Lectures on Finite Dimensional Optimization Theory.

for M6361; Applicable Analysis, Spring 2010.

Giles Auchmuty
Department of Mathematics,
University of Houston.
auchmuty@uh.edu September 14, 2010

Table of Contents.

Introduction

These notes assume that you have had courses in metric space topology and linear algebra. Please review your previous material on metric spaces, especially topics regarding compactness and convergence. Also the geometry and analysis of $\mathbb{R}^n$. A good reference for the analysis background is the calculus text of Fleming [3]; a thorough treatment of advanced multivariate calculus is in Ortega and Rheinboldt [5]. The linear algebra background is covered very nicely in the text of Strang [7]. A thorough coverage of much of this material may be found in Ortega and Rheinboldt [5].

Much of the material covered in these notes can be found in Berkovitz [1] in a different order and with a somewhat different emphasis. For a more computational view of this material see the book [2] on Convex Optimaization by Boyd and Vanderberghe. A nice introduction to the geometry of convex sets is in the book [4] by Lay. Both of these last two texts treat a lot of different applications of these theories.

Two conventions that will be used are that when a term is defined the term is italicized. The expression " $:=$ " means that the left hand side of the formula is defined to be equal to the right hand side. Occasionally " $=:$ " is used and it means that the right hand expression is being defined.

1. Metric spaces, infs and sups

This semester we will mostly work with functions and operators defined on metric spaces. Here I'll collect the basic definitions and notation that will be used. If you've done more analysis, you may have seen more general definitions than the one's given here - and the definitions given below sometimes are theorems associated with more general definitions in topology.

Let X be a nonempty set and (X, d) be a metric space. An element of X will be called a point in X; a sequence in X is a set that is indexed by the set of natural numbers $\mathbb{N} := \{1, 2, 3, \ldots\}$. Note that any nonempty subset of a metric space is again a metric space.

The open ball $B_r(x)$ of radius r about a point $x \in X$ is the set

$$B_r(x) \ := \ \{y \in X : d(y, x) < r\}$$

The closed ball $\mathbb{B}_r(x)$ has $\leq$ in place of $<$ in the preceding definition. The sphere of radius r, center x is

$$S_r(x) \ := \ \{y \in X : d(y, x) = r\}$$

A subset U of X is said to be *open* if for each $x \in U$, there is an $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U$. When U is an open set and $x \in U$ then U is said to be a *neighborhood* of x. A subset U of $\mathbb{R}^n$ is *closed* if $\mathbb{R}^n \setminus U \ := \ \{x \in \mathbb{R}^n : x \notin U\}$ is open.

A sequence $\Gamma := \{x_k : k \geq 1\}$ *converges* to an element x of X provided

$$\lim_{k \to \infty} d(x_k, x) \ = \ 0.$$

In this case we say that x is the *limit* of the sequence Γ. To prove that this holds one generally shows that

$$\text{Given } \epsilon > 0, \text{ there is an } K(\epsilon) \text{ such that } k \geq K(\epsilon) \Rightarrow \ d(x_k, x) < \epsilon. \tag{1.1}$$

1.1. **Completeness and Compact Sets.** A sequence $\Gamma := \{x_k : k \geq 1\}$ is a *Cauchy sequence* in X provided

$$\text{Given } \epsilon > 0, \text{there is an } K(\epsilon) \text{ such that } m, k \geq K(\epsilon) \ \Rightarrow \ d(x_k, x_m) < \epsilon.$$

A metric space (X,d) is *complete* provided any Cauchy sequence in X has a limit in X. That is, when $\Gamma := \{x_k : k \geq 1\}$ is a Cauchy sequence in X then there is an $\hat{x} \in X$ such that $\lim_{k\to\infty} x_k = \hat{x}$.

A subset W of X is an $\epsilon-$net for X provided $X \ = \ \cup_{y \in W} B_\epsilon(y)$.

A subset K of X is *totally bounded* provided for each $\epsilon > 0$ there is a finite $\epsilon-$net of K.

A metric space (X,d) is said to be *compact* provided it is complete and totally bounded.

This may not be the definition of compactness that you have seen before; but it is the criterion that is commonly used in proving that a specific set is compact. Try proving that the unit interval [0,1] is compact using another defintion and compare the proof with one that verifies this criterion.

This definition of compactness implies immediately that a closed, totally bounded, subset of a complete metric space is compact. Also a compact metric space is complete. A basic result in metric space topology is that (X,d) is compact if and only if every sequence in X has a subsequence that converges to an element in X. A basic question in optimization theory is when does a sequence of points in X converge to an optimizer of a problem - so we will often need some compactness properties in the problem.

1.2. **The Extended Real Line.** The *extended real line* is the set

$$\overline{\mathbb{R}} \ := \ \mathbb{R} \cup \{\pm\infty\} \ = \ [-\infty, \infty]. \tag{1.2}$$

It is a totally ordered set with the usual order on $\mathbb{R}$ together with $-\infty < x < \infty$ for each real number $x \in \mathbb{R}$. The elements of $\mathbb{R}$ are called *real numbers* and an element of $\overline{\mathbb{R}}$ is a *number.*

$\overline{\mathbb{R}}$ is a metric space with respect to the metric

$$d_e(x, y) \ := \ \left| \int_x^y \frac{dt}{1+t^2} \right| \ = \ |\arctan(y) - \arctan(x)|. \tag{1.3}$$

The proof that this is a metric is straightforward. (Please verify this).

Convergence on $\mathbb{R}$ may be compared with with convergence on $\overline{\mathbb{R}}$. A basic result is the following

Theorem 1.1. *$\overline{\mathbb{R}}$ with the metric (1.3) is a compact metric space. When $\{x_k : k \geq 1\}$ is a sequence of points in $\mathbb{R}$ which converges to $\hat{x}$ in $\mathbb{R}$, then it also converges to $\hat{x}$ in $\overline{\mathbb{R}}$.*

Proof. The proof that this space is compact is an exercise in metric space topology. You should show that $\overline{\mathbb{R}}$ is complete and totally bounded by verifying each of these properties. The proof that the sequence converges wrt the metric d_e follows as

$$d_e(\hat{x}, x_k) \ := \ \left| \int_{x_k}^{\hat{x}} \frac{dt}{1+t^2} \right| \ \leq \ |\hat{x} - x_k| \quad \text{as} \quad \frac{1}{1+t^2} \leq 1 \quad \text{for all t.}$$

$\square$

Note that while $\mathbb{R}$ is complete - but not compact, $\overline{\mathbb{R}}$ is both complete and compact. The following example shows that a sequence may converge in $\overline{\mathbb{R}}$ but not in $\mathbb{R}$.

Example 1.1. The sequence $\{1, 2, 3, ..., n, ...\}$ does not converge in $\mathbb{R}$. It converges to ∞ in $\overline{\mathbb{R}}$ since

$$d_e(n, \infty) = \int_n^{\infty} \frac{dt}{1+t^2} = \frac{\pi}{2} - \arctan \ n$$

Given $\epsilon > 0$, choose an $N_\epsilon > \tan{(\pi - 2\epsilon)}/2$. Then for $n > N_\epsilon$, $d_e(n, \infty) < d_e(N_\epsilon, \infty) < \epsilon$. So $d_e(n, \infty) \to 0$ as $n \to \infty$ in $\mathbb{R}$.

The following is a special case of the general theorem that a subset of a compact metric space is compact if and only if it is closed.

Corollary 1.2. *A subset K of $\overline{\mathbb{R}} \ := \ [-\infty, \infty]$ is compact if and only if K is a closed subset of $\overline{\mathbb{R}}$.*

Example 1.2: $[1, \infty]$ is compact in $\overline{\mathbb{R}}$. However the sequence of example 1.1 is a subset of $\mathbb{R}$, that does not have a convergent subsequence.

1.3. **Algebra of $\overline{\mathbb{R}}$.** The usual definition of scalar multiplication may be extended to products of $\pm\infty$ and a real number c by

$$c \ .(\infty) \ := \ \begin{cases} \infty & \text{if} \quad c > 0 \\ 0 & \text{if} \quad c = 0 \\ -\infty & \text{if} \quad c < 0. \end{cases} \tag{1.4}$$

and $c.(-\infty) \ := \ (-c).(\infty)$. These formulae also hold for $c = \pm\infty$ - so multiplication is defined for any two elements of $\overline{\mathbb{R}}$.

Addition is given by the following conventions. For $c \in \mathbb{R}$.

(i.) $c + \infty \ := \ \infty, \qquad \infty + \infty \ := \ \infty.$

(ii.) $c + (-\infty) \ := \ -\infty, \qquad -\infty + (-\infty) \ := \ -\infty.$

(iii.) The expression $\infty + (-\infty)$ is not defined.

In view of (iii), $\overline{\mathbb{R}}$ is not a vector space. Since addition is not always defined on $\overline{\mathbb{R}}$, the distributive laws for $c.(a + b)$ do not extend to $\overline{\mathbb{R}}$.

When A is a subset of $\overline{\mathbb{R}}$, then the set cA is defined by $cA := \{cx \ : \ x \in A\}$. In particular, $-A := \{-x : \ x \in A\}$. When A, B are subsets of $\mathbb{R}$, then the set A+B is defined by $A + B := \{x + y : \ x \in A, \ y \in B\}$. Similarly $A - B := \{x - y : \ x \in A, \ y \in B\}$. The set theoretic difference will be denoted $A \setminus B := \{x : \ x \in A, \ \& \ x \notin B\}$.

The following special notations will be used in these notes. $\mathbb{R}^{\#} := (-\infty, \infty]$ $\mathbb{R}_+ := [0, \infty)$ and $\mathbb{R}_{++} := (0, \infty)$.

Exercises. .

Exercise 1.1. Show that multiplication on $\overline{\mathbb{R}}$ is commutative and associative. Is $\overline{\mathbb{R}}$ a commutative group with respect to this operation?

Exercise 1.2 Define the function $m : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ by $m(x, y) := x \cdot y$.
(i) Show that $m(., y)$ is a continuous function on $\overline{\mathbb{R}}$ for each $y \in \overline{\mathbb{R}}$.
(ii) Show that there are points $(\hat{x}, \hat{y}) \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$ at which m is not continuous. Construct sequences that converge to such a $(\hat{x}, \hat{y})$ (and do not contain the limit point) that have a limit other than $m(\hat{x}, \hat{y})$.
(iii) Find all the points in $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ at which m is not continuous.

Exercise 1.3: Show that $\mathbb{R}^{\#}$ is an open subset of $\overline{\mathbb{R}}$. Define $a : \mathbb{R}^{\#} \times \mathbb{R}^{\#} \to \mathbb{R}^{\#}$ by $a(x, y) := x + y$. Show that a is continuous.

Exercise 1.4: Show that $d(x, y) := |\ln(y/x)|$ is a metric on $\mathbb{R}_{++}$. Is $(\mathbb{R}_{++}, d)$ a complete metric space? What is a necessary and sufficient condition for a subset E of $\mathbb{R}_{++}$ to be bounded? What is a necessary and sufficient condition for E to be compact?

2. Infima, Suprema, Minima and Maxima.

Let E be a non-empty subset of $\overline{\mathbb{R}}$. $\gamma \in \overline{\mathbb{R}}$ is a *lower bound* for E provided $\gamma \leq x$ for all $x \in E$. Similiarly, γ is an *upper bound* for E provided $\gamma \geq x$ for all $x \in E$. Thus

- $-\infty$ is a lower bound for every nonempty subset E of $\overline{\mathbb{R}}$.
- $+\infty$ is a upper bound for every nonempty subset E of $\overline{\mathbb{R}}$.
- Every element of $\overline{\mathbb{R}}$ is both a lower bound and an upper bound of the empty set $\emptyset$.

We say that α is the *infimum* of E if α is a lower bound for E and when γ is any lower bound of E, then $\gamma \leq \alpha$.

Similarly, β is the *supremum* of E if β is an upper bound for E and when γ is any upper bound of E, then $\beta \leq \gamma$.

When E is any subset of $\overline{\mathbb{R}}$, its infimum will generally be denoted $\alpha(E)$ and equals $\inf\{x : x \in E\}$. Similarly the supremum will be denoted $\beta(E)$. The infimum and supremum are unique; lower and upper bounds generally are not unique.

Example 2.1: Suppose $E := (0, 1)$. Then any $\gamma \leq 0$ will be a lower bound for E and 0 is the infimum of E. Similarly, any $\delta \geq 1$ will be a upper bound for E and 1 is the supremum of E.

Theorem 2.1. *If E is a subset of $\overline{\mathbb{R}}$, then E has a unique infimum $\alpha(E)$ and a unique supremum $\beta(E)$ in $\overline{\mathbb{R}}$. Moreover, when E is non-empty, $\alpha(E) \leq \beta(E)$.*

Proof. When E is a non-empty bounded subset of $\mathbb{R}$, this is often an regarded as an axiom of the real number system. Alternatively (as in Rudin) the real numbers are constructed in a manner that ensures this property holds. When $-\infty \in E$, then $\alpha(E) = -\infty$. If E is a subset of $\mathbb{R}$ which is not bounded below, the definition implies that $\alpha(E) = -\infty$. Similarly for suprema.

Thus for any $x \in E$, we have $\alpha(E) \leq x \leq \beta(E)$. □

Applying these definitions to the empty set, you see that $\alpha(\emptyset) = +\infty$ and $\beta(\emptyset) = -\infty$ - so the last inequality in this theorem does not hold. When E is non-empty we say that $\alpha(E)$ is the *minimum* of E if $\inf\{x : x \in E\}$ is in E. Similarly, if $\beta(E) \in E$, then $\beta(E)$ is the *maximum* of E. When E is a finite set, then the minimum and maximum of E exist, but in general there need not be a minimum or maximum of a set E that is not a closed subset of $\overline{\mathbb{R}}$.

