we say that f *is continuous*. We use similar terminology for right-continuous, left-continuous, upper semicontinuous, and lower semicontinuous functions.

If $f : X \mapsto \Re^m$ is continuous at every point of a subset of its domain X, we say that f *is continuous over that subset*. If $f : X \mapsto \Re^m$ is continuous at every point of its domain X, we say that f *is continuous*. We use similar terminology for right-continuous, left-continuous, upper semicontinuous, and lower semicontinuous functions.

Proposition 0.1.9:

(a) The composition of two continuous functions is continuous.

(b) Any vector norm on $\Re^n$ is a continuous function.

(c) Let $f : \Re^n \mapsto \Re^m$ be continuous, and let $Y \subset \Re^m$ be open (respectively, closed). Then the inverse image of Y, $\big\{x \in \Re^n \mid f(x) \in Y\big\}$, is open (respectively, closed).

(d) Let $f : \Re^n \mapsto \Re^m$ be continuous, and let $X \subset \Re^n$ be compact. Then the forward image of X, $\big\{f(x) \mid x \in X\big\}$, is compact.

0.1.3 Continuous Correspondences

Let X be a subset of re^n and Y be a compact subset of $\Re^k$. A correspondence from X to Y is a map that assigns each element of $x \in X$ a subset $\Gamma(x) \subset Y$. To distinguish a correspondence from a function, we will use the notation $\Gamma : X \rightrightarrows Y$. A correspondence Γ is said to be closed-(compact-)valued at x if $\Gamma(x)$ is a closed (compact) set.

Definition 0.1.6: A compact-valued correspondence $\Gamma : X \mapsto Y$ is upper semicontinuous (hemicontinuous) (usc) at the point $\bar{x}$ if $\Gamma(\bar{x})$ is nonempty and for every sequence $x^k \to \bar{x}$ and every sequence $\{y^k\}$ with $y^k \in \Gamma(x^k)$, y^k converges to $\bar{y} \in \Gamma(\bar{x})$.

Note that our definition of upper semicontinuity applies only to compact-valued correspondences, which is sufficient for our purposes.

Definition 0.1.6: A correspondence Γ is lower semicontinuous (hemicontinuous) (lsc) at the point $\bar{x}$ if if $\Gamma(\bar{x})$ is nonempty and for every sequence $x^k \to \bar{x}$ and every $\bar{y} \in \Gamma(\bar{x})$, there exists a sequence $\{y^k\}$ such that $y^k \in \Gamma(x^k)$ and $y^k \to \bar{y}$.

The next example illustrates the semicontinuity definitions.

Example: Let $X = Y = [0, 2]$. Define the correspondence $\Gamma_1 : X \rightrightarrows Y$ by

$$\Gamma_1(x) = \begin{cases} 1, & 0 \leq x < 1, \\ [0, 2], & 1 \leq x \leq 2. \end{cases}$$

Γ_1 is both usc and lsc at all $x \neq 1$. At $x = 1$, Γ_1 is usc but not lsc. Next consider the correspondence $\Gamma_2 : X \rightrightarrows Y$ defined by

$$\Gamma_1(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ [0, 2], & 1 < x \leq 2. \end{cases}$$

Γ_2 is both usc and lsc at all $x \neq 1$. At $x = 1$, Γ_1 is lsc but not usc.

The correspondence Γ is continuous at the point $\bar{x}$ if it is upper and lower semicontinuous at $\bar{x}$. Semicontinuities and continuity on X are defined as semicontinuities and continuity at every point of S. It is immediate from the definitions that a single-valued correspondence that is semicontinuous (usc or lsc) is continuous when viewed as a function. The notions of usc and lsc for functions, defined in Definition 0.1.5, are different from those for correspondences. Therefore, the terminology is unfortunately sometimes misleading. This is perhaps why in the economics literature the term "hemicontinuity" is reserved when referring to correspondences. In these notes, we will use the term "semicontinuity" for both correspondences and functions.

The graph of a correspondence $\Gamma : X \Rightarrow Y$ is defined by

$$Graph(\Gamma) = \{(x, y) \in X \times Y \mid y \in \Gamma(x)\}.$$

It is immediate from the definition that the correspondence Γ is upper semicontinuous on X if and only if its graph is a closed set in $X \times Y$. (Note that this equivalence is valid only when the range of the correspondence is compact and the correspondence itself is compact-valued. In general, it is not true that an usc correspondence necessarily has a closed graph, nor it is true that a correspondence with a closed graph is usc.)

We next present the Theorem of the Maximum. The Theorem of the Maximum is a statement about the continuity properties of the solution set of a parametric optimization problem provided that the underlying optimization problem possesses some degree of continuity.

Proposition 0.1.10: (Theorem of the Maximum) Let $X \subset \Re^n$ and $Y \subset \Re^k$, let $f : X \times Y \mapsto \Re$ be a continuous function, and $\Gamma : X \to Y$ be a compact-valued and continuous correspondence. Then the function $h : X \mapsto \Re$ given by

$$h(x) = \max_{y \in \Gamma(x)} f(x, y),$$

is continuous, and the correspondence $G : X \to Y$ given by

$$G(x) = \arg \max_{y \in \Gamma(x)} f(x, y) = \{y \in \Gamma(x) \mid f(x, y) = h(x)\},$$

is nonempty, compact-valued, and upper semicontinuous.

0.1.4 Fixed Point Theorems

In most problems, establishing the existence of an equilibrium of a system is cast as finding a fixed point of a suitably constructed function or correspondence. We next state the classical fixed point theorems (proofs will be included later).

Proposition 0.1.11: (Brouwer's Fixed Point Theorem) Suppose that $A \subset \Re^n$ is a nonempty, compact, convex set, and that $f : A \mapsto A$ is a continuous function. Then $f(\cdot)$ has a fixed point; that is there exists some $x \in A$ such that $x = f(x)$.

In most applications, the following extension of Brouwer's fixed point theorem to correspondences is useful.

Proposition 0.1.12: (Kakutani's Fixed Point Theorem) Suppose that $A \subset \Re^n$ is a nonempty, compact, convex set, and that $f : A \Rightarrow A$ is an usc correspondence that is nonempty and convex-valued for all $x \in A$. Then $f(\cdot)$ has a fixed point; that is there exists some $x \in A$ such that $x \in f(x)$.

0.1.5 Square Matrices

Definition 0.1.8: A square matrix A is called *singular* if its determinant is zero. Otherwise it is called *nonsingular* or *invertible*.

Proposition 0.1.13:

(a) Let A be an $n \times n$ matrix. The following are equivalent:

  (i) The matrix A is nonsingular.

  (ii) The matrix A' is nonsingular.