Example 2.2: (0,1) has no maximum or minimum. [0,1] has both a maximum and a minimum.

With the above conventions for a $-A$, we see that

$$\alpha(-A) \ := \ \inf\{x : x \in -A\} \ = \ -\sup\{x : x \in A\} \ = \ -\beta(A)$$

provide $A \subset \overline{\mathbb{R}}$ is non-empty. Similarly $\beta(-A) \ = \ -\alpha(A)$.

There is an algebra for infs and sups with respect to unions and intersections. Let $\{E_j : j \in J\}$ be a family of subsets of $\overline{\mathbb{R}}$. Then

$$\alpha(\cup_{j\in J} E_j) \ = \ \inf\{x : x \in E_j \text{ for some } j \in J\} \ = \ \inf_{j \in J} \ \alpha(E_j).$$

Also

$$\alpha(\cap_{j\in J} E_j) \ = \ \sup_{j\in J} \ \alpha(E_j) \qquad \text{if} \quad \cap_{j\in J} E_j \neq \emptyset, \text{ and similarly}$$

$$\beta(\cup_{j\in J} E_j) \ = \ \sup_{j\in J} \ \beta(E_j), \quad \text{and} \quad \beta(\cap_{j\in J} E_j) \ = \ \inf_{j\in J} \ \beta(E_j) \qquad \text{if} \quad \cap_{j\in J} E_j \neq \emptyset.$$

An infinite sequence $\Gamma := \{x^{(k)} : k \geq 1\} \subset \overline{\mathbb{R}}$ is said to be *increasing (resp. decreasing)* provided $x^{(k+1)} \geq x^{(k)}$, (resp. $x^{(k+1)} \leq x^{(k)}$). If it is either increasing or decreasing, then the sequence is said to be *monotone*. When Γ is an increasing (decreasing) sequence, then

$$\lim_{k\to\infty} x^{(k)} \ = \ \sup_{k\geq 1} x^{(k)} \quad (\text{or } \inf_{k\geq 1} x^{(k)}.) \tag{2.1}$$

These limits may be $\pm\infty$.

In general a bounded infinite sequence of real numbers need not have a limit. However every bounded sequence does have a *liminf* and a *limsup* defined as follows.

$$\liminf_{k\to\infty} x^{(k)} \ := \ \sup_{l\geq 1} \ \inf_{k\geq l} \ x^{(k)} \quad \text{and} \quad \limsup_{k\to\infty} x^{(k)} \ := \ \inf_{l\geq 1} \ \sup_{k\geq l} \ x^{(k)} \tag{2.2}$$

These numbers are finite when Γ is bounded. The definition (2.2) remains valid when $\Gamma \subset \overline{\mathbb{R}}$. In this case the $\liminf$ or $\limsup$ could be $\pm\infty$ and every such sequence has a liminf and a limsup with

$$\liminf \ \Gamma \ \leq \ \limsup \ \Gamma$$

Note that these definitions of liminf and limsup do NOT involve a metric or any topology so they are not limits in the sense of topology. They only involve the order relation on $\mathbb{R}$ or $\overline{\mathbb{R}}$.

Definition 2.2. When a sequence $\Gamma \subset \overline{\mathbb{R}}$, has

$$\gamma \;=\; \liminf_{k\to\infty}\; x^{(k)} \;=\; \limsup\;_{k\to\infty}\, x^{(k)}. \tag{2.3}$$

then we say that γ is the *limit* of Γ as $k \to \infty$ and write $\gamma \;=\; \lim_{k\to\infty}\; x^{(k)}$.

Example 2.3. Define a sequence by $x_{2m+1} = 1 - (1/m)$ and $x_{2m} := -1 + (1/m)$. Then $-1 < x_m < 1$ for all m,

$$\liminf_{m\to\infty} x_m \;=\; \lim_{m\to\infty}\; x_{2m} \;=\; -1 \quad \text{and also}$$

$$\limsup_{m\to\infty} x_m \;=\; \lim_{m\to\infty}\; x_{2m+1} \;=\; 1$$

When $E \subset \overline{\mathbb{R}}$ and $f : E \to \overline{\mathbb{R}}$ is a function we say that f is *increasing* on E provided

$$x, y \;\in E \;\&\; x \;<\; y \quad \Rightarrow \quad f(x) \;\leq\; f(y). \tag{2.4}$$

f is *strictly increasing* on E if this holds with $<$ in place of $\leq$ in the last inequality. Similarly for *decreasing* and *strictly decreasing* functions on E.

3. Norms on $\mathbb{R}^n$.

Let $\mathbb{N}$ be the set of all strictly positive integers (or natural numbers). Given $n \in \mathbb{N}$ define $I_n \;:=\; \{1, 2, \ldots, n\}$ and let $\mathbb{R}^n$ be the set of all n-tuples of real numbers. Elements of $\mathbb{R}^n$ are called n-vectors and $\mathbb{R}^n$ is a vector space with respect to the operations + and . defined for $x, y \;\in \mathbb{R}^n, c \in \mathbb{R}$ by

$$x + y \;:=\; (x_1 + y_1, x_2 + y_2, ..., x_n + y_n) \;\; \text{and} \;\; cx \;:=\; (cx_, cx_2, ..., cx_n)$$

A *norm* on $\mathbb{R}^n$ is a function $p : \mathbb{R}^n \to [0, \infty)$ with the properties;

(N1): $p(x) \geq 0$ and $p(x) = 0$ implies $x = 0$.
(N2): (homogeniety) $p(cx) \;:=\; |c|p(x)$.
(N3): (triangle inequality) $p(x + y) \;\leq\; p(x) \;+\; p(y)$.

Often norms are denoted $\|x\|$ instead of $p(x)$. The p-norm on $\mathbb{R}^n$ is defined by

$$\|x\|_p \;:=\; \{\sum_{j=1}^{n} |x_j|^p\}^{1/p} \qquad 1 \leq p < \infty. \tag{3.1}$$

When $p = 2$ this is called the Euclidean norm and arises from the inner product

$$\langle\, x, y\,\rangle \;:=\; \sum_{j=1}^{n} \quad x_j y_j. \tag{3.2}$$

When the subscript p is omitted in a norm, $p = 2$ is implied. We'll also often just write $|x|$ for $\|x\|_2$.

The $\infty-$ (or max-) norm is

$$\|x\|_\infty \;:=\; \max_{1\,\leq\, j\,\leq\, n} |x_j|. \tag{3.3}$$

It is straightforward to show that the properties (N1) & (N2) of a norm hold for this range of p. The fact that $\|x+y\|_p \leq \|x\|_p + \|y\|_p$ holds is easy for $p = 1, 2$ and ∞. For general $1 < p < \infty$), this inequality is Minkowski's inequality which will be proved later. When $0 < p < 1$, $\|x\|_p$ is not a norm as the triangle inequality (N3) does not hold and the unit ball with respect to $\|.\|_p$ is not a convex subset of $\mathbb{R}^n$.

These norms induce a metric on $\mathbb{R}^n$ via $d_p(x, y) \;:=\; \|x \;-\; y\|_p$

For each $1 \leq p \leq \infty$, the *open ball* $B^p_r(a)$, the *sphere* $S^p_r(a)$, and the *closed ball* $\overline{B}^p_r(a)$ of radius r and center $a \in \mathbb{R}^n$ with respect to the p-norm are the sets defined by

$$\begin{aligned} B^p_r(a) \;&:=\; \{x \in \mathbb{R}^n : \|x - a\|_p < r\} && (3.4)\\ S^p_r(a) \;&:=\; \{x \in \mathbb{R}^n : \|x - a\|_p = r\}, \text{ and} && (3.5)\\ \overline{B}^p_n(a) \;&:=\; \{x \in \mathbb{R}^n : \|x - a\|_p \leq r\} \;=\; B^p_r(a) \cup S^p_r(a) && (3.6)\end{aligned}$$

Here $r > 0$ and a is usually omitted when $a = 0$. For $p = 2$, we often omit the superscripts in these expressions and write $|x|$ in place of $\|x\|_2$.

Theorem 3.1. *For each $x \in \mathbb{R}^n \setminus \{0\}$, $p > 0$, $\|x\|_p$ is a decreasing function of p. It is strictly decreasing when x has more than 1 non-zero component. Moreover*

$$\lim_{p\to\infty} \|x\|_p \;=\; \inf_{p\geq 1} \|x\|_p \;=\; \|x\|_\infty. \tag{3.7}$$

Proof. The result is true for $x = 0$. Given $x \in \mathbb{R}^n \setminus \{0\}$ consider

$$f(p) \;:=\; p^{-1} \ln \left(\sum_{j=1}^{n} |x_j|^p \right) \;=\; \ln\, \|x\|_p.$$

Let $m = \|x\|_\infty$ and define $y \;:=\; x/m$. Express $f(p)$ in terms of y and differentiate with respect to p, then a computation shows that $f'(p) \leq 0$ for all $p > 0$. The first equality in (3.7) holds as the sequence $\{\|x\|_p : p \in \mathbb{N}\}$ is a decreasing sequence. The second part follows from the properties of $f(p)$. □

Corollary 3.2. *For $0 < p_1 < p_2 \leq \infty$, $B^{p_1}_r(a) \subset B^{p_2}_r(a)$.*

Given an "index" $p \in (1, \infty)$, the number $p^* \;:=\; p/(p-1)$ is called the *conjugate index* of p. Note that $p^{**} = p$. When $p = 1,\ \infty$, then $p^* := \infty,\ 1$ respectively. Another common way of stating this is

$$\frac{1}{p} + \frac{1}{p^*} \;=\; 1 \qquad \text{for } p \in [1, \infty]. \tag{3.8}$$

There are two other inequalities involving p-norms that are used throughout analysis and geometry. Given $x, y \in \mathbb{R}^n$, *Young's inequality* is that

$$|\langle x, y \rangle| \;\le\; \frac{1}{p}\,\|x\|_p^p \;+\; \frac{1}{p^*}\,\|y\|_{p^*}^{p^*} \qquad \text{for} \quad 1 < p < \infty. \tag{3.9}$$

Hölder's inequality is that

$$|\langle x, y \rangle| \;\le\; \|x\|_p\,\|y\|_{p^*} \qquad \text{for} \quad 1 \;\le\; p \;\le\; \infty. \tag{3.10}$$

Cauchy's inequality is the special case of Hölders inequality with $p = p^* = 2$. Hölder's inequality is easily verified for $p = 1, \infty$. These inequalities will be proved later in section 10. In particular Hölder's inequality yields the following useful inequalities.

Theorem 3.3. *Suppose* $1 \le p < q \le \infty$ *and* $x \in \mathbb{R}^n \backslash \{0\}$, *then*

$$0 < \;\|x\|_q \;\le\; \|x\|_p \;\le\; n^{1/p-1/q}\,\|x\|_q. \tag{3.11}$$

Proof. The result is straightforward if $q = \infty$, with the central inequality following from the montonicity result of theorem 3.1. and the second being a simple estimate. For finite q, let $z_j := |x_j|^p$ and $r := q/p > 1$. Then

$$\|x\|_p^p \;=\; \|z\|_1 \;=\; \langle z, e \rangle \quad \text{where} \quad e := (1, 1, \ldots, 1). \tag{3.12}$$

Now

$$\|z\|_r^r \;=\; \|x\|_q^q \qquad \text{so} \quad \|z\|_r \;=\; \|x\|_q^p.$$

Apply Holder's inequality to (3.12), then

$$\|x\|_p^p \;\le\; \|z\|_r\,\|e\|_{r^*} \;=\; n^{1-1/r}\,\|x\|_q^p. \tag{3.13}$$

Take p-th roots of both sides of this, then (3.11) follows. □

This result shows that the different p-norms actually generate equivalent topologies on the metric space $\mathbb{R}^n$. This will be used in the following form.

Corollary 3.4. *If a sequence* $\Gamma \subseteq \mathbb{R}^n$ *converges to* $\hat{x}$ *with respect to* $\|.\|_p$ *for some* $p \in [1, \infty)$, *then it converges to* $\hat{x}$ *with respect to any* $\|.\|_q$, $q \in [1, \infty]$.

Thanks to this result, you may prove the convergence, or otherwise, of a particular sequence using any convenient choice of p in $[1, \infty]$.

Exercises. .

Exercise 3.1 Prove that if $\langle a, x \rangle \ge 0$ for all $x \in \mathbb{R}^n$, then $a = 0$

Exercise 3.2 Prove that the Euclidean norm on $\mathbb{R}^n$ obeys the triangle inequality. Show that equality equality holds in (N3) for the Euclidean norm if and only if $y = cx$ for some $c \ge 0$.

Exercise 3.3 Suppose p_1 is a norm on $\mathbb{R}^m$ and p_2 is a norm on $\mathbb{R}^n$. Let $\mathbb{R}^{m+n} := \mathbb{R}^m \times \mathbb{R}^n$, show that

(a) $p((x,y)) \ := \ p_1(x) \ + \ p_2(y)$ is a norm on $\mathbb{R}^{m+n}$.
(b) $p((x,y)) \ := \ \max\{p_1(x), p_2(y)\}$ is a norm on $\mathbb{R}^{m+n}$.