  (iii) For every nonzero $x \in \Re^n$, we have $Ax \neq 0$.

  (iv) For every $y \in \Re^n$, there is a unique $x \in \Re^n$ such that $Ax = y$.

  (v) There is an $n \times n$ matrix B such that $AB = I = BA$.

  (vi) The columns of A are linearly independent.

  (vii) The rows of A are linearly independent.

(b) Assuming that A is nonsingular, the matrix B of statement (v) (called the *inverse* of A and denoted by A^{-1}) is unique.

(c) For any two square invertible matrices A and B of the same dimensions, we have $(AB)^{-1} = B^{-1}A^{-1}$.

Definition 0.1.9: The *characteristic polynomial* ϕ of an $n \times n$ matrix A is defined by $\phi(\lambda) = \det(\lambda I - A)$, where I is the identity matrix of the same size as A. The n (possibly repeated or complex) roots of ϕ are called the *eigenvalues* of A. A nonzero vector x (with possibly complex coordinates) such that $Ax = \lambda x$, where λ is an eigenvalue of A, is called an *eigenvector* of A associated with λ.

Proposition 0.1.14: Let A be a square matrix.

(a) A complex number λ is an eigenvalue of A if and only if there exists a nonzero eigenvector associated with λ.

(b) A is singular if and only if it has an eigenvalue that is equal to zero.

Proposition 0.1.15: Let A be an $n \times n$ matrix.

(a) If T is a nonsingular matrix and $B = TAT^{-1}$, then the eigenvalues of A and B coincide.

(b) For any scalar c, the eigenvalues of $cI + A$ are equal to $c + \lambda_1, \ldots, c + \lambda_n$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

(c) The eigenvalues of A^k are equal to $\lambda_1^k, \ldots, \lambda_n^k$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

(d) If A is nonsingular, then the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A.

(e) The eigenvalues of A and A' coincide.

Symmetric and Positive Definite Matrices

Symmetric matrices have several special properties, particularly regarding theireigenvalues and eigenvectors. In what follows in this section, $\|\cdot\|$ denotes the Euclidean norm.

Proposition 0.1.16: Let A be a symmetric $n \times n$ matrix. Then:

(a) The eigenvalues of A are real.

(b) The matrix A has a set of n mutually orthogonal, real, and nonzero eigenvectors $x_1, \dots, x_n$.

(c) Suppose that the eigenvectors in part (b) have been normalized so that $\|x_i\| = 1$ for each i. Then

$$A = \sum_{i=1}^{n} \lambda_i x_i x_i',$$

where λ_i is the eigenvalue corresponding to x_i.

Proposition 0.1.17: Let A be a symmetric $n \times n$ matrix, and let $\lambda_1 \le \cdots \le \lambda_n$ be its (real) eigenvalues. Then:

(a) $\|A\| = \max\big\{|\lambda_1|, |\lambda_n|\big\}$, where $\|\cdot\|$ is the matrix norm induced by the Euclidean norm.

(b) $\lambda_1 \|y\|^2 \le y'Ay \le \lambda_n \|y\|^2$ for all $y \in \Re^n$.

(c) If A is nonsingular, then

$$\|A^{-1}\| = \frac{1}{\min\big\{|\lambda_1|, |\lambda_n|\big\}}.$$

Proposition 0.1.18: Let A be a square matrix, and let $\|\cdot\|$ be the matrix norm induced by the Euclidean norm. Then:

(a) If A is symmetric, then $\|A^k\| = \|A\|^k$ for any positive integer k.

(b) $\|A\|^2 = \|A'A\| = \|AA'\|$.

Definition 0.1.10: A symmetric $n \times n$ matrix A is called *positive definite* if $x'Ax > 0$ for all $x \in \Re^n$, $x \ne 0$. It is called *positive semidefinite* if $x'Ax \ge 0$ for all $x \in \Re^n$.

In these notes, the notion of positive definiteness applies exclusively to symmetric matrices. Thus *whenever we say that a matrix is positive (semi)definite, we implicitly assume that the matrix is symmetric.*

Proposition 0.1.19:

(a) The sum of two positive semidefinite matrices is positive semidefinite. If one of the two matrices is positive definite, the sum is positive definite.

(b) If A is a positive semidefinite $n \times n$ matrix and T is an $m \times n$ matrix, then the matrix TAT' is positive semidefinite. If A is positive definite and T is invertible, then TAT' is positive definite.

Proposition 0.1.20:

(a) For any $m \times n$ matrix A, the matrix $A'A$ is symmetric and positive semidefinite. $A'A$ is positive definite if and only if A has rank n. In particular, if $m = n$, $A'A$ is positive definite if and only if A is nonsingular.

(b) A square symmetric matrix is positive semidefinite (respectively, positive definite) if and only if all of its eigenvalues are nonnegative (respectively, positive).

(c) The inverse of a symmetric positive definite matrix is symmetric and positive definite.

Proposition 0.1.21: Let A be a symmetric positive semidefinite $n \times n$ matrix of rank m. There exists an $n \times m$ matrix C of rank m such that

$$A = CC'.$$

Furthermore, for any such matrix C:

(a) A and C' have the same null space: $N(A) = N(C')$.

(b) A and C have the same range space: $R(A) = R(C)$.

Proposition 0.1.22: Let A be a square symmetric positive semidefinite matrix.

(a) There exists a symmetric matrix Q with the property $Q^2 = A$. Such a matrix is called a *symmetric square root* of A and is denoted by $A^{1/2}$.

(b) There is a unique symmetric square root if and only if A is positive definite.

(c) A symmetric square root $A^{1/2}$ is invertible if and only if A is invertible. Its inverse is denoted by $A^{-1/2}$.

(d) There holds $A^{-1/2}A^{-1/2} = A^{-1}$.

(e) There holds $AA^{1/2} = A^{1/2}A$.

0.2 CONVEX SETS AND FUNCTIONS

We now introduce some of the basic notions relating to convex sets and functions.

0.2.1 Basic Properties

The notion of a convex set is defined below.

Definition 0.2.1: Let C be a subset of $\Re^n$. We say that C is *convex* if

$$\alpha x + (1-\alpha)y \in C, \qquad \forall\ x, y \in C,\ \forall\ \alpha \in [0,1]. \tag{0.0}$$

Note that the empty set is by convention considered to be convex. Generally, when referring to a convex set, it will usually be apparent from the context whether this set can be empty, but we will often be specific in order to minimize ambiguities.

The following proposition lists some operations that preserve convexity of a set.

Proposition 0.2.1:

(a) The intersection $\cap_{i \in I} C_i$ of any collection $\{C_i \mid i \in I\}$ of convex sets is convex.

(b) The vector sum $C_1 + C_2$ of two convex sets C_1 and C_2 is convex.

(c) The cartesian product $C_1 \times C_2$ of two convex sets C_1 and C_2 is convex.

(d) The set $x + \lambda C$ is convex for any convex set C, vector x, and scalar λ. Furthermore, if C is a convex set and λ_1, λ_2 are positive scalars, we have

$$(\lambda_1 + \lambda_2)C = \lambda_1 C + \lambda_2 C.$$

(e) The closure and the interior of a convex set are convex.

(f) The image and the inverse image of a convex set under an affine function are convex.

Proof: The proof is straightforward using the definition of convexity, cf. Eq. (0.0). For example, to prove part (a), we take two points x and y from $\cap_{i \in I} C_i$, and we use the convexity of C_i to argue that the line segment connecting x and y belongs to all the sets C_i, and hence, to their intersection. The proofs of parts (b)-(f) are similar and are left as exercises for the reader. **Q.E.D.**

A set C is said to be a *cone* if for all $x \in C$ and $\lambda > 0$, we have $\lambda x \in C$. A cone need not be convex and need not contain the origin (although the origin always lies in the closure of a nonempty cone). Several of the results of the above proposition have analogs for cones.

Convex Functions

The notion of a convex function is defined below.

Definition 0.2.2: Let C be a convex subset of $\Re^n$. A function $f : C \mapsto \Re$ is called *convex* if

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha)f(y), \qquad \forall\, x,y \in C,\ \forall\, \alpha \in [0,1]. \tag{0.1}$$

The function f is called *concave* if $-f$ is convex. The function f is called *strictly convex* if the above inequality is strict for all $x, y \in C$ with $x \neq y$, and all $\alpha \in (0,1)$. For a function $f : X \mapsto \Re$, we also say that f is *convex over the convex set* C if the domain X of f contains C and Eq. (0.1) holds, i.e., when the domain of f is restricted to C, f becomes convex.

If C is a convex set and $f : C \mapsto \Re$ is a convex function, the level sets $\{x \in C \mid f(x) \le \gamma\}$ and $\{x \in C \mid f(x) < \gamma\}$ are convex for all scalars γ. To see this, note that if $x, y \in C$ are such that $f(x) \le \gamma$ and $f(y) \le \gamma$, then for any $\alpha \in [0,1]$, we have $\alpha x + (1-\alpha)y \in C$, by the convexity of C, and $f\big((\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha)f(y) \le \gamma$, by the convexity of f. However, the converse is not true; for example, the function $f(x) = \sqrt{|x|}$ has convex level sets but is not convex.

We generally prefer to deal with convex functions that are real-valued and are defined over the entire space $\Re^n$ (rather than over just a convex subset). However, in some situations, prominently arising in the context of duality theory, we encounter functions f that are convex over a convex subset C and cannot be extended to functions that are convex over $\Re^n$. In such situations, it may be convenient, instead of restricting the domain of f to the subset C where it takes real values, to extend its domain to the entire space $\Re^n$ but allow f to take the value ∞.

We are thus motivated to introduce *extended real-valued* convex functions that can take the value of ∞ at some points. In particular, if $C \subset \Re^n$ is a convex set, a function $f : C \mapsto (-\infty, \infty]$ is called *convex* over C (or simply convex when $C = \Re^n$) if the condition

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha)f(y), \qquad \forall\, x,y \in C,\ \forall\, \alpha \in [0,1]$$

holds. The function f is called *strictly convex* if the above inequality is strict for all x and y in C such that $f(x) < \infty$ and $f(y) < \infty$. It can be seen that if f is convex, the level sets $\{x \in C \mid f(x) \le \gamma\}$ and $\{x \in C \mid f(x) < \gamma\}$ are convex for all scalars γ.

One complication when dealing with extended real-valued functions is that we sometimes need to be careful to exclude the unusual case where $f(x) = \infty$ for all $x \in C$, in which case f is still convex, since it trivially satisfies the above inequality (such functions are sometimes called *improper* in the literature of convex functions, but we will not use this terminology here). Another area of concern is working with functions that can take both values $-\infty$ and ∞, because of the possibility of the forbidden expression $\infty - \infty$ arising when sums of function values are considered. For this reason, we will try to minimize the occurances of such functionz. On occasion we will deal with extended real-valued functions that can take the value $-\infty$ but not the value ∞. In particular, a function $f : C \mapsto [-\infty, \infty)$, where C is a convex set, is called *concave* if the function $-f : C \mapsto (-\infty, \infty]$ is convex as per the preceding definition.

We define the *effective domain* of an extended real-valued function $f : X \mapsto (-\infty, \infty]$ to be the set

$$\mathrm{dom}(f) = \big\{x \in X \mid f(x) < \infty\big\}.$$

Note that if X is convex and f is convex over X, then $\mathrm{dom}(f)$ is a convex set. Similarly, we define the effective domain of an extended real-valued function $f : X \mapsto [-\infty, \infty)$ to be the set $\mathrm{dom}(f) = \big\{x \in X \mid f(x) > -\infty\big\}$. Note that by replacing the domain of an extended real-valued convex function with its effective domain, we can convert it to a real-valued function. In this way, we can use results stated in terms of real-valued functions, and we can also avoid calculations with ∞. Thus, the entire subject of convex functions can be developed without resorting to extended real-valued functions. The reverse is also true, namely that extended real-valued functions can be adopted as the norm; for example, the classical treatment of Rockafellar [70] uses this approach.

Epigraphs

An extended real-valued function $f : X \mapsto (-\infty, \infty]$ is called *lower semicontinuous* at a vector $x \in X$ if $f(x) \leq \liminf_{k\to\infty} f(x_k)$ for every sequence $\{x_k\}$ converging to x. This definition is consistent with the corresponding definition for real-valued functions [cf. Def. 0.1.5(c)]. If f is lower semicontinuous at every x in a subset $U \subset X$, we say that f is *lower semicontinuous over* U.

The *epigraph* of a function $f : X \mapsto (-\infty, \infty]$, where $X \subset \Re^n$, is the subset of $\Re^{n+1}$ given by

$$\mathrm{epi}(f) = \big\{(x, w) \mid x \in X,\ w \in \Re,\ f(x) \leq w\big\}.$$

Note that if we restrict f to its effective domain $\big\{x \in X \mid f(x) < \infty\big\}$, so that it becomes real-valued, the epigraph remains unaffected. Epigraphs are useful for our purposes because of the following proposition, which shows that questions about convexity and lower semicontinuity of functions can be reduced to corresponding questions of convexity and closure of their epigraphs.