4. Continuity and semi-continuity of functions.

Let (X,d) and (Y, d_1) be metric spaces and $F : X \to Y$ be a given function. F is said to be *continuous* at a point $x \in X$ provided for each $\epsilon > 0$, there is a δ such that

$$d(x,y) < \delta \quad \text{implies that} \quad d_1(F(x), F(y)) < \epsilon \tag{4.1}$$

It is continuous on a subset E of X provided it is continuous at each point of E. It is *Lipschitz continuous* on a subset E of X provided there is a constant L such that

$$d_1(F(x), F(y)) \ \leq \ L\, d(x,y) \qquad \text{for all } x, y \in E. \tag{4.2}$$

The function F is said to be *locally Lipschitz* on X provided it is Lipschitz continuous on each compact subset of X.

The above considerations apply to functions between general metric spaces. When f is a real valued, or extended real valued, there are some further defininitions that will be used frequently. Suppose $f : X \to \overline{\mathbb{R}}$ is a function and $c \in \mathbb{R}$, then a *level (or contour) set* of f is the set

$$L_c(f) := \{x \in X : f(x) = c\}. \tag{4.3}$$

The *synoptic (or sublevel) sets* of f are

$$S_c(f) := \{x \in X : f(x) \leq c\}. \tag{4.4}$$

When f is continuous these are closed subsets of X. The *effective domain of f* is the set $\mathrm{dom}(f) \ := \ \{x \in X \ : \ f(x) \in \mathbb{R}\}$ and the *graph of f* is the set

$$G(f) \ := \ \{(x, f(x)) : x \in X\} \qquad \subset X \times \overline{\mathbb{R}}. \tag{4.5}$$

The function f is said to be *proper* provided dom(f) is nonempty and $f(x) \neq -\infty$ for any $x \in X$. That is, a function is proper provided its range is a subset of $\mathbb{R}^{\#}$.

Example 4.1. The Newtonian potential is the function $f : \mathbb{R}^3 \to [0, \infty]$ defined by $f(x) \ := \ \frac{1}{|x|}$ for $x \neq 0$ and $f(x) := \infty$ for $x = 0$. This function is a continuous and proper e.r.v. function on $\mathbb{R}^3$. Its synoptic sets are annular regions in space for $c > 0$ and its level sets are spheres.

When a function defined on $\mathbb{R}^n$ or a ball in $\mathbb{R}^n$ is a function of $r := |x|$, so that $f(x) = \varphi(|x|)$ for all x, then the function is said to be *radial.* The level sets of such a function are (unions of) spheres.

Example 4.2. Consider the function $Ln : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by $Ln(x) \ := \ \ln |x|$ for $x \neq 0$ and $Ln(x) := -\infty$ for $x = 0$. This function is a continuous and e.r.v. function on $\mathbb{R}^n$. It is not proper, its synoptic sets are balls centered at the origin and its level sets are spheres. The function $-Ln(x)$ is a proper continuous function on $\mathbb{R}^n$. The function $f : \mathbb{R}^n \to [0, \infty)$ defined by $f(x) \ := \ \ln 1 + |x|$ is a positive continuous real valued function on $\mathbb{R}^n$.

Suppose E is a nonempty subset of the metric space X and $f : E \to \mathbb{R}$ is a function. Consider the functions

$$f_k(x) := \inf\ \{f(y) : y \in E \text{ and } 0 < d(x,y) < 1/k\}$$

For each $x \in E$, the sequence $\{f_k(x) : k \geq 1\}$ is an increasing sequence, so

$$\liminf_{y \to x} f(y) \ := \ \lim_{k \to \infty}\ f_k(x) \ = \ \sup_{k \geq 1}\ f_k(x).$$

is well-defined either as a real number or $\pm\infty$. Similarly define

$$g_k(x) := \sup\ \{f(y) : y \in E \text{ and } 0 < d(x,y) < 1/k\}, \qquad \text{and}$$

$$\limsup_{y \to x} f(y) \ := \ \lim_{k \to \infty}\ g_k(x) \ = \ \inf_{k \geq 1}\ g_k(x).$$

and observe that since $f_k(x) \ \leq \ g_k(x)$ for all k you have

$$\liminf_{y \to x} f(y) \ \leq \ \limsup_{y \to x} f(y).$$

A function $f : E \to \mathbb{R}$ is said to be *lower semi-continuous (l.s.c.)* at the point $x \in E$ provided

$$f(x) \ \leq \ \liminf_{y \to x}\ f(y).$$

It is l.s.c. on E provided it is l.s.c. at each point of E. f is *upper semi-continuous (u.s.c.)* at $x \in E$ provided

$$f(x) \ \geq \ \limsup_{y \to x}\ f(y).$$

It is u.s.c. on E provided it is u.s.c. at each point of E.

Theorem 4.1. *Let (X,d) be a metric space and $f : X \to \overline{\mathbb{R}}$ be a function. Then f is l.s.c. on X if and only if the synoptic sets $S_c(f)$ defined by (4.4) are closed for each $c \in \mathbb{R}$.*

Proof. Good exercise. □

It is straightforward to show that f is continuous at x if and only if f is both l.s.c and u.s.c. at x.

Example 4.3. The function $f(x) \ := \ \sin(x^{-1})$ does not have a limit as $x \to 0,\ x \neq 0$. Define $f(0) = c$. The set of all limit points of $f(x)$ as $x \to 0$ is [-1,1], and the above deinitions imply that $\liminf_{x \to 0}\ \sin(x^{-1}) \ = \ -1$ and $\limsup_{x \to 0}\ \sin(x^{-1}) \ = \ 1$. Thus f is l.s.c. at 0 provided $c \leq -1$ and it is u.s.c at 0 provided $c \geq 1$.

4.1. **Envelopes of a family of Functions.** Suppose $\mathcal{F} \ := \ \{f_j : j \in \mathcal{J}\}$ is a family of e.r.v. functions on X. Then the *lower envelope* F and the *upper envelope* G of $\mathcal{F}$ are the functions defined by

$$F(x) \ := \ \inf_{j \in \mathcal{J}}\ f_j(x), \qquad \text{and} \qquad G(x) \ := \ \sup_{j \in \mathcal{J}}\ f_j(x). \tag{4.6}$$

When each of the functions in $\mathcal{F}$ is continuous and $\mathcal{J}$ is a finite set, then both F and G will be continuous functions on X. When this is an infinite family the following theorem holds.

Theorem 4.2. *$\mathcal{F}$ is a family of e.r.v. functions on X. If each f_j is l.s.c. on X, then G is l.s.c on X. When each f_j is u.s.c. on X, then F is u.s.c. on X.*

Proof. The characterization of l.s.c. in terms of synoptic sets is used. Given $c \in \mathbb{R}$ we have $S_c(G) = \cap_{j \in \mathcal{J}} S_c(f_j)$ and each $S_c(f_j)$ is closed since f_j is l.s.c. Thus $S_c(G)$ is closed so G is l.s.c on X. Similarly for F. $\square$

Example 4.4: Consider the family of functions $f_m : [0,2] \to [0,1]$ defined by $f_m(x) = \frac{x^m}{1+x^m}; \quad m \in \mathbb{N}$. This family has the following envelopes

$$F(x) := \inf_{m \geq 1} f_m(x) = \begin{cases} 0 & x < 1 \\ \frac{x}{1+x} & x \geq 1 \end{cases}$$

$$G(x) := \sup_{m \geq 1} f_m(x) = \begin{cases} \frac{x}{1+x} & x \leq 1 \\ 1 & x > 1. \end{cases}$$

Note that F is u.s.c but not continuous, while G is l.s.c but not continuous at $x = 1$. The function $\lim_{m \to \infty} f_m(x)$ is a discontinuous function and neither l.s.c nor u.s.c at $x = 1$.

Exercises. .

Exercise 4.1 For $k \in \mathbb{N}$, define $f_k : [-1,1] \to [0,1]$ be the piecewise linear function that interpolates the values of $f(x) := 1 - x^2$ at the points $E_k := \{\frac{l}{2^k} : -2^k \leq l \leq 2^k\}$. Show that $f_k(x) \leq f_{k+1}(x)$ for all $x \in [-1,1]$ and that the upper envelope of this sequence is $f(x)$. Find an explicit expression for the area under f_k for k=1,2.

Exercise 4.2 For $m \in \mathbb{N}$, define $f_m(x) := \frac{x}{m} \exp(-x/m)$ for $x \geq 0$ and $G(x) := \sup_{m \in \mathbb{N}} f_m(x)$. Show that $G(x)$ is continuous and bounded on $[0,\infty)$. What is $\limsup_{x \to \infty} G(x)$?

Exercise 4.3. Prove theorem 4.1.

5. Minimization Problems and Minimizers.

Let (X,d) be a complete metric space and $f : X \to \overline{\mathbb{R}}$ be a given function.

The *minimization problem* for f on X is to find

$$\alpha(f;X) := \inf_{x \in X} f(x) \quad \text{and} \tag{5.1}$$

$$\mathcal{M}(f) := \{x \in X : f(x) = \alpha(f;X)\}. \tag{5.2}$$

The set $\mathcal{M}(f)$ is called the set of *minimizers of f on X*. $\alpha(f;X)$ is the *value* of this problem. Sometimes f is called the objective functions and in many economic and engineering applications it is called the cost, or value, function.

The *maximization problem* for f on X is to find

$$\beta(f;X) := \sup_{x \in X} f(x) \quad \text{and } \{x \in X : f(x) = \beta(f;X)\}. \tag{5.3}$$

One sees that $\beta(f;X) = -\alpha(-f;X)$ and a point $\hat{x}$ maximizes f on X if and only if it minimizes $-f$ on X. Thus a maximization problem can be converted into a minimization problem and vice versa. In these notes minimization problems will be emphasized.

When $X = \mathbb{R}^n$, we often write $\alpha(f)$ for $\alpha(f;\mathbb{R}^n)$.

Some simple 1-dimensional examples of optimization problems include the following.

Example 5.1. The function $f : \mathbb{R} \to \overline{\mathbb{R}}$ defined by $f(x) = e^x$ is continuous. $\alpha(f;\mathbb{R}) = \inf_{x\in\mathbb{R}} \quad e^x = 0$, but there is no minimizer of f on $\mathbb{R}$.

Example 5.2. If $f(x) := ax^2 + bx + c, \quad x \in \mathbb{R},\ a > 0$, then

$$\inf_{x\in\mathbb{R}} \quad f(x) = c - \frac{b^2}{4a}$$

and this function has a unique minimizer $\hat{x} = -b/2a$.

Example 5.3. Let $K := [-\pi/2, \pi/2]$. Define $f : K \to \overline{\mathbb{R}}$ by

$$f(x) = \begin{cases} -\infty & \text{if} \quad x = -\pi/2 \\ \tan\ x & \text{if} \quad |x| < \pi/2 \\ \infty & \text{if} \quad x = \pi/2 \end{cases}$$

This function f is continuous with respect to the metric d_e on $\overline{\mathbb{R}}$. The value $\alpha(f;K) = -\infty$ and $-\pi/2$ is the minimizer of f on K.

Example 5.4. Consider the function $Ln(x)$ defined as in example 4.2 and with $n = 1$. This function is even and continuous as a map of $\mathbb{R}$ into the metric space $(\overline{\mathbb{R}}, d_e)$. The value $\alpha(f) = -\infty$ and 0 is the (unique) minimizer of f on $\mathbb{R}$. The supremum of f on $\mathbb{R}$ is ∞, but there is no maximizer.

5.1. **Existence of Minimizers of Optimization Problems.** The following definitions are used to describe the properties of optimization problems for minimizing an e.r.v. function f on a non-empty set X.

A sequence $\Gamma := \{x^{(k)} : k \geq 1\} \subseteq X$ is a *descent sequence* for f on X provided $f(x^{(k+1)}) \leq f(x^{(k)})$ for all $k \geq 1$.

Γ is a *strict descent sequence* if $f(x^{(k+1)}) < f(x^{(k)})$ for all $k \geq 1$.

Γ is a *minimizing sequence* for f on X if it is a descent sequence and also

$$\lim_{k\to\infty} f(x^{(k)}) = \inf_{k\geq 1} f(x^{(k)}) = \alpha_(f;X).$$

Note that for any minimization problem, one may construct a minimizing sequence by induction. Given a finite number l (greater than or equal to 1) of elements of a strict descent sequence, then either $f(x^{(l)})$ is the infimum in which case we define $x^{(k)} = x^{(l)}$ for all $k \geq l$, or else there is another element $\tilde{x}$ of X with $f(\tilde{x}) < f(x^{(l)})$. In this case put $x^{(k+1)} = \tilde{x}$ and continue. This process generates either

(i) a sequence that attains the infimum in a finite number of steps, or

(ii) is a strict descent sequence.

In case (i) we are done. In case (ii) let γ be the infimum of $f(x^{(k)})$. If γ is the infimum of f on X we have a minimizing sequence for f on X. If not, there is an $\tilde{x} \in X$ with $f(\tilde{x}) < \gamma$. Start a new sequence, just as above, from this $\tilde{x}$ and continue.

Note that this construction does not assume any continuity of f, or any topology or metric on X and that f could take the values $\pm\infty$ at points in X

To prove the existence of minimizers we usually invoke the following classical theorem of real analysis. Berkovitz, Chapter 1, section 9. calls it the *fundamental existence theorem* for finite dimensional optimization.

Theorem 5.1. *(Weierstrass) Suppose K is a non-empty compact subset of a metric space (X, d) and $f : K \to \mathbb{R}$ is continuous, then*

(i) $\alpha(f; K) := inf_{x\in K} f(x)$ *and* $\beta(f : K) := sup_{x\in K} f(x)$ *are finite,*

(ii) *there are points $\hat{x}, \hat{y}$ in K such that $f(\hat{x}) = \alpha(f; K)$ and $f(\hat{y}) = \beta(f; K)$.*

Proof. (i) When f is continuous, K is compact, then $f(K)$ is a compact subset of $\mathbb{R}$. Hence it is closed and bounded and $\alpha(f; K), \beta(f, K)$ are finite.
(ii). Select a minimizing sequence $\{x^{(k)} : k \geq 1\}$ for f on K so that $f(x^{(k)})) \to \alpha(f; K)$ as described above. Since K is compact there is a convergent subsequence $\{x^{(k_j)} : j \geq 1\}$ and there is an $\hat{x}$ in K such that $x^{(k_j)} \to \hat{x} \in K$. Then $f(x^{(k_j)}) \longrightarrow f(\hat{x})$ as f is continuous, so $f(\hat{x}) = \alpha(f; K)$ and this is finite. Hence (i) and (ii) hold for minimization.

Similar results for maximization may be proved using these arguments with $-f$ in place of f. □

This existence theorem is not "constructive". We have not provided an an explicit algorithm or construction of the minimizing sequence - we only know that there must be one. For any particular optimization problem one would like to have specific algorithms for finding the minimizers or maximizers. Also note that, from the definition of compact set in section 1, (K, d) is a complete metric space since it is compact.

For minimizations problems there is a more general result that holds for lower semi-continuous e.r.v. functions.

Corollary 5.2. *Suppose K is a non-empty compact subset of a metric space (X, d) and $f : K \to \overline{\mathbb{R}}$ is lower semi-continuous, then there is an $\hat{x}$ in K that minimizes f on K.*

Proof. $\alpha(f; K) := inf_{x\in K} f(x)$ is a number in $\overline{\mathbb{R}}$. A minimizing sequence for f on K may be selected as in the preceding proof. This sequence has a convergent subsequence with a limit $\hat{x}$ in K, since K is compact. Thus

$$f(\hat{x}) \;\leq\; \liminf_{k\to\infty} \; f(x^{(k)}) \qquad \text{as } f \text{ is l.s.c on K.}$$

Since this is a minimizing sequence, the right hand side is $\alpha(f; K)$, and $\hat{x}$ is the desired minimizer. □

Example 5.5 The function $E : \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ defined by $E(x) = e^x$ for $x \in \mathbb{R}, E(-\infty) := 0, E(\infty) := \infty$ is continuous. It is the continuous extension of e^x to $\overline{\mathbb{R}}$. Thus, from corollary 3.3, since $K := \overline{\mathbb{R}}$ is compact, there is a minimizer of E on $\overline{\mathbb{R}}$. In fact $\alpha(E; \overline{\mathbb{R}}) := inf_{x\in\overline{\mathbb{R}}} \;\; e^x = 0$, and $-\infty$ is the unique minimizer of E on $\overline{\mathbb{R}}$. Similarly ∞ is the maximizer of E on $\overline{\mathbb{R}}$.

Note that the infimum and supremum of e^x on $\mathbb{R}$ are $0, \infty$ respectively - but they are not attained on $\mathbb{R}$.

Example 5.6. If $f(x) \; := \; \langle\, a, x \,\rangle + b, \quad x \in \mathbb{R}^n$, then

$$\inf_{x\in\mathbb{R}^n} \; f(x) = \begin{cases} -\infty & \text{if } a \neq 0 \\ b & \text{if } a = 0 \end{cases}$$

When $a = 0$, the set of all minimizers of f on $\mathbb{R}^n$ is $\mathbb{R}^n$. This is also the set of all maximizers of f on $\mathbb{R}^n$.

This semester we will primarily consider problems where functions to be extremized are defined on closed subsets of $\mathbb{R}^n$ and use the following characterization of compact sets in $\mathbb{R}^n$. With the definition provided in section 1.1, the following theorem is easily proved.

Theorem 5.3. *(Heine-Borel) A subset K of $\mathbb{R}^n$ is compact if and only if K is closed and bounded.*

Quite often we wish to prove the existence of minimizers of functions on non-compact metric spaces - such as $\mathbb{R}$ or $\mathbb{R}^n$. The simplest way to obtain such results is by showing that the minimizers must lie in a compact subset of the domain so the following definitions will be used repeatedly.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *weakly coercive* provided

$$f(x) \to +\infty \quad \text{as} \quad |x| \to \infty.$$

f is *coercive* on $\mathbb{R}^n$ provided

$$\lim_{|x|\to\infty} \; |x|^{-1}.f(x) \; = \; \infty.$$

When f is coercive on $\mathbb{R}^n$, then the function $g(x) \; := \; f(x) + \langle\, a, x \,\rangle + b$ is also coercive on $\mathbb{R}^n$ for any $a \in \mathbb{R}^n, \;\; b \in \mathbb{R}$. There are simple examples of functions f that are weakly coercive on $\mathbb{R}^n$ but a similar g is not weakly coercive.

Theorem 5.4. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semi-continuous and there is a $c \in \mathbb{R}$ such that $S_c(f)$ is non-empty and bounded. Then there is a $\hat{x} \in \mathbb{R}^n$ that minimizes f on $\mathbb{R}^n$. When f is also proper then $\alpha(f)$ is finite.*

Proof. Since f is l.s.c, $S_c(f)$ is closed in $\mathbb{R}^n$ for each real c. Choose c so that this set is bounded and non-empty then it is compact and we have $\alpha(f) = \inf_{x\in S_c(f)} f(x)$. (Why?). Use corollary 5.2 with $K = S_c(f)$ then f attains its value on $S_c(f)$ so there is an $\hat{x} \in S_c(f)$ such that

$$f(\hat{x}) \;=\; \inf_{x\in S_c(f)} \; f(x) \;=\; \alpha(f). \tag{5.4}$$

When f is proper then $f(\hat{x}) \neq -\infty$ so this value must be finite. □

Corollary 5.5. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is lower semi-continuous and weakly coercive. Then there is a minimizer of f on $\mathbb{R}^n$.*

Proof. Under these assumptions on f, $S_c(f)$ is closed and bounded for any $c \in \mathbb{R}$. If one such set is nonempty then, since the set is compact, there will be a minimizer of f in that set from theorem 5.2. If all the sets are empty then $f(x) \equiv \infty$ and the value of the problem is ∞ and the set of minimizers is $\mathbb{R}^n$. $\square$

It is worth noting that the problem of minimizing a l.s.c. e.r.v. function f defined on a non-empty closed subset K of $\mathbb{R}^n$, is equivalent to minimizing the l.s.c. e.r.v. function f_e on $\mathbb{R}^n$ where

$$f_e(x) \;:=\; f(x) \quad \text{for } x \in K \qquad \text{and} \quad f_e(x) \;:=\; \infty \quad \text{for } x \in \mathbb{R}^n \backslash K. \tag{5.5}$$

In particular the values of these two problems are the same and, provided the value is not ∞, the minimizers will also be the same. Thus the existence theorem above also applies to problems defined on proper closed subsets of $\mathbb{R}^n$.

Exercises. .

Exercise 5.1. If $p(x) = x^{2m} + a_1 x^{2m-1} + \ldots + a_{2m-1}x + a_{2m}$ is a polynomial of even degree on $\mathbb{R}$, then $\alpha(p) := \inf_{x\in\mathbb{R}}(p(x))$ is finite and there is minimizer $\hat{x}$ of p on $\mathbb{R}$. When $m \geq 2$, this minimizer need not be unique.

Exercise 5.2. If $p(x) = x^{2m+1} + a_1 x^{2m} + \ldots + a_{2m}x + a_{2m+1}$ is a polynomial of odd degree, then $\alpha(p) = -\infty$ and there is no minimizer, or maximizer, of p in $\mathbb{R}$.

Exercise 5.3. Suppose that $f(x) := x^2 + 2a\cos x - 2bx$ with a, b constants. Show that this function has minimizers on $\mathbb{R}$ and find the equations that they satisfy. Find bounds on this minimizer and show that the minimimum value is less than or equal to $(a+1)^2 - b^2 - 1$.

Exercise 5.4. Suppose y is a real number and $x^2 \;\leq\; \epsilon^2\, y^2 \;+\; C$ for some ϵ and $C > 0$. Prove that $|x| \;\leq\; \epsilon\,|y| \;+\; \sqrt{C}$. More generally show that $|x|^p \;\leq\; \epsilon^p\,|y|^p \;+\; C$ for some $C > 0, p > 1$ implies that there is a $C_1 > 0$ such that $|x| \;\leq\; \epsilon\,|y| \;+\; C_1$.

Exercise 5.5: Suppose E is a non-empty closed subset of $\mathbb{R}$ and $f : E \to \mathbb{R}$ is continuous. Let f_e be the extension of f to $\mathbb{R}$ defined by equation (5.5) above. Prove that f_e is l.s.c. on $\mathbb{R}$. Give a counterexample that shows this result need not hold when E is an open interval.

Exercise 5.6. Suppose K_1, K_2 are compact sets in $\mathbb{R}^m, \mathbb{R}^n$ respectively and $K := K_1 \times K_2$. When $F : K \to \mathbb{R}$ is a continuous function, prove that
(i) $G_1 : K_1 \to \mathbb{R}$ defined by $G_1(x) \;:=\; \inf_{y\in K_2}\; F(x,y)$ is a continuous function on K_1.
(The same proof shows that $G_2 : K_2 \to \mathbb{R}$ defined by $G_2(y) \;:=\; \inf_{x\in K_2}\; F(x,y)$. is continuous.)
(ii) Prove that

$$\inf_{x\in K_1}\; G_1(x) \;=\; \inf_{y\in K_2}\; G_2(y) \;=\; \inf_{(x,y)\in K}\; F(x,y)$$

(iii) Do the preceding equalities hold when K_1, K_2 are arbitrary non-empty subsets of $\mathbb{R}^m, \mathbb{R}^n$ respectively? If so, prove it. If not find a counterexample.

6. 1-dimensional Local Minimizers and Critical Points

In your first calculus course, you were taught some elementary results about finding the local minimizers, and maximizers of differentiable functions of a real variable. This section will summarize this material - and point out some issues with what you were probably told.
.

Definitions. When $f : \mathbb{R} \to \overline{\mathbb{R}}$ is a given function, a point $\tilde{x}$ in $\mathbb{R}$ is said to be
(i) a *local minimizer* of f if there is a $\delta > 0$, such that $f(x) \geq f(\hat{x})$ for $|x - \hat{x}| < \delta$.
(ii) a *strict local minimizer* of f if it is a local minimizer and $f(x) > f(\hat{x})$ when $0 < |x - \hat{x}| < \delta$.
(iii) an *isolated local minimizer* of f if $\hat{x}$ is a local minimizer of f and there is a $\delta_1 > 0$ such that f has no other local minimizer in $|x - \hat{x}| < \delta_1$.

If the above conditions hold for all $\delta > 0$, then the adjective local may be omitted. $\tilde{x}$ is a local maximizer when the inequality for f in (i) is reversed.

Note that a local minimizer of f may also be a local and/or a global maximizer of f. (Construct an example). However, a strict (or isolated) local minimizer cannot be a local maximizer. Clearly (ii) implies (i) and Proposition 6.1 below shows that $(iii) \Rightarrow (ii) \Rightarrow (i)$ above.

Example 6.1: Consider $f(x) := \begin{cases} x^2\,[2 + \cos(x^{-1})] & x \neq 0 \\ 0 & x = 0 \end{cases}$
Then f is an even function and $x^2 \leq f(x) \leq 3x^2$ for all x. That is, the function lies between two parabolas yet it has infinitely many positive local minimizers $\{x_k : k \geq 1\} \subset (0, 2]$. They have a limit point at 0 and each is an isolated local minimizer. 0 is the unique global minimizer and is a strict minimizer of f on $\mathbb{R}$ - but 0 is not a isolated local minimizer of f.

Proposition 6.1. *If $\hat{x}$ is an isolated local minimizer of f on $\mathbb{R}$, then $\hat{x}$ is a strict local minimizer.*

Proof. Choose $\delta_2 = \min(\delta, \delta_1)$ where δ, δ_1 as in the definitions (i) and (iii) of an isolated local minimizer. Let $I_2 = (\hat{x} - \delta_2, \hat{x} + \delta_2)$. If $\hat{x}$ is not a strict local minimizer, then there exists $\tilde{x} \in I_2$ such that $f(\tilde{x}) = f(\hat{x})$. But this implies $\tilde{x}$ is a local minimizer of f as $f(x) \geq f(\hat{x})$ on $(\hat{x} - \delta, \hat{x} + \delta)$. Contradiction. So $\hat{x}$ is a strict local minimizer. ☐

There is a straightforward existence theorem for local minimizers of f on an interval.

Theorem 6.2. *Suppose $f : [a, b] \to \mathbb{R}$ is continuous and there are points x_1, x_2, x_3 with*

$$a \leq x_1 < x_3 < x_2 \leq b \quad \text{and} \quad f(x_3) < \min\,(f(x_1), f(x_2)), \tag{6.1}$$

then there is a local minimizer of f in (x_1, x_2).

Proof. From Weierstrass' theorem, $\alpha = \inf_{[x_1,x_2]} f(x)$ is finite and there is a point $\hat{x} \in [x_1, x_2]$ such that $f(\hat{x}) = \alpha$. Now (6.1) implies that $\hat{x}$ is neither x_1 nor x_2, so there will be an $\hat{x}$ in (x_1, x_2) which is a local minimizer of f on $[a, b]$. ☐

In calculus I you are told that if a function f is differentiable on an interval (a,b) and it has a local minimizer at a point c in (a,b), then $f'(c) = 0$. This is sometimes called Fermat's rule - and he described this result for polynomials more than 50 years before calculus was invented. The following results describe criteria for a point to be a local minimizer of f. In many applications the minimum of an (objective) function occurs at corners or cusps where the function is not differentiable. In these cases, the following criteria may be used.