Proposition 0.2.2: Let $f : X \mapsto (-\infty, \infty]$ be a function. Then:

(a) $\text{epi}(f)$ is convex if and only if the set X is convex and f is convex over X.

(b) Assuming $X = \Re^n$, the following are equivalent:

 (i) $\text{epi}(f)$ is closed.

 (ii) f is lower semicontinuous over $\Re^n$.

 (iii) The level sets $\{x \mid f(x) \leq \gamma\}$ are closed for all scalars γ.

Proof: (a) Assume that X is convex and f is convex over X. If (x_1, w_1) and (x_2, w_2) belong to $\text{epi}(f)$ and $\alpha \in [0,1]$, we have

$$f(x_1) \leq w_1, \qquad f(x_2) \leq w_2,$$

and by multiplying these inequalities with α and $(1-\alpha)$, respectively, by adding, and by using the convexity of f, we obtain

$$f\big(\alpha x_1 + (1-\alpha)x_2\big) \leq \alpha f(x_1) + (1-\alpha) f(x_2) \leq \alpha w_1 + (1-\alpha) w_2.$$

Hence the vector $\big(\alpha x_1 + (1-\alpha)x_2,\, \alpha w_1 + (1-\alpha)w_2\big)$, which is equal to $\alpha(x_1, w_1) + (1-\alpha)(x_2, w_2)$, belongs to $\text{epi}(f)$, showing the convexity of $\text{epi}(f)$.

Conversely, assume that $\text{epi}(f)$ is convex, and let $x_1, x_2 \in X$ and $\alpha \in [0,1]$. The pairs $\big(x_1, f(x_1)\big)$ and $\big(x_2, f(x_2)\big)$ belong to $\text{epi}(f)$, so by convexity, we have

$$\big(\alpha x_1 + (1-\alpha)x_2,\, \alpha f(x_1) + (1-\alpha) f(x_2)\big) \in \text{epi}(f).$$

Therefore, by the definition of $\text{epi}(f)$, it follows that $\alpha x_1 + (1-\alpha)x_2 \in X$, so X is convex, while

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

so f is convex over X.

(b) We first show that (i) and (ii) are equivalent. Assume that f is lower semicontinuous over $\Re^n$, and let (x, w) be the limit of a sequence $\{(x_k, w_k)\} \subset \text{epi}(f)$. We have $f(x_k) \leq w_k$, and by taking limit as $k \to \infty$ and by using the lower semicontinuity of f at x, we obtain $f(x) \leq \liminf_{k\to\infty} f(x_k) \leq w$. Hence $(x, w) \in \text{epi}(f)$ and $\text{epi}(f)$ is closed.

Conversely, assume that $\text{epi}(f)$ is closed, choose any $x \in \Re^n$, let $\{x_k\}$ be a sequence converging to x, and let $w = \liminf_{k\to\infty} f(x_k)$. We will show that $f(x) \leq w$. Indeed, if $w = \infty$, we have $f(x) \leq w$. If $w < \infty$, for each positive integer n, let $w_n = w + 1/n$, and let $k(n)$ be an integer such that $k(n) \geq n$ and $f\big(x_{k(n)}\big) \leq w_n$. The sequence $\big\{\big(x_{k(n)}, w_n\big)\big\}$ belongs to $\text{epi}(f)$ and converges to (x, w), so by the closure of $\text{epi}(f)$, we must have $f(x) \leq w$. Thus, f is lower semicontinuous at x.

We next show that (i) implies (iii), and that (iii) implies (ii). Assume that $\text{epi}(f)$ is closed and let $\{x_k\}$ be a sequence that converges to some x and belongs to the level set $\big\{z \mid f(z) \leq \gamma\big\}$, where γ is a scalar. Then $(x_k, \gamma) \in \text{epi}(f)$ for all k, and by closure of $\text{epi}(f)$, we have $\big(f(x), \gamma\big) \in \text{epi}(f)$. Hence x belongs to the level set $\big\{x \mid f(x) \leq \gamma\big\}$, implying that this set is closed. Therefore (i) implies (iii).

Finally, assume that the level sets $\big\{x \mid f(x) \leq \gamma\big\}$ are closed, fix an x, and let $\{x_k\}$ be a sequence converging to x. If $\liminf_{k\to\infty} f(x_k) < \infty$, then for each γ with $\liminf_{k\to\infty} f(x_k) < \gamma$ and all sufficiently large k, we have $f(x_k) < \gamma$. From the closure of the level sets $\big\{x \mid f(x) \leq \gamma\big\}$, it follows that x belongs to all the levels with $\liminf_{k\to\infty} f(x_k) < \gamma$, implying that $f(x) \leq \liminf_{k\to\infty} f(x_k)$, and that f is lower semicontinuous at x. Therefore, (iii) implies (ii). **Q.E.D.**

If the epigraph of a function $f : X \mapsto (-\infty, \infty]$ is a closed set, we say that f is a *closed* function. Thus, if we extend the domain of f to $\Re^n$ and consider the function $\tilde{f}$ given by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in X, \\ \infty & \text{if } x \notin X, \end{cases}$$

we see that according to the preceding proposition, *f is closed if and only $\tilde{f}$ is lower semicontinuous over $\Re^n$*. Note that if f is lower semicontinuous over $\text{dom}(f)$, it is not necessarily closed; take for example f to be constant for x in some nonclosed set and ∞ otherwise. Furthermore, if f is closed it is not necessarily true that $\text{dom}(f)$ is closed; for example, the function

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x > 0, \\ \infty & \text{otherwise}, \end{cases}$$

is closed but $\text{dom}(f)$ is the open half-line of positive numbers. On the other hand, *if* $\text{dom}(f)$ *is closed and f is lower semicontinuous over* $\text{dom}(f)$*, then f is closed* because $\text{epi}(f)$ is closed, as can be seen by reviewing the relevant part of the proof of Prop. 0.2.2(b).

Common examples of convex functions are affine functions and norms; this is straightforward to verify, using the definition of convexity. For example, for any $x, y \in \Re^n$ and any $\alpha \in [0, 1]$, we have by using the triangle inequality,

$$\|\alpha x + (1 - \alpha)y\| \leq \|\alpha x\| + \|(1 - \alpha)y\| = \alpha\|x\| + (1 - \alpha)\|y\|,$$

so the norm function $\|\cdot\|$ is convex. The following proposition provides some means for recognizing convex functions, and gives some algebraic operations that preserve convexity of a function.

Proposition 0.2.3:

(a) Let $f_1, \ldots, f_m : \Re^n \mapsto (-\infty, \infty]$ be given functions, let $\lambda_1, \ldots, \lambda_m$ be positive scalars, and consider the function $g : \Re^n \mapsto (-\infty, \infty]$ given by

$$g(x) = \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x).$$

If $f_1, \ldots, f_m$ are convex, then g is also convex, while if $f_1, \ldots, f_m$ are closed, then g is also closed.