For the rest of this section, $I := (a, b)$ is an open interval and $f : I \to \mathbb{R}$ is assumed to be a continuous function. f is said to have a *left, (right), derivative* at $x \in (a, b)$ provided the following limits exist (in $\overline{\mathbb{R}}$)

$$D_- f(x) \ := \ \lim_{t \to 0^+} \ t^{-1} \left[f(x) - f(x - t) \right], \quad D_+ f(x) \ := \ \lim_{t \to 0^+} \ t^{-1} \left[f(x + t) - f(x) \right] \tag{6.2}$$

When $x \in (a, b)$, and these limits exist, are finite and equal then f is said to be differentiable at x. This common value is denoted $f'(x)$ or $Df(x)$. If $D_+ f(a)$ exists and is finite we say that f is differentiable at a. Similarly for b. When f is differentiable at each point in I and the derivative f' is continuous on I, then f is said to be *continuously differentiable* or C^1 on I. More generally, f is said to be C^k for $k \in \mathbb{N}$, provided the functions $f, Df, \ldots, D^k f$ exist and are continuous on I.

The following result is a sharp version of what you learn, or teach, in Calculus I. The first criterion involves the *lower derivative* of f at $\hat{x}$, while the second involves the left and right derivatives of f at $\hat{x}$.

Theorem 6.3. *(Necessary conditions for a local minimizer) Suppose $f : (a, b) \to \mathbb{R}$ is continuous and $\hat{x}$ is a local minimizer of f on (a, b), then*

(i) $\liminf_{|h| \to 0} \ |h|^{-1} \left[f(\hat{x} + h) - f(\hat{x}) \right] \ \geq \ 0.$

(ii) *If f has left and right derivatives at $\hat{x}$, then*

$$D_- f(\hat{x}) \ \leq \ 0 \qquad \text{and} \qquad D_+ f(\hat{x}) \ \geq \ 0. \tag{6.3}$$

(iii) *If f is differentiable at $\hat{x}$ then $f'(\hat{x}) = 0$.*

Proof. (i) From the definition of a local minimizer, there is a $\delta > 0$, such that $|h| < \delta \Rightarrow f(\hat{x} + h) - f(\hat{x}) \geq 0$. Thus

$$|h|^{-1} [f(\hat{x} + h) - f(\hat{x})] \geq 0 \qquad \text{for} \quad 0 < |h| \leq \delta.$$

Take lim inf of this as $|h| \to 0$, then (i) follows.
(ii) When $\hat{x}$ is a local minimizer of f, then from the criterion (i)

$$h^{-1} [f(\hat{x} + h) - f(\hat{x})] \ \geq \ 0 \qquad \text{for} \quad 0 < h \leq \delta.$$

Take the limit of this as $h \to 0$, then $D_+ f(\hat{x}) \geq 0$ - provided this limit exists. When h is negative then the inequality sign here is reversed so now $D_- f(\hat{x}) \leq 0$. whenever this limit exists.
(iii) If f is differentiable at $\hat{x}$, then $D_- f(\hat{x}) \ = \ D_+ f(\hat{x})$, so (6.3) show that this common value must be 0 or (iii) holds. □

These criteria are not sufficient conditions for a point to be a local minimizer since $\hat{x}$ in (a, b) may satisfy each of these conditions without being a local minimizer (or maximizer) of f on (a,b). A simple example is $f(x) = x^3$ which has $f'(0) = 0$ and (i) - (iii) hold at $\hat{x} = 0$ but 0 is not a local minimizer.

The following corollary provides a sufficient condition that is useful and simple, especially for numerical work - but only applies when f is not differentiable at the point $\hat{x}$.

Corollary 6.4. *(Sufficient conditions for a local minimizer) Suppose $f : (a, b) \to \mathbb{R}$ is continuous and f has left and right derivatives at $\hat{x}$ with*

$$D_- f(\hat{x}) \ < \ 0 \qquad \textit{and} \quad D_+ f(\hat{x}) \ > \ 0. \tag{6.4}$$

then $\hat{x}$ is a strict local minimizer of f on (a, b).

Proof. Straightforward use of the definitions. □

Note that if $\hat{x}$ is a local maximizer of f on (a, b), then conditions (i) and (ii) in theorem 6.3 change (to what?) - but condition (iii) is again the necessary condition for a local maximizer. If the signs are changed in the preceding corollary, you have a sufficient condition for a local maximizer.

The following examples satisfy (i) and (ii) of the theorem but the function does not have a derivative at its minimizers.

Example 6.2. Consider the function $f : \mathbb{R} \to [0, \infty)$ defined by $f(x) := \sqrt{|x|}$. This function is C^∞ on any interval that does not contain 0. It is not differentiable at 0, but you can evaluate $D_- f(0)$ and $D_+ f(0)$ and use the corollary to conclude that 0 is the unique minimizer of f on $\mathbb{R}$.

Example 6.1 (continued). The function f described in example 6.1 above obeys (i) but not (ii) or (iii) at its global minimizer $x = 0$. (Determine the derivative and see what happens as x approaches zero.) It has infinitely many other local minimizers - all of which satisfy (iii).

Example 6.3: The function $f(x) = |x|$ has a unique minimizer at $\hat{x} = 0$. f is not differentiable at 0, but $D_- f(0) = -1$ and $D_+ f(0) = 1$. So (i) and (ii) of the theorem hold - but not (iii).

In this course we will use the following terminology - which may be different to the definitions used in your Calculus text.

Definition. A point $\tilde{x} \in (a, b)$ is said to be a *critical point* of f provided f is differentiable at $\tilde{x}$ and $f'(\tilde{x}) = 0$. $\tilde{x}$ is an *isolated critical point* provided there is a $\delta > 0$ such that $f'(x) \neq 0$ for all $x \in B_\delta(\tilde{x}) \backslash \{\tilde{x}\}$.

Definition. A critical point $\tilde{x} \in (a, b)$ of f is said to be a *saddle point* of f provided that for any $\delta > 0$, there are points x_1, x_2 in $B_\delta(\tilde{x})$ such that

$$(f(x_1) \ - \ f(\tilde{x})).(f(x_2) \ - \ f(\tilde{x})) \ < \ 0$$

When f is C^1 on $I := (a, b)$ and continuous on $[a, b]$, then the following existence result is commonly used.

Theorem 6.5. *Suppose $f : (a, b) \to \mathbb{R}$ is continuous and C^1 on (a, b) and there are $x_1, x_2 \in I$ such that $a < x_1 < x_2 < b$ and $f'(x_1) \ < \ 0 \ < \ f'(x_2)$. Then there is at least one local minimizer $\hat{x}$ of f in (x_1, x_2).*

Proof. Since $f'(x_1) < 0$, there is $\delta > 0$ such that $f(x_1+\delta) < f(x_1)$ and similarly $f(x_2-\delta) < f(x_2)$, as $f'(x_2) > 0$. Take x_3 to be either $x_1+\delta$ or $x_2-\delta$ so that

$$f(x_3) = \min\ (f(x_1+\delta), f(x_2-\delta))$$

Then $f(x_3) < \min\ (f(x_1), f(x_2))$. From Theorem 6.2 there is a local minimizer $\hat{x}$ of f in (x_1, x_2). □

You should work out the corresponding criterion for local maximizers. What changes in the assumptions yield an existence result?

Example 6.4: Let $f : \mathbb{R} \to \mathbb{R}$ be the piecewise linear continuous function whose slope only changes at a finite number of points (nodes) $\Gamma := \{x_1, \ldots, x_m\}$. Let $x_0 := -\infty$, $x_{m+1} := \infty$ and $I_j := (x_{j-1}, x_j)$ be the j-th interval. Suppose the function has slope s_j on I_j with $s_1 < 0$ and $s_{m+1} > 0$. Then the function f is weakly coercive and has at least one local minimizer. If none of the slopes are zero, then the local minimizers of f are a subset of Γ. What is the maximum number of local minimizers of the function?

Exercises. .

Exercise 6.1: Suppose $f : I := [a, b] \to \mathbb{R}$ is a continuous function and $c_1, c_2 \in I$ are local maximizers of f with $c_1 < c_2$. Show that there is a local minimizer $\hat{x}$ of f in (c_1, c_2).

Exercise 6.2: Suppose f is defined as in exercise 5.3. Show that there is a unique critical point of this function when $|a| < 1$. Show that there are multiple critical points of this function when $|a| > 1$.

7. Error Estimates and Second Order Conditions.

In your first calculus course, you learnt about criteria for whether a critical point was a local minimizer or maximizer by checking the sign of the second derivative of a function at the critical point. Here these conditions will be described in terms of properties of the first derivative of the function. The condition that a local minimizer is isolated is sometimes called a local uniqueness result. The following theorem is the basic *error estimate* for 1-d optimization.

Theorem 7.1. *Suppose $f : (a, b) \to \mathbb{R}$ is C^1 and there are points $c_1 < c_2$ in (a, b) and a $c_0 > 0$, such that*

(i) $f'(x_2) - f'(x_1) \ \geq\ c_0(x_2 - x_1)$ *for all $a < x_1 < x_2 < b$, and*
(ii) $f'(c_1)f'(c_2) <\ 0$.

Then there is a unique minimizer $\hat{x}$ of f in (a, b) with

$$\hat{x} \in (c_1, c_2) \quad \text{and} \qquad |x - \hat{x}| \ \leq\ {c_0}^{-1}\ |f'(x)| \qquad \text{for all } x \in (a, b). \tag{7.1}$$

Proof. From (i), $f'(x_2) \geq f'(x_1) +\ c_0(x_2 - x_1) >\ f'(x_1)$ from the fact that $x_2 > x_1$ and $c_0 > 0$. Thus (ii) yields that $f'(c_1) < 0$ and $f'(c_2) > 0$. Theorem 6.5 then says there is a local minimizer $\hat{x}$ of f in (c_1, c_2) and $f'(\hat{x}) = 0$ as f is C^1. When $a < x_1 < \hat{x}$, then (i) yields that $f'(\hat{x}) - f'(x_1) \ \geq \ c_0 \ (\hat{x} - x_1) > \ 0$. Thus $f'(x_1) < 0$ as $f'(\hat{x}) = 0$ and $(\hat{x} - x_1) < {c_0}^{-1} \ |f'(x_1)|$ which is (7.1). Similarly, $\hat{x} < x_2 < b$, implies that $f'(x_2) > 0$. Thus $\hat{x}$ is the unique critical point of f in (a, b) - so it will also be the unique local minimizer. When $x \in (\hat{x}, b)$ substitute $x_1 = \hat{x}$ in (i) to again see that (7.1) holds. $\square$

Note that condition (i) holds provided f is C^2 on (a, b) and $f''(x) \geq c_0 > 0$ for all $x \in (a, b)$ from the differential mean value theorem. To use this as an error estimate one must have bounds on c_0. Then (7.1) provides an upper bound on the distance to the local minimizer in (c_1, c_2).

The condition (7.1) is often used as a stopping criterion in numerical optimization. That is a point x in I is taken to be the approximate minimizer whenevr $f'(x)$ is sufficiently small.

The following two results provide some necessary and sufficient conditions for a point to be a local minimizer of a C^1-function.

Theorem 7.2. *(Second-order necessary condition) Suppose $f : (a, b) \to \mathbb{R}$ is continuously differentiable and $\hat{x}$ is a local minimizer of f.*
(i) If $\hat{x}$ is an isolated critical point of f, then

$$\liminf_{h \to 0} \ h^{-1} \, f'(\hat{x} + h) \ = \ c \ \geq \ 0. \tag{7.2}$$

(ii) If $f''(\hat{x})$ exists, then $f''(\hat{x}) \geq 0$.

Proof. (i) Suppose $\hat{x}$ is an isolated critical point, then there is a $\delta > 0$, such that $0 < |h| < \delta$ implies that $|f'(\hat{x} + h)| \neq 0$. Thus $h^{-1} f'(\hat{x} + h)$ is of constant sign on the intervals $(-\delta, 0)$ and $(0, \delta)$. If $c < 0$, in (7.2), then there is a h_1 such that

$${h_1}^{-1} \ f'(\hat{x} + h_1) \ \leq \ \frac{c}{2} \ < \ 0 \quad \text{and} \quad 0 < |h_1| < \delta.$$

Suppose $h_1 > 0$, then $f(\hat{x} + h) - f(\hat{x}) = \int_0^h f'(\hat{x} + s) ds \ < \ 0$ for $0 < h \leq h_1$. This contradicts the assumption that $\hat{x}$ is a local minimizer. Similarly, if $h_1 < 0, \ c < 0, \ \hat{x}$ will not be a local minimizer, so we must have $c \geq 0$.
(ii) If $f''(\hat{x})$ exists, then $f''(\hat{x}) = \lim_{h \to 0} \ h^{-1} \ f'(\hat{x} + h) \ = \ c$. for some real c as $f'(\hat{x}) = 0$. If $c < 0$ repeat the argument used in the proof of part (i) to obtain a contradiction. Thus $c \geq 0$. $\square$

When a point $\hat{x}$ is a critical point of f on an interval then the following result guarantees that it is a local minimizer.

Theorem 7.3. *(2nd order sufficient condition) Suppose f is C^1 on (a, b), $\hat{x} \in (a, b)$ is a critical point of f and (7.2) holds with $c > 0$. Then $\hat{x}$ is an isolated strict local minimizer of f on (a, b).*

Proof. When (7.2) holds with $c > 0$, there is a $\delta > 0$ such that

$$|h| < \delta \quad \Rightarrow \quad h^{-1} f'(\hat{x} + h) \geq c/2.$$

$$\text{Then} \qquad f(\hat{x}+h) - f(\hat{x}) = \int_0^h f'(\hat{x}+s)ds \ \geq (c/2) \int_0^h \ s\ ds \ = \ (c/4)h^2.$$

for $0 \leq |h| < \ \delta$. Similarly for $f(\hat{x}) - f(\hat{x} - |h|)$ and thus $\hat{x}$ is an isolated strict local minimizer of f. □

In particular, the conditions of this theorem hold whenever $f''(\hat{x}) > 0$ - which is a result you surely have used before.

Example 6.2 (again). The function $f(x) \ := \ \sqrt{|x|}$ is continuous with $f'(x) \neq 0$ and $f''(x) < 0$ except at $x = 0$. This function is not differentiable at 0. So there is no critical point of this function and no point at which $f''(x) \geq 0$. But this function has a minimizer and no local maximizers! In particular, the usual second derivative test as described in Theorem 4.3.5 of Salas, Hille and Etgen (page 214, 8th edition) does not help find the unique minimizer.

This example shows that the "rules" for C^2- minimization do not work when the function is not C^2 - even at just one point. A different calculus is needed for "non-smooth" optimization. The basis of this calculus is a careful analysis of the minimization of convex functions and the possible extensions from convex functions to non-convex functions.

Exercises. .

Exercise 7.1: State and prove an analog of theorem 7.2 when $\hat{x}$ is an isolated local maximizer of f on (a, b).

Exercise 7.2 Suppose $\psi : (-1, 1) \to \mathbb{R}$ is continuous, $\psi(0) = 0$ and ψ is differentiable at 0 with $\psi'(0) = a > 0$, prove that there is a $\delta > 0$ such that

$$\psi(t) \ \leq \ at/2 \quad \text{on } (-\delta, 0) \qquad \text{and} \quad \psi(t) \ \geq \ at/2 \quad \text{on } (0, \delta).$$

8. Univariate Convex Functions.

Let I be a nontrivial interval in $\mathbb{R}$. That is, I is a connected subset of $\mathbb{R}$ containing at least two points - or $|I| > 0$. A function $f : I \to \mathbb{R}$ is *convex* provided

$$f((1-t)x + ty) \ \leq \ (1-t)f(x) \ + \ tf(y) \qquad \text{for all} \quad 0 \leq t \leq 1, \ x, y \in I. \qquad \textbf{(C)}$$

It is *strictly convex* if inequality holds in (C) whenever $x \neq y$ and $0 < t < 1$.

This definition can also be used for functions whose range is a subset of $(-\infty, \infty]$. Note that this definition just involves a pointwise (algebraic) inequality. It implies, or is equivalent to, a number of other inequalities for, or analytical results about, f.

Theorem 8.1. *Let $f : I \to \mathbb{R}$ be such that, for each $x \in I$, there is a $\eta \in \mathbb{R}$ satisfying*

$$f(y) \ \geq \ f(x) \ + \ \eta\ (y - x) \qquad \textit{for all} \quad y \in I \tag{8.1}$$

then f is convex on I.

Proof. Given $x \in I$ and a $y > x$ put $z := (1-t)x + ty$ with $0 < t < 1$. Then (8.1) implies

$$f(x) \;\geq\; f(z) + \eta(x-z) \;=\; f(z) + \eta t(x-y)$$

where $\eta := \eta_z$. Similarly

$$f(y) \;\geq\; f(z) + \eta(y-z) \;=\; f(z) + \eta(1-t)(y-x)$$

Multiply the first inequality by $(1-t)$, the second by t and add to find

$$(1-t)f(x) + tf(y) \;\geq\; f(z)$$

so f is convex on I. A similar argument holds when $y < x$. □

Corollary 8.2. *Suppose $f : I \to \mathbb{R}$ is such that for each $x \in I$ there is a $\eta \in \mathbb{R}$ satisfying*

$$f(y) \;>\; f(x) \;+\; \eta\;(y-x) \qquad \textit{for all} \quad y \in I,\; y \neq x, \tag{8.2}$$

then f is strictly convex on I.

Proof. Just as for the theorem - but with strict inequalities. □

Corollary 8.3. *Suppose I is an open interval and $f : I \to \mathbb{R}$ is differentiable at each point $x \in I$. Then f is convex on I if and only if*

$$f(y) \;\geq\; f(x) \;+\; f'(x)\;(y-x) \qquad \textit{for all} \quad x, y \in I. \tag{8.3}$$

It is strictly convex on I iff (8.3) holds with $>$ in place of $\geq$ and $y \neq x$.

Proof. When (8.3) holds then f is convex on I from the theorem. Conversely if f is convex on I, $0 < t \leq 1$ then

$$f(y) \;\geq\; t^{-1}[f((1-t)x + ty) - (1-t)f(x)] \;=\; f(x) \;+\; t^{-1}[f(x + t(y-x)) - f(x)].$$

for all $x, y \in I$. Let $t \to 0^+$ in this inequality then (8.3) holds. Similarly for strict convexity. □

This last corollary says that the graph of the function $z = f(x)$ lies on, or above, its tangent line at each point in I. Note that no continuity of the derivative is involved here.

Example 8.1. $f_p(x) := |x|^p$ is convex on $\mathbb{R}$ when $p \geq 1$. $-f_p(x)$ is convex on $[0, \infty)$ for $0 < p \;\leq\; 1$. This is usually proved by verifying the conditions of Corollary 8.3 - not by proving (C) directly. Sketch the graphs of these functions on $[0, \infty)$ for a number of values of p.

Example 8.2. Define $f_\infty,\; \chi_1$ on $\mathbb{R}$ by

$$f_\infty(x) \;:=\; \sup_{p \geq 1}\; f_p(x) \;=\; \begin{cases} |x| & |x| \;\leq\; 1 \\ \infty & |x| > 1 \end{cases} \tag{8.4}$$

$$\chi_1(x) \;:=\; \lim_{p \to \infty}\; f_p(x) \;:=\; \begin{cases} 0 & |x| \;\leq\; 1 \\ 1 & |x| = 1 \\ \infty & |x| > 1 \end{cases} \tag{8.5}$$

Both of these are even convex functions on $\mathbb{R}$. Note that the limit and the supremum differ on $(0, 1)$.

Example 8.3 $\quad f(x) := \; e^{\alpha x} \quad$ is convex on $\mathbb{R}$ $\quad$ for any choice of $\alpha \in \mathbb{R}$.

Example 8.4. $\quad h(x) := \begin{cases} x \ln x & x > 0 \\ 0 & x = 0 \end{cases}$ $\quad$ is convex and continuous on $[0, \infty)$. h is the *entropy* function on $[0, \infty)$.

For any function $f : I \to \mathbb{R}$, the *(first-order Newton) divided difference* is the slope of the straight line (or chord) joining points $(x_0, f(x_0))$ and $(x_1, f(x_1))$ on the graph of f. When $x_0 \neq x_1$, this is denoted

$$f[x_0, x_1] \;:=\; \frac{f(x_1) \;-\; f(x_0)}{x_1 \;-\; x_0}. \tag{8.6}$$

Theorem 8.4. *(3-slopes) Suppose $f : I \to \mathbb{R}$ is convex and $x_1 < x_2 < x_3$ are in I, then*

$$f[x_2, x_1] \;\leq\; f[x_3, x_1] \;\leq\; f[x_3, x_2] \tag{8.7}$$

or $\quad S_{12} \leq S_{13} \leq S_{23}$ *where* S_{ij} *= slope of* f *on the interval* $[x_i, x_j]$.

Proof. Since $x_2 \in (x_1, x_3)$, there is a $t \in (0, 1)$, such that $x_2 = (1-t)x_1 + tx_3$. From (C), $f(x_2) \leq (1-t)f(x_1) + tf(x_3)$. Rearrange this, then $f(x_2) - f(x_1) \leq t(f(x_3) - f(x_1))$, so

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{t(f(x_3) - f(x_1))}{t(x_3 - x_1)}$$

as $x_2 - x_1 = t(x_3 - x_1)$. Thus $S_{12} < S_{13}$.

A different rearrangement yields

$$f(x_3) - f(x_2) \geq (1-t)(f(x_3) - f(x_1)).$$

So, using the expression for x_2 again, $x_3 - x_2 = (1-t)(x_3 - x_1)$ $\quad$ and thus,

$$\frac{f(x_3) - f(x_2)}{x_3 - x_2} \geq \frac{(1-t)(f(x_3) - f(x_1))}{(1-t)(x_3 - x_1)}.$$

This yields $S_{23} \geq S_{13}$. $\square$

A consequence of this ordering is that if $f[x - \sigma, x]$ $\quad$ and $\quad f[x, x+s]$ are divided differences for f, then the left divided difference of a convex function f is an decreasing function of σ as $\sigma \to 0^+$. as , and the right divided difference is also a decreasing function of s as s decreases to 0. Taking limits of this observation leads to the following important regularity result.

Theorem 8.5. *(Differentiation of Convex Functions) Let $I := [a, b]$ be a non-trivial closed interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ be convex.*
(i) $D_+f(a)$ exists, it may be $-\infty$ but cannot be $+\infty$.
(ii) If x is in the interior (a, b) of I then $D_-f(x)$, $D_+f(x)$ exist and are finite, with $D_-f(x) \leq D_+f(x)$. Moreover, $D_-f(x)$ and $D_+f(x)$ are increasing functions on (a, b).
(iii) $D_-f(b)$ exists, it cannot be $-\infty$ but may be $+\infty$.
(iv) If $J := [c, d]$ is a closed subset of (a, b), then f is Lipschitz continuous on J with

$$|f(y) \;-\; f(x)| \;\leq\; L\,|y - x| \qquad \text{with } L := \max\{|D_+f(c)|, |D_-f(d)|\}. \tag{8.8}$$

Proof. (i) As noted above $f[a, a+s]$ is a decreasing function of s, hence $D_+f(a)$ exists as a number in $[-\infty, \infty)$.
(ii) Suppose $x \in (a,b), x_1 = x - \sigma, x_3 = x + s$. Take $x_2 = x$ in (8.7), then $f[x-\sigma, x] \leq f[x, x+s]$ Let $\sigma \to 0^+$, then this LHS is increasing and is bounded above by the expression on the right. Hence $D_-f(x)$ is the limit of these divided differences and is finite. Similarly as $s \to 0+$, the RHS is a decreasing function of s and the divided differences are bounded below - so $D_+f(x)$ exists and is finite. Moreover $D_-f(x) \leq D_+f(x)$.

To prove that $D_-f(x)$ is increasing choose two points $x_1 < x_2 \leq b$ and extend the above result to verify that

$$f[x_1 - s, x_1] \leq f[x_2 - s, x_2]$$

Take limits of both sides as $s \to 0^+$ then $D_-f(x_1) \leq D_-f(x_2)$. A similar proof holds for $D_+f(x)$.
(iii) the proof of (iii) is similar to that of (i).
(iv) When f is continuously differentiable on an interval containing $[c, d]$, then this result follows from the integral mean value theorem and the fundamental theorem of calculus. When f is not C^1-, then a similar proof holds with 1-sided derivatives in place of the usual derivative. This needs a more detailed proof. □

Corollary 8.6. *Suppose $I := (a,b)$ is an open interval in $\mathbb{R}$, and $f : I \to \mathbb{R}$ is convex then f is locally Lipschitz continuous on I.*

Proof. Suppose f is convex and $x \in I$. Then there is a $\delta > 0$ such that $[x-\delta, x+\delta] \subset I$. From (iv) of the theorem f is Lipschitz continuous on this interval. Hence f is locallly Lipschitz on I. □

When $D_f(x) = D_+f(x)$, then f is differentiable at x. When inequality holds here, f is said to have a *corner* at x. In real analysis you may see a proof that if f is convex on an interval I, then the set of points at which f is not differentiable is a set of measure zero. Equivalently convex functions have (classical) derivatives almost everywhere on their (essential) domain.

When f is C^1 or C^2 there are some simpler criteria for convexity in terms of the derivatives of f.

Theorem 8.7. *Let I be an open interval and assume $f : I \to \mathbb{R}$ is differentiable at each $x \in I$. Then f is convex on I if and only if f' is an increasing function on I. If $f''(x)$ exists, then $f''(x) \geq 0$.*

Proof. Theorem 8.5 shows that when f is convex and differentiable on I, then $f'(x)$ is increasing on I. When $f'(x)$ is increasing and $x < y \in I$, then

$$f(y) = f(x) + f'(\xi)(y-x) \geq f(x) + f'(x)(y-x) \text{ with } x < \xi < y.$$

from the mean value theorem. Similarly when $y < x$. Hence corollary 8.3 implies that f is convex.

If $f''(x)$ exists, then since $f'(x)$ is increasing on I, $f''(x)$ must be positive. □

In undergraduate calculus classes a function is defined to be convex, (or concave up), on an open interval I provided the assumptions of the following corollary hold. Note that of the examples given above only example 8.3 always satisfies the "undergraduate" definition.