(b) Let $f : \Re^n \mapsto (-\infty, \infty]$ be a given function, let A be an $m \times n$ matrix, and consider the function $g : \Re^n \mapsto (-\infty, \infty]$ given by

$$g(x) = f(Ax).$$

If f is convex, then g is also convex, while if f is closed, then g is also closed.

(c) Let $f_i : \Re^n \mapsto (-\infty, \infty]$ be given functions for $i \in I$, where I is an index set, and consider the function $g : \Re^n \mapsto (-\infty, \infty]$ given by

$$g(x) = \sup_{i \in I} f_i(x).$$

If f_i, $i \in I$, are convex, then g is also convex, while if f_i, $i \in I$, are closed, then g is also closed.

Proof: (a) Let $f_1, \ldots, f_m$ be convex. We use the definition of convexity to write for any $x, y \in \Re^n$ and $\alpha \in [0, 1]$,

$$\begin{aligned} f\big(\alpha x + (1-\alpha)y\big) &= \sum_{i=1}^m \lambda_i f_i\big(\alpha x + (1-\alpha)y\big) \\ &\le \sum_{i=1}^m \lambda_i \big(\alpha f_i(x) + (1-\alpha) f_i(y)\big) \\ &= \alpha \sum_{i=1}^m \lambda_i f_i(x) + (1-\alpha) \sum_{i=1}^m \lambda_i f_i(y) \\ &= \alpha f(x) + (1-\alpha) f(y). \end{aligned}$$

Hence f is convex.

Let $f_1, \ldots, f_m$ be closed. Then the f_i are lower semicontinuous at every $x \in \Re^n$ [cf. Prop. 0.2.2(b)], so for every sequence $\{x_k\}$ converging to x, we have $f_i(x) \le \liminf_{k \to \infty} f_i(x_k)$ for all i. Hence

$$g(x) \le \sum_{i=1}^m \lambda_i \liminf_{k \to \infty} f_i(x_k) \le \liminf_{k \to \infty} \sum_{i=1}^m \lambda_i f_i(x_k) = \liminf_{k \to \infty} g(x_k).$$

where we have used Prop. 0.1.4(d) (the sum of the limit inferiors of sequences is less or equal to the limit inferior of the sum sequence). Therefore, g is lower semicontinuous at all $x \in \Re^n$, so by Prop. 0.2.2(b), it is closed.

(b) This is straightforward, along the lines of the proof of part (a).

(c) A pair (x, w) belongs to the epigraph

$$\mathrm{epi}(g) = \big\{(x, w) \mid g(x) \leq w\big\}$$

if and only if $f_i(x) \leq w$ for all $i \in I$, or $(x, w) \in \cap_{i \in I} \mathrm{epi}(f_i)$. Therefore,

$$\mathrm{epi}(g) = \cap_{i \in I} \mathrm{epi}(f_i).$$

If the f_i are convex, the epigraphs $\mathrm{epi}(f_i)$ are convex, so $\mathrm{epi}(g)$ is convex, and g is convex. If the f_i are closed, then the epigraphs $\mathrm{epi}(f_i)$ are closed, so $\mathrm{epi}(g)$ is closed, and g is closed. **Q.E.D.**

Characterizations of Differentiable Convex Functions

For differentiable functions, there is an alternative characterization of convexity, given in the following proposition.

Proposition 0.2.4: Let $C \subset \Re^n$ be a convex set and let $f : \Re^n \mapsto \Re$ be differentiable over $\Re^n$.

(a) f is convex over C if and only if

$$f(z) \geq f(x) + (z - x)'\nabla f(x), \qquad \forall\, x, z \in C. \tag{0.2}$$

(b) f is strictly convex over C if and only if the above inequality is strict whenever $x \neq z$.

Proof: We prove (a) and (b) simultaneously. Assume that the inequality (0.2) holds. Choose any $x, y \in C$ and $\alpha \in [0, 1]$, and let $z = \alpha x + (1 - \alpha)y$. Using the inequality (0.2) twice, we obtain

$$f(x) \geq f(z) + (x - z)'\nabla f(z),$$

$$f(y) \geq f(z) + (y - z)'\nabla f(z).$$

We multiply the first inequality by α, the second by $(1 - \alpha)$, and add them to obtain

$$\alpha f(x) + (1 - \alpha)f(y) \geq f(z) + \big(\alpha x + (1 - \alpha)y - z\big)'\nabla f(z) = f(z),$$

which proves that f is convex. If the inequality (0.2) is strict as stated in part (b), then if we take $x \neq y$ and $\alpha \in (0,1)$ above, the three preceding inequalities become strict, thus showing the strict convexity of f.

Conversely, assume that f is convex, let x and z be any vectors in C with $x \neq z$, and for $\alpha \in (0,1)$, consider the function

$$g(\alpha) = \frac{f\big(x + \alpha(z-x)\big) - f(x)}{\alpha}, \qquad \alpha \in (0,1].$$

We will show that $g(\alpha)$ is monotonically decreasing with α, and is strictly monotonically decreasing if f is strictly convex. This will imply that

$$(z-x)'\nabla f(x) = \lim_{\alpha \downarrow 0} g(a) \leq g(1) = f(z) - f(x),$$

with strict inequality if g is strictly monotonically decreasing, thereby showing that the desired inequality (0.2) holds (and holds strictly if f is strictly convex). Indeed, consider any α_1, α_2, with $0 < \alpha_1 < \alpha_2 < 1$, and let

$$\overline{\alpha} = \frac{\alpha_1}{\alpha_2}, \qquad \overline{z} = x + \alpha_2(z-x). \tag{0.3}$$

We have

$$f\big(x + \overline{\alpha}(\overline{z} - x)\big) \leq \overline{\alpha} f(\overline{z}) + (1-\overline{\alpha}) f(x),$$

or

$$\frac{f\big(x + \overline{\alpha}(\overline{z} - x)\big) - f(x)}{\overline{\alpha}} \leq f(\overline{z}) - f(x), \tag{0.4}$$

and the above inequalities are strict if f is strictly convex. Substituting the definitions (0.3) in Eq. (0.4), we obtain after a straightforward calculation

$$\frac{f\big(x + \alpha_1(z-x)\big) - f(x)}{\alpha_1} \leq \frac{f\big(x + \alpha_2(z-x)\big) - f(x)}{\alpha_2},$$

or

$$g(\alpha_1) \leq g(\alpha_2),$$

with strict inequality if f is strictly convex. Hence g is monotonically decreasing with α, and strictly so if f is strictly convex. **Q.E.D.**

Note a simple consequence of Prop. 0.2.4(a): if $f : \Re^n \mapsto \Re$ is a convex function and $\nabla f(x^*) = 0$, then x^* minimizes f over $\Re^n$. This is a classical sufficient condition for unconstrained optimality, originally formulated (in one dimension) by Fermat in 1637.