Corollary 8.8. *Let I be an open interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ is such that $f''(x) \geq 0$ for each $x \in I$, then f is C^1 and convex on I.*

Proof. When $f''(x)$ exists then f' is continuous at x from a standard result in calculus, so f' will be continuous on I. Moreover f' is an increasing function on I as $f''(x) \geq 0$. Hence the result follows from the preceding theorem 8.7. □

Exercises. .

Exercise 8.1: (a) Complete the proof of theorem 3.1 by evaluating the derivative of the function f(p) and showing that $f'(p) \leq 0$ for all $p > 0$.
(b) Evaluate $f''(p)$ and prove that f is a convex function on $(0, \infty)$.

Exercise 8.2: Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex with $f(0) = 0$. Show that $f(y) \geq f(1)y$ for $y \geq 1$ and $f(y) \geq f(-1)|y|$ for $y \leq -1$. Hence show that f is weakly coercive on $\mathbb{R}$ provided $m := \min\{f(-1), f(1)\} > 0$.

Exercise 8.3: Suppose I is a non-trivial interval in $\mathbb{R}$ and $f, g : I \to [0, \infty)$ are convex, increasing functions on I. Show that $h(x) := f(x)\, g(x)$ is a convex increasing function on I. Give a counterexample to this result if g is allowed to be negative on a subset of I.

Exercise 8.4 Suppose a person saves \$$P$ a year and the funds accummulate at an interest rate of $r\%$ per year. After m years the balance in the account is given by

$$A_m = \frac{P}{i}\,[(1 + i)^m - 1] \qquad \text{where } i = r/100.$$

Show that A_m is an increasing, convex function of i.

Exercise 8.5: Suppose $a : [0, 1] \to \mathbb{R}$ is continuous and C^2- on $(0, 1)$. Describe conditions on the function a that imply that $f(x) := \frac{a(x)}{1-x}$ is convex on $[0, 1)$.

Exercise 8.6: (a) Show that the function $f : (0, \infty) \to [0, \infty)$ defined by $f(x) := |\ln x|$ is convex. At what points is f differentiable? Is it continuously differentiable on $(0, \infty)$?
(b) Show that $d(x, y) := |\ln(y/x)|$ is a metric on $\mathbb{R}_{++}$.
(c) Is $(\mathbb{R}_{++}, d)$ a complete metric space? Give reasons for your answer.
(d) What is a necessary and sufficient condition for a subset E of $\mathbb{R}_{++}$ to be bounded? What is a necessary and sufficient condition for E to be compact?

9. 1-d Optimization and Inequalities

Suppose I is a compact subset of $\mathbb{R}$ and $f : I \to \mathbb{R}$ is a continuous function. Then Weierstrass' theorem, says that

$$\alpha(f, I) \ := \ \inf_{x \in I} \ f(x) \qquad \text{and} \quad \beta(f, I) \ := \ \sup_{x \in I} \ f(x)$$

are finite and there are minimizers and maximizers where these values are attained. These points sometimes are called the *global minimizers (or maximizers)* of f on I. The conditions satisfied by such a minimizer may be described by the following extension of theorem 6.3 - which uses divided differneces rather than derivatives.

Theorem 9.1. *(Necessary conditions for minimizer) Suppose $I := [a, b]$ is bounded and $f : I \to \mathbb{R}$ is continuous and minimized at a point $\hat{x}$. When*
(i) $\hat{x} = a$, then $f[a, a + s] \ \geq \ 0$, for all $s \in (0, b - a]$,
(ii) $\hat{x} \in (a, b)$, then

$$f[x, \hat{x}] \ \leq \ 0 \ \textit{for} \quad a < x < \hat{x} \qquad \textit{and} \quad f[\hat{x}, x] \ \geq \ 0, \ \textit{for} \quad \hat{x} < x < b. \tag{9.1}$$

(iii) $\hat{x} = b$, then $f[b - s, b] \ \leq \ 0$, for all $s \in (0, b - a]$.

Proof. The proofs of these are all similar and straightforward. To prove (i) observe that since a minimizes f on I one has that

$$s^{-1} \ [f(a + s) - f(a)] \ \geq \ 0 \qquad \text{for all} \ \ 0 \ < \ s \ < \ b - a.$$

which is (i). □

In particular, the theorem shows that the global minimizers of a function f on a closed interval satisfies **inequalities** involving the difference quotients of f. We don't need to take limits! If $\tilde{x}$ is a maximizer of f on I, then analogous (reversed) inequalities hold.

When the function f is convex on the closed interval $I := [a, b]$, there are stronger results as we have that the left and right derivatives exist and are increasing functions. In particular the following result shows that these conditions are sufficient for a point to be a minimizer of f on I.

Theorem 9.2. *(Sufficient conditions for convex 1-d minimization) Let $I = [a, b]$ be a bounded closed interval and $f : I \to \mathbb{R}$ be continuous and convex. If*
(i) $D_+ f(a) \ \geq \ 0$, then a minimizes f on I,
(ii) $\hat{x} \in (a, b)$ satisfies (9.1) then $\hat{x}$ minimizes f on I,
(iii) $D_- f(b) \ \leq \ 0$ then b minimizes f on I.
Moreover the set of all minimizers of f is a closed subinterval of $[a, b]$.

Proof. The first 3 parts are a good exercise. The last sentence holds as the set of all minimizers of f is $\mathcal{S} := \{x \in [a, b] : f(x) = \alpha \ \}$. This is closed as f is continuous. If $x_1, x_2 \in \mathcal{S}$ and $x_t = (1 - t)x_1 + tx_2$, then

$$f((1 - t)x_1 + tx_2) \leq (1 - t)f(x_1) + tf(x_2) = \alpha.$$

Thus $\mathcal{S}$ is a closed subinterval of I. □

Corollary 9.3. *(Uniqueness) If $f : [a,b] \to \mathbb{R}$ is strictly convex, then there is a unique minimizer $\hat{x}$ of f on $[a,b]$.*

Proof. Suppose $\hat{x_1}, \hat{x_2}$ are two distinct minimizers, then $f(\hat{x_1}) = f(\hat{x_2}) = \alpha$, so that

$$f\left(\frac{\hat{x_1}+\hat{x_2}}{2}\right) < [f(\hat{x_1}) + f(\hat{x_2})]/2 \;=\; \alpha \qquad \text{as f is strictly convex.}$$

This contradicts the definition of α so must have $\hat{x_1} = \hat{x_2}$ or the minimizer is unique. □

When f is convex on a closed interval, there also are special properties of the maximizers of f on I. The maximum principles described below are the prototypes of some very important theorems for differential inequalities and solutions of elliptic partial differential equations.

Theorem 9.4. *Suppose $f : [a,b] \to \mathbb{R}$ is continuous and convex, then*
(i) (weak maximum principle) $\beta(f,I) \;=\; \max\{f(a), f(b)\}$ *and,*
(ii). (strong maximum principle) if there is a point $c \in (a,b)$ that maximizes f on [a,b], then $f(x)$ is constant on $[a,b]$.

Proof. (i) Suppose $D_+f(a) \geq 0$, then $D_+f(x) \geq 0$ for all $x \in [a,b)$, so f is increasing on $[a,b)$ and either f is constant on the whole interval or the maximum of f is at b. If $D_+f(a) < 0$, then a is a local maximum of f. If $D_-f(b) \leq 0$, then since $D_+f(x), D_-f(x)$ are increasing functions on (a,b), we have $D_-f(x) \leq 0$ for all $x \in (a,b)$, so f decreases on (a,b) and thus $\{a\}$ is the maximizer. If $D_-f(b) > 0$, then f is increasing near b and there cannot be an interior maximizer. The maximizer is either at a or b and (ii) holds.

(ii) A good exercise is to derive this from the previous arguments. □

Many important inequalities of analysis are based on elementary inequalities for functions of a real variable. These inequalities may often be proved using a minimization argument. Essentially the inequality

$$f(x) \;\leq\; g(x) \quad \text{holds iff} \quad g(x) \;-\; f(x) \;\geq\; 0 \qquad \text{or} \quad \frac{g(x)}{f(x)} \;\geq\; 1$$

for all $x \in I$. For the second case, we must also require that f be positive on I. In either case the inequality is proved if we can obtain an appropriate lower bound. Another common method is to look at whether functions of more than 1 variable may be factorized to the form

$$f(tx) \;=\; g(t)\; h(e_x) \qquad \text{where} \quad e_x = x/|x| \qquad \text{and} \quad t \;\geq\; 0.$$

Then the minimization of f is reduced to a minimization with respect to t, e_x separately. Since the unit vectors lie on a compact subset of $\mathbb{R}^n$, it is the minimization of g on $(0,\infty)$ that often is more difficult.

Exercises. .

Exercise 9.1. Prove (i) - (iii) of theorem 9.2.

Exercise 9.2. Prove (ii) of theorem 9.4.

Exercise 9.3. Prove that $x^{1-t}\,y^t \;\leq\; (1-t)x \;+\; ty$ for all $x, y > 0$ and $0 \leq t \leq 1$. When $t = 1/2$, this is the arithmetic-geometric mean inequality.

Exercise 9.4. Suppose $f : [0, L] \to \mathbb{R}$ is a continuous function with $D_+f(0) < 0,\ D_-f(L) > 0$. Show that the global minimizers of f on $[0, L]$ are attained at points in $(0, L)$.

Exercise 9.5. Prove the following inequalities for the function $f_p(x) := x^p$.
(i) Given $p > 1, s > 0$, find a positive constant C(s,p) such that $x^p \;\geq\; sx \;-\; C(s,p)$ for all $x > 0$.
(ii) Given $p > 1$, find the smallest positive constant $C(p) \in (0, 1)$ such that $x^{p-1} \;\leq\; x^p \;+\; C(p)$ for all $x > 0$.
(iii) Suppose $p > 1,\ x, y > 0$, prove that

$$x^p \;+\; p\,x^{p-1}\,y \;\leq\; [x \;+\; y]^p \;\leq\; x^p \;+\; p\,[x \;+\; y]^{p-1}\,y \qquad \text{and}$$

$$[x \;+\; y]^p \;\leq\; 2^{p-1}\,(x^p \;+\; y^p).$$

(iv) (general parallelogram inequality) Prove that if $p \;\geq\; 2,\; x, y > 0$, then

$$|x + y|^p \;+\; |x - y|^p \;\geq\; 2\,[\,x^p \;+\; y^p\,].$$

(v) (Clarkson's first inequality) Show that when $p \;\geq\; 2,\; x, y > 0$, then

$$\left|\frac{x+y}{2}\right|^p \;+\; \left|\frac{x-y}{2}\right|^p \;\leq\; \frac{1}{2}[\,x^p \;+\; y^p\,].$$

(vi) Given $p \in (0, 1), \epsilon > 0$ show that there is a constant $C(\epsilon, p) > 0$ such that

$$x^p \;\leq\; \epsilon\,x \;-\; C(\epsilon, p) \qquad \text{for all } x > 0.$$

(vii) Given $p \in (0, 2),\ \epsilon > 0$, show that there is a constant $C(\epsilon, p) > 0$ such that

$$x^p \;\leq\; \epsilon\,x^2 \;-\; C(\epsilon, p) \qquad \text{for all } x > 0.$$

Exercise 9.6. Given x, y in $(0, \infty)$, their *harmonic mean* is $h(x, y) := \frac{2xy}{x+y}$. Prove the following
(i) If $0 < x < y$, then $x < h(x, y) < y$.
(ii) $h(x, y) \;\leq\; \sqrt{xy} \;\leq\; (x + y)/2$

Exercise 9.7. Show that the function $f(x) := |x|\,\ln(1 + |x|)$ is an even, positive, coercive, strictly convex function on $\mathbb{R}$ and f is strictly increasing on $[0, \infty)$. Prove that for each $\epsilon > 0, p > 0$, there is a constant $C > 0$ such that

$$\ln(1 + x) \;\leq\; \epsilon\,x^p \;+\; C. \qquad \text{for } x \geq 0.$$

This implies that while f grows faster than $|x|$ as $|x| \to \infty$, it grows more slowly than $|x|^p$ for every $p > 1$.

Exercise 9.8. Suppose I is an open interval in $\mathbb{R}$ and $F : I \times (0, \infty) \to \mathbb{R}$ is a C^2-function. Assume the function $F(x, .)$ is weakly coercive and $D_t^2 F(x, t) \geq c > 0$ for all $x \in I$. Define

$f : I \to \mathbb{R}$ by

$$f(x) \ := \ \inf_{t>0} \ F(x,t).$$

Prove that f is continuously differentiable on I and that $f'(x) \ = \ D_x\, F(x, T(x))$ for any $x \in I$, where $T(x)$ is a minimizer of $F(x, .)$ on $(0, \infty)$.

10. Fundamental Inequalities for $\mathbb{R}^n$

For the rest of this course, we shall mostly work in $\mathbb{R}^n$ with n a natural number. A standard method for proving many results in multivariable calculus is to reduce the problem to a simpler, lower dimensional problem. In this section some fundamental results about norms and inner products in $\mathbb{R}^n$ will be proved using elementary one-dimensional calculus.

First note that the triangle inequality $|x \ + \ y| \ \leq \ |x| \ + \ |y|$ holds in $\mathbb{R}$, so absolute values provide a norm on $\mathbb{R}$. Thus the 1-norm and the $\infty-$norm on $\mathbb{R}^n$ satisfy the triangle inequality (N3) so they are norms on $\mathbb{R}^n$.

For other values of $p \in (1, \infty)$, the basic inequalities for the p-norm described in section 3 follow from either the following simple lemma or the observation that $|x|^p$ is convex on $\mathbb{R}$. As earlier, $p^* = (p-1)/p$ is the conjugate index of p.