For twice differentiable convex functions, there is another characterization of convexity as shown by the following proposition.

Proposition 0.2.5: Let $C \subset \Re^n$ be a convex set and let $f : \Re^n \mapsto \Re$ be twice continuously differentiable over $\Re^n$.

(a) If $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$, then f is convex over C.

(b) If $\nabla^2 f(x)$ is positive definite for all $x \in C$, then f is strictly convex over C.

(c) If $C = \Re^n$ and f is convex, then $\nabla^2 f(x)$ is positive semidefinite for all x.

Proof: (a) For all $x, y \in C$ we have

$$f(y) = f(x) + (y - x)'\nabla f(x) + \tfrac{1}{2}(y - x)'\nabla^2 f\big(x + \alpha(y - x)\big)(y - x)$$

for some $\alpha \in [0, 1]$. Therefore, using the positive semidefiniteness of $\nabla^2 f$, we obtain

$$f(y) \geq f(x) + (y - x)'\nabla f(x), \qquad \forall\, x, y \in C.$$

From Prop. 0.2.4(a), we conclude that f is convex.

(b) Similar to the proof of part (a), we have $f(y) > f(x) + (y - x)'\nabla f(x)$ for all $x, y \in C$ with $x \neq y$, and the result follows from Prop. 0.2.4(b).

(c) Suppose that $f : \Re^n \mapsto \Re$ is convex and suppose, to obtain a contradiction, that there exist some $x \in \Re^n$ and some $z \in \Re^n$ such that $z'\nabla^2 f(x)z < 0$. Using the continuity of $\nabla^2 f$, we see that we can choose the norm of z to be small enough so that $z'\nabla^2 f(x + \alpha z)z < 0$ for every $\alpha \in [0, 1]$. Then, we obtain $f(x + z) < f(x) + z'\nabla f(x)$, which, in view of Prop. 0.2.4(a), contradicts the convexity of f. **Q.E.D.**

If f is convex over a strict subset $C \subset \Re^n$, it is not necessarily true that $\nabla^2 f(x)$ is positive semidefinite at any point of C [take for example $n = 2$, $C = \{(x_1, 0) \mid x_1 \in \Re\}$, and $f(x) = x_1^2 - x_2^2$]. A strengthened version of Prop. 0.2.5 is given in the exercises. It can be shown that the conclusion of Prop. 0.2.5(c) also holds if C is assumed to have nonempty interior instead of being equal to $\Re^n$.

0.2.2 Convex and Affine Hulls

Let X be a subset of $\Re^n$. A *convex combination* of elements of X is a vector of the form $\sum_{i=1}^m \alpha_i x_i$, where m is a positive integer, $x_1, \ldots, x_m$ belong to X, and $\alpha_1, \ldots, \alpha_m$ are scalars such that

$$\alpha_i \geq 0, \quad i = 1, \ldots, m, \qquad \sum_{i=1}^m \alpha_i = 1.$$

Note that if X is convex, then the convex combination $\sum_{i=1}^m \alpha_i x_i$ belongs to X, and for any function $f : \Re^n \mapsto \Re$ that is convex over X, we have

$$f\left(\sum_{i=1}^m \alpha_i x_i\right) \le \sum_{i=1}^m \alpha_i f(x_i). \tag{0.5}$$

This follows by using repeatedly the definition of convexity. The preceding relation is a special case of *Jensen's inequality* and can be used to prove a number of interesting inequalities in applied mathematics and probability theory.

The *convex hull* of a set X, denoted $\mathrm{conv}(X)$, is the intersection of all convex sets containing X, and is a convex set by Prop. 0.2.1(a). It is straightforward to verify that the set of all convex combinations of elements of X is convex, and is equal to the convex hull $\mathrm{conv}(X)$ (see the exercises). In particular, if X consists of a finite number of vectors $x_1, \ldots, x_m$, its convex hull is

$$\mathrm{conv}\big(\{x_1, \ldots, x_m\}\big) = \left\{ \sum_{i=1}^m \alpha_i x_i \;\middle|\; \alpha_i \ge 0,\, i = 1, \ldots, m,\, \sum_{i=1}^m \alpha_i = 1 \right\}.$$

We recall that an affine set M is a set of the form $x + S$, where S is a subspace, called the *subspace parallel to* M. If X is a subset of $\Re^n$, the *affine hull* of X, denoted $\mathrm{aff}(X)$, is the intersection of all affine sets containing X. Note that $\mathrm{aff}(X)$ is itself an affine set and that it contains $\mathrm{conv}(X)$. It can be seen that the affine hull of X, the affine hull of the convex hull $\mathrm{conv}(X)$, and the affine hull of the closure $\mathrm{cl}(X)$ coincide (see the exercises). For a convex set C, the *dimension* of C is defined to be the dimension of $\mathrm{aff}(C)$.

Given a subset $X \subset \Re^n$, a *nonnegative combination* of elements of X is a vector of the form $\sum_{i=1}^m \alpha_i x_i$, where m is a positive integer, $x_1, \ldots, x_m$ belong to X, and $\alpha_1, \ldots, \alpha_m$ are nonnegative scalars. If the scalars α_i are all positive, the combination $\sum_{i=1}^m \alpha_i x_i$ is said to be *positive*. The *cone generated by* X, denoted by $\mathrm{cone}(X)$, is the set of nonnegative combinations of elements of X. It is easily seen that $\mathrm{cone}(X)$ is a convex cone, although it need not be closed [$\mathrm{cone}(X)$ can be shown to be closed in special cases, such as when X is a finite set – this is one of the central results of polyhedral convexity].

The following is a fundamental characterization of convex hulls.

Proposition 0.2.6: (Caratheodory's Theorem) Let X be a subset of $\Re^n$.

(a) Every x in cone(X) can be represented as a positive combination of vectors $x_1, \ldots, x_m \in X$ that are linearly independent, where m is a positive integer with $m \leq n$.

(b) Every x in conv(X) can be represented as a convex combination of vectors $x_1, \ldots, x_m \in X$ such that $x_2 - x_1, \ldots, x_m - x_1$ are linearly independent, where m is a positive integer with $m \leq n+1$.

Proof: (a) Let x be a nonzero vector in the cone(X), and let m be the smallest integer such that x has the form $\sum_{i=1}^m \alpha_i x_i$, where $\alpha_i > 0$ and $x_i \in X$ for all $i = 1, \ldots, m$. If the vectors x_i were linearly dependent, there would exist scalars $\lambda_1, \ldots, \lambda_m$, with $\sum_{i=1}^m \lambda_i x_i = 0$ and at least one of the λ_i is positive. Consider the linear combination $\sum_{i=1}^m (\alpha_i - \overline{\gamma}\lambda_i) x_i$, where $\overline{\gamma}$ is the largest γ such that $\alpha_i - \gamma\lambda_i \geq 0$ for all i. This combination provides a representation of x as a positive combination of fewer than m vectors of X – a contradiction. Since any linearly independent set of vectors contains at most n elements, we must have $m \leq n$.

(b) The proof will be obtained by applying part (a) to the subset of $\Re^{n+1}$ given by

$$Y = \big\{(x, 1) \mid x \in X\big\}.$$

Indeed, if $x \in \text{conv}(X)$, then $x = \sum_{i=1}^m \alpha_i x_i$ for some positive α_i with $1 = \sum_{i=1}^m \alpha_i$, or equivalently, $(x, 1) \in \text{cone}(Y)$. By part (a), this is true if for some positive $\alpha_1, \ldots, \alpha_m$ and vectors $(x_1, 1), \ldots, (x_m, 1)$, which are linearly independent (implying that $m \leq n+1$) we have $(x, 1) = \sum_{i=1}^m \alpha_i (x_i, 1)$, or equivalently

$$x = \sum_{i=1}^m \alpha_i x_i, \qquad 1 = \sum_{i=1}^m \alpha_i.$$

Finally, to show that $x_2 - x_1, \ldots, x_m - x_1$ are linearly independent, assume to arrive at a contradiction, that there exist $\lambda_2, \ldots, \lambda_m$, not all 0, such that

$$\sum_{i=2}^m \lambda_i (x_i - x_1) = 0$$

Equivalently, defining $\lambda_1 = -(\lambda_2 + \cdots + \lambda_m)$, we have

$$\sum_{i=1}^m \lambda_i (x_i, 1) = 0,$$

which contradicts the linear independence of $(x_1, 1), \ldots, (x_m, 1)$. **Q.E.D.**

It is not generally true that the convex hull of a closed set is closed [take for instance the convex hull of the set consisting of the origin and the subset $\{(x_1, x_2) \mid x_1 x_2 = 1,\ x_1 \geq 0,\ x_2 \geq 0\}$ of $\Re^2$]. We have, however, the following.

Proposition 0.2.7: The convex hull of a compact set is compact.

Proof: Let X be a compact subset of $\Re^n$. By Caratheodory's Theorem, a sequence in $\text{conv}(X)$ can be expressed as $\left\{\sum_{i=1}^{n+1} \alpha_i^k x_i^k\right\}$, where for all k and i, $\alpha_i^k \geq 0$, $x_i^k \in X$, and $\sum_{i=1}^{n+1} \alpha_i^k = 1$. Since the sequence

$$\left\{(\alpha_1^k, \ldots, \alpha_{n+1}^k, x_1^k, \ldots, x_{n+1}^k)\right\}$$

belongs to a compact set, it has a limit point $\left\{(\alpha_1, \ldots, \alpha_{n+1}, x_1, \ldots, x_{n+1})\right\}$ such that $\sum_{i=1}^{n+1} \alpha_i = 1$, and for all i, $\alpha_i \geq 0$, and $x_i \in X$. Thus, the vector $\sum_{i=1}^{n+1} \alpha_i x_i$, which belongs to $\text{conv}(X)$, is a limit point of the sequence $\left\{\sum_{i=1}^{n+1} \alpha_i^k x_i^k\right\}$, showing that $\text{conv}(X)$ is compact. **Q.E.D.**

E X E R C I S E S

0.2.1

Let C be a nonempty subset of $\Re^n$, and let λ_1 and λ_2 be positive scalars. Show by example that the sets $(\lambda_1 + \lambda_2)C$ and $\lambda_1 C + \lambda_2 C$ may differ when C is not convex [cf. Prop. 0.2.1(c)].

0.2.2 (Properties of Cones)

(a) For any collection $\{C_i \mid i \in I\}$ of cones, the intersection $\cap_{i \in I} C_i$ is a cone.

(b) The Cartesian product $C_1 \times C_2$ of two cones C_1 and C_2 is a cone.

(c) The vector sum $C_1 + C_2$ of two cones C_1 and C_2 is a cone.

(d) The closure of a cone is a cone.

(e) The image and the inverse image of a cone under a linear transformation is a cone.

0.2.3 (Lower Semicontinuity under Composition)

(a) Let $f : \Re^n \mapsto \Re^m$ be a continuous function and $g : \Re^m \mapsto \Re$ be a lower semicontinuous function. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is lower semicontinuous.

(b) Let $f : \Re^n \mapsto \Re$ be a lower semicontinuous function, and $g : \Re \mapsto \Re$ be a lower semicontinuous and monotonically nondecreasing function. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is lower semicontinuous. Give an example showing that the assumption that g is monotonically nondecreasing is essential.

0.2.4 (Convexity under Composition)

(a) Let C be a convex subset of $\Re^n$, $f : C \mapsto \Re$ be a convex function, and $g : \Re \mapsto \Re$ be a convex monotonically nondecreasing function. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is convex over C. In addition, if g is monotonically increasing and f is strictly convex, then h is strictly convex.

(b) Let $f = (f_1, \ldots, f_m)$ where each $f_i : \Re^n \mapsto \Re$ is a convex function, and let $g : \Re^m \mapsto \Re$ be a function that is convex and monotonically nondecreasing in the sense that for all $u, \overline{u} \in \Re^m$ such that $u \leq \overline{u}$, we have $g(u) \leq g(\overline{u})$. Show that the function h defined by $h(x) = g\big(f(x)\big)$ is convex.

0.2.5 (Examples of Convex Functions)

Show that the following functions from $\Re^n$ to $(-\infty, \infty]$ are convex:

(a)

$$f_1(x_1, \ldots, x_n) = \begin{cases} -(x_1 x_2 \cdots x_n)^{\frac{1}{n}}, & \text{if } x_1 > 0, \ldots, x_n > 0, \\ \infty & \text{otherwise.} \end{cases}$$

(b) $f_2(x) = \ln\big(e^{x_1} + \cdots + e^{x_n}\big)$.

(c) $f_3(x) = \|x\|^p$ with $p \geq 1$.

(d) $f_4(x) = \frac{1}{f(x)}$ where $f : \Re^n \mapsto \Re$ is a concave function and $f(x) > 0$ for all x.

(e) $f_5(x) = \alpha f(x) + \beta$ where $f : \Re^n \mapsto \Re$ is a convex function, and α and β are scalars such that $\alpha \geq 0$.

(f) $f_6(x) = e^{\beta x' A x}$, where A is an $n \times n$ positive semidefinite symmetric matrix and β is a positive scalar.

(g) $f_7(x) = f(Ax + b)$, where $f : \Re^m \mapsto \Re$ is a convex function, A is an $m \times n$ matrix, and b is a vector in $\Re^m$.

0.2.6 (Posynomials)

A *posynomial* of scalar variables $y_1, \ldots, y_n$ is a function of the form

$$g(y_1, \ldots, y_n) = \sum_{i=1}^{m} \beta_i y_1^{a_{i1}} \cdots y_n^{a_{in}},$$

where the scalars y_i and β_i are positive, and the exponents a_{ij} are real numbers. Show the following:

(a) A posynomial need not be convex.

(b) By a logarithmic change of variables, where we set

$$f(x) = \ln\big(g(y_1, \ldots, y_n)\big), \qquad b_i = \ln \beta_i,\ \forall\ i, \qquad x_j = \ln y_j,\ \forall\ j,$$

we obtain a convex function

$$f(x) = \ln \exp(Ax + b), \qquad \forall\ x \in \Re^n,$$

where $\exp(z) = e^{z_1} + \cdots + e^{z_m}$ for all $z \in \Re^m$, A is an $m \times n$ matrix with entries a_{ij}, and $b \in \Re^m$ is a vector with components b_i.

(c) Every function $g : \Re^n \mapsto \Re$ of the form

$$g(y) = g_1(y)^{\gamma_1} \cdots g_r(y)^{\gamma_r},$$

where g_k is a posynomial and $\gamma_k > 0$ for all k, can be transformed by a logarithmic change of variables into a convex function f given by

$$f(x) = \sum_{k=1}^{r} \gamma_k \ln \exp(A_k x + b_k),$$

with the matrix A_k and the vector b_k being associated with the posynomial g_k for each k.

0.2.7 (Arithmetic-Geometric Mean Inequality)

Show that if $\alpha_1, \ldots, \alpha_n$ are positive scalars with $\sum_{i=1}^{n} \alpha_i = 1$, then for every set of positive scalars $x_1, \ldots, x_n$, we have

$$x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \le \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n,$$

with equality if and only if $x_1 = x_2 = \cdots = x_n$. *Hint*: Show that $-\ln x$ is a strictly convex function on $(0, \infty)$.

0.2.8 (Convex Hulls, Affine Hulls, and Generated Cones)

Let X be a nonempty set.

(a) Show that X, $\text{conv}(X)$, and $\text{cl}(X)$ have the same affine hull.

(b) Show that $\text{cone}(X) = \text{cone}\big(\text{conv}(X)\big)$.

(c) Show that $\text{aff}\big(\text{conv}(X)\big) \subset \text{aff}\big(\text{cone}(X)\big)$. Give an example in which the inclusion is strict, i.e., $\text{aff}\big(\text{conv}(X)\big)$ is a strict subset of $\text{aff}\big(\text{cone}(X)\big)$.

(d) Assuming that the origin belongs to $\text{conv}(X)$, show that $\text{aff}\big(\text{conv}(X)\big) = \text{aff}\big(\text{cone}(X)\big)$.

0.2.9

Let $\{f_i \mid i \in I\}$ be an arbitrary collection of proper convex functions $f_i : \Re^n \mapsto (-\infty, \infty]$. Define

$$f(x) = \inf\big\{w \mid (x,w) \in \text{conv}\big(\cup_{i\in I}\text{epi}(f_i)\big)\big\}, \qquad \forall\, x \in \Re^n.$$

Show that $f(x)$ is given by

$$f(x) = \inf\bigg\{\sum_{i\in\bar{I}} \alpha_i f_i(x_i) \;\bigg|\; \sum_{i\in\bar{I}} \alpha_i x_i = x,\ x_i \in \Re^n,\ \sum_{i\in\bar{I}} \alpha_i = 1,\ \alpha_i \geq 0,\ \forall\, i \in \bar{I},\ \bar{I} \subset I,\ \bar{I}: \text{ finite}\bigg\},$$

where the infimum is taken over all representations of x as a convex combination of elements x_i, such that only finitely many coefficients α_i are nonzero.

0.2.10 (Convexification of Nonconvex Functions)

Let X be a nonempty subset of $\Re^n$ and let $f : X \mapsto \Re$ be a function that is bounded below over X. Define the function $F : \text{conv}(X) \mapsto \Re$ by

$$F(x) = \inf\big\{w \mid (x,w) \in \text{conv}\big(\text{epi}(f)\big)\big\}.$$

(a) Show that F is convex over $\text{conv}(X)$ and it is given by

$$F(x) = \inf\bigg\{\sum_i \alpha_i f(x_i) \;\bigg|\; \sum_i \alpha_i x_i = x,\ x_i \in X,\ \sum_i \alpha_i = 1,\ \alpha_i \geq 0,\ \forall\, i\bigg\},$$

where the infimum is taken over all representations of x as a convex combination of elements of X, such that only finitely many coefficients α_i are nonzero.

(b) Show that

$$\inf_{x\in\text{conv}(X)} F(x) = \inf_{x\in X} f(x).$$

(c) Show that every global minimum of f over X is also a global minimum of F over $\text{conv}(X)$.

0.2.11 (Minimization of Linear Functions)

Show that minimization of a linear function over a set is equivalent to minimization over its convex hull. In particular, if $X \subset \Re^n$ and $c \in \Re^n$, then

$$\inf_{x \in \mathrm{conv}(X)} c'x = \inf_{x \in X} c'x.$$

Furthermore, the infimum in the left-hand side above is attained if and only if the infimum in the right-hand side is attained.

0.2.12 (Properties of Cartesian Products)

Given sets $X_i \subset \Re^{n_i}$, $i = 1, \ldots, m$, let $X = X_1 \times \cdots \times X_m$ be their Cartesian product.

(a) Show that the convex hull (closure, affine hull) of X is equal to the Cartesian product of the convex hulls (closures, affine hulls, respectively) of the X_i.

(b) Assuming that all the sets $X_1, \ldots, X_m$ contain the origin, show that $\mathrm{cone}(X) = \mathrm{cone}(X_1) \times \cdots \times \mathrm{cone}(X_m)$. Show that the result fails if one of the sets does not contain the origin.

(c) Assuming that all the sets $X_1, \ldots, X_m$ are convex, show that the relative interior (recession cone) of X is equal to the Cartesian product of the relative interiors (recession cones) of the X_i.