Lemma 10.1. *Suppose $p \in (1, \infty), \ y \in \mathbb{R}$, then*

$$f_p(x) \ := \ \frac{1}{p}\,|x|^p \ - \ x\,y \ \geq \ -\,\frac{1}{p^*}\,|y|^{p^*} \qquad \text{for all } x \in \mathbb{R}. \tag{10.1}$$

If $y \neq 0$ then equality holds here if and only if $y \ = \ |x|^{p-2}\,x$.

Proof. When $p > 1$, the function f_p is strictly convex and coercive on $\mathbb{R}$, so it has a unique minimizer on $\mathbb{R}$. This minimizer occurs at the unique solution of $f_p'(x) = 0$ and the results follow by evaluation. □

The generalization of this to $\mathbb{R}^n$ is called *Young's inequality.*

Theorem 10.2. *If $p \in (1, \infty)$ and $p^* := (p-1)/p$, then*

$$|\langle\, x, y \,\rangle| \ \leq \ \frac{1}{p}\,\|x\|_p^p \ + \ \frac{1}{p^*}\,\|y\|_{p^*}^{p^*} \qquad \text{for all } x, y \in \mathbb{R}^n. \tag{10.2}$$

For non-zero x, y equality holds here if and only if either x_j, y_j are both zero or else $y_j \ = \ |x_j|^{p-2}\ x_j$

Proof. Use the preceding lemma for each x_j, y_j and add to obtain

$$|\langle\, x, y \,\rangle| \ \leq \ \sum_{j=1}^{n} \ |x_j\ y_j| \ \leq \ \sum_{j=1}^{n} \ \left(\frac{|x_j|^p}{p} \ + \ \frac{|y_j|^{p^*}}{p^*} \right)$$

which is (10.2). □

Theorem 10.3. *(Hölder's inequality) If $p \in (1, \infty)$ and $p^* := (p-1)/p$, then*

$$|\langle x, y \rangle| \;\leq\; \|x\|_p \, \|y\|_{p^*} \qquad \text{for all } x, y \in \mathbb{R}^n. \tag{10.3}$$

When $x \neq 0$, equality holds here when $y_j \;=\; c|x_j|^{p-2} \, x_j$ for some $c \in \mathbb{R}$ and all j such that $x_j \neq 0$.

Proof. Equality holds here if either x or y is the zero vector. For $x, y \in \mathbb{R}^n \setminus \{0\}$, define $u := x/\|x\|_p$ and $v := y/\|y\|_{p^*}$. Then $\|u\|_p \;=\; \|v\|_{p^*} = 1$. From Young's inequality,

$$|\langle u, v \rangle| \;=\; \sum_{j=1}^{n} \quad \frac{|x_j \; y_j|}{\|x\|_p \; \|y\|_{p^*}} \;\leq\; \frac{1}{p} \;+\; \frac{1}{p^*} \;=\; 1.$$

This implies (10.3) and the equality condition here follows from that of theorem 10.2. $\square$

Corollary 10.4. *If $p \in [1, \infty]$, and $x \in \mathbb{R}^n$ then*

$$\| \, x \, \|_p \;=\; \sup_{\|y\|_{p^*} \leq 1} \; \langle \, x, y \, \rangle. \tag{10.4}$$

Proof. This is easily verified when $p = 1$ or ∞. For $p \in (1, \infty)$, this follows from the above form of Hölder's inequality. Note that the sup will always be attained. $\square$

As noted earlier, the triangle inequality for the p-norm on $\mathbb{R}^n$ follows from the 1-dimensional triangle inequality when $p = 1, \infty$ and it doesn't hold when $0 < p < 1$ and $y \neq cx$. The following is a proof for the cases $1 < p < \infty$ that only uses the convexity of $|x|^p$.

Theorem 10.5. *(Minkowski's inequality) If $p \in (1, \infty)$, $x, y \in \mathbb{R}^n$, then*

$$\| \; x \;+\; y \; \|_p \;\leq\; \| \; x \; \|_p \;+\; \| \; y \; \|_p. \tag{10.5}$$

For nonzero x, equality holds here if and only if $y = cx$ for some $c \;\geq\; 0$.

Proof. When either x or y is the zero vector then equality holds in (10.5) . The fact that $|u|^p$ is convex on $\mathbb{R}$ implies that

$$|\,(1-t)u + tv\,|^p \;\leq\; (1-t)\,|u|^p \;+\; t\,|v|^p \qquad \text{for all } 0 \leq t \leq 1. \tag{10.6}$$

Given non-zero vectors x, y, let

$$u = x_j/\|x\|_p, \quad v = y_j/\|y\|_p, \quad t \;=\; \frac{\|y\|_p}{\|x\|_p + \|y\|_p}$$

$$\text{Then} \qquad (1-t)u + tv \;=\; \frac{x_j + y_j}{\|x\|_p + \|y\|_p}.$$

Use (10.6) for each j and sum to obtain

$$(\|x\|_p + \|y\|_p)^{1-p} \sum_{j=1}^{n} \; |x_j + y_j|^p \;\leq\; \|x\|_p{}^{1-p} \sum_{j=1}^{n} \; |x_j|^p \;+\; \|y\|_p{}^{1-p} \sum_{j=1}^{n} \; |y_j|^p$$

Change these sums to norms and (10.5) follows. The equality claim is proved by checking when equality holds in (10.6). $\square$

11. Convex Sets in $\mathbb{R}^n$.

A subset C of $\mathbb{R}^n$ is said to be *convex* provided

$$x, y \in C \quad \text{implies } (1-t)x + ty \in C \quad \text{for } 0 \le t \le 1.$$

We often write $[x, y] := \{(1-t)x + ty : 0 \le t \le 1\}$ and this is the *closed (line) interval* from x to y in $\mathbb{R}^n$. A convex set is said to be *non-trivial* if it contains at least two distinct points. A convex subset C of $\mathbb{R}^n$ is said to be a *convex cone* provided $cx \in C$ whenever $c \ge 0$ and $x \in C$.

In the following the topology on $\mathbb{R}^n$ will be assume to be that induced by the Euclidean metric - or an equivalent p-norm. Thus a convex set is said to be open, respectively closed, provided it is open (closed) in the Euclidean metric.

The empty set and singletons are convex sets - all other convex sets are said to be *nontrivial convex sets*. When C contains 2 distinct points, then C is convex if and only if the closed line interval joining any two points in C is a subset of C.

Example 11.1. The only convex subsets of $\mathbb{R}$ are intervals - including the empty set and singletons.

Example 11.2. If V is a subspace of $\mathbb{R}^n$ then V is a closed convex set. When $a \notin V$, then the set $\{a + v : v \in V\}$ is called an affine subspace of $\mathbb{R}^n$ and is a convex set.

Example 11.3. Define $H := \{x \in \mathbb{R}^n : \langle a, x \rangle = c\}$ where a is a unit vector in $\mathbb{R}^n$. Then H is called a *hyperplane* in $\mathbb{R}^n$ and it is a closed convex set.

Example 11.4. The *positive orthant* in $\mathbb{R}^n$ is the set of all vectors $x \in \mathbb{R}^n$ whose components satisfy $x_j \ge 0$ for all $j \in I_n$. The set of all strictly positive vectors in $\mathbb{R}^n$ is denoted $\mathbb{R}^n_{++} := (0, \infty)^n$. Both are convex sets in $\mathbb{R}^n$.

Let $\Gamma := \{a^{(j)} : 1 \le j \le J\}$ be a finite set of vectors in $\mathbb{R}^n$. $x \in \mathbb{R}^n$ is said to be a *convex combination* of vectors in Γ provided

$$x = \sum_{j=1}^{J} t_j \, a^{(j)} \quad \text{with each} \quad t_j \ge 0 \qquad \text{and} \quad \sum_{j=1}^{J} t_j = 1. \tag{11.1}$$

The set of all convex combinations of points in a set Γ will be denoted $co(\Gamma)$ and is called the convex hull of Γ. When Γ is a finite set, then $co(\Gamma)$ is called a *polyhedron* and it is a bounded closed convex subset of $\mathbb{R}^n$. When $J = n+1$, such a set is called a *simplex* whose vertices are Γ. For n =2, a simplex is a triangle; for $n = 3$ it is a tetrahedron.

Example 11.5. For $n \ge 2$, let $e^{(j)}$ be the j-th unit vector in $\mathbb{R}^n$. $e^{(j)}$ has k-th component δ_{jk} for $k \in I_n$. The unit simplex in $\mathbb{R}^n$ is the closed convex hull of the origin and the set of unit vectors $\{0, e^{(1)}, \ldots, e^{(n)}\}$. That is $\Delta_n := co(\{0, e^{(1)}, e^{(2)}, \ldots e^{(n)}\})$. It is the set of vectors in $\mathbb{R}^n$ whose components x_j satisfy

$$x_j \ge 0 \qquad \text{and} \quad \sum_{j=1}^{n} x_j \le 1. \tag{11.2}$$

This simplex has $n+1$ vertices and also $n+1$ faces (or sides). How many edges does it have?

The set of probability vectors in $\mathbb{R}^n$ is the closed convex hull of $\{e^{(1)}, \dots, e^{(n)}\}$. It is a face of the simplex Δ_n and an n-vector x is in Δ'_n whenever its components obey

$$x_j \geq 0 \qquad \text{and} \qquad \sum_{j=1}^{n} x_j \ = \ 1 \tag{11.3}$$

Δ'_n may be viewed as the intersection of the positive orthant with the unit sphere, with respect to the 1-norm, of $\mathbb{R}^n$. It is the only face of Δ_n that is not a subset of one of the coordinate hyperplanes $H_j := \{x \in \mathbb{R}^n : x_j = 0\}$.

When $x \in \mathbb{R}^n$, we will write $|x| := (|x_1|, |x_2|, \dots, |x_n|)$ so $|x| = \|x\|_1 \, d$ for some $d \in \Delta'_n$.

When C_1, C_2 are non-empty convex subsets of $\mathbb{R}^n$, $\lambda \in \mathbb{R}$, then the sets

$$\lambda C_1, \quad C_1 \cap C_2, \quad C_1 \ + \ C_2, \quad C_1 \ - C_2$$

will again be convex subsets of $\mathbb{R}^n$. The union of two convex sets need not be convex. When $\{C_k : k \in \mathcal{K}\}$ is an arbitrary family of convex sets in $\mathbb{R}^n$, then the intersection $\cap_{k \in \mathcal{K}} C_k$ will again be convex. If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear, or affine, transformation and C is a convex subset of $\mathbb{R}^n$, then $L(C)$ is a convex subset of $\mathbb{R}^m$.

The proofs of each of these statements are quite straightforward - but it is worthwhile verifying them yourself.

Let C be a non-empty convex set in $\mathbb{R}^n$ which contains the origin. Then the *dimension of C* is the maximum number of linearly independent vectors in C. An equivalent statement is that it is the dimension of the subspace spanned (or generated) by C. When C does not contain the origin choose $a \in C$ and let $C_1 := C - \{a\}$. C_1 is said to be a translate of C and contains the origin. The dimension of C is defined to be the dimension of C_1 and is denoted $dim\, C$. This dimension is independent of the choice of a.

Example 11.6. Define $B := \{a^{(0)} \ + \ \sum_{j=1}^{m} t_j \, d^{(j)} \ : \ 0 \leq t_j \leq 1 \text{ for } 1 \leq j \leq m\}$. B is an *m-dimensional parallelogram* in $\mathbb{R}^n$ whenever the vectors $\{d^{(j)} \ : \ 1 \leq j \leq m\}$ are linearly independent. If the vectors $d^{(j)}$ are orthogonal, then it is an *m-dimensional box* in $\mathbb{R}^n$.

The m-dimensional volume of such a box will be $|B|_m \ := \ \prod_{j=1}^{m} \ |d^{(j)}|$.

The volume of a simplex $\Delta := co(\Gamma)$ defined as above is

$$|\Delta|_n \ = \ \frac{1}{n!} \, \det[d^{(1)}, \dots, d^{(n)}]$$

where $d^{(j)} := a^{(j+1)} - a^{(1)}$ for $j \in I_n$. These vectors $d^{(j)}$ represent the edges of the simplex. This formulae generalizes the area of a triangle, or volume of a tetrahedron, to n-dimensions. If this volume is zero then the set Γ lies in a proper affine subspace of $\mathbb{R}^n$. The volume of a general set in $\mathbb{R}^n$ is usually defined to be the limit of the volumes of approximating, or covering, simplices, boxes or balls.

Example 11.7. The open and closed balls $B_r^p(a), \overline{B}_r^p(a)$ of radius r, center a in the p-norm are closed convex sets from Minkowski's inequality. When $p = 2$ they each have volume $|B| \ = \ \omega_n.(r^n/n)$. Here ω_n is the Euclidean surface area of the unit sphere in $\mathbb{R}^n$.

Exercises. .

Exercise 11.1 Suppose C is a subset of the plane defined as the interior of the continuous closed curve defined by the equation $r = f(\theta)$ in polar coordinates. Assume f is differentiable and periodic of period 2π. What conditions on f imply that C is convex?

Exercise 11.2 Suppose C is a subset of the plane containing the origin and $A : \mathbb{R}^2 \to \mathbb{R}^2$ is the map defined by $A(x) = (|x_1|, |x_2|)$, is A(C) convex? Prove this or find a counterexample.

Exercise 11.3 For $1 \leq p < \infty$, find an integral expression for the volume of the closed unit ball centered at 0 in the p-norm. When $p = 2$ show that the volume in example 11.7 has $\omega_n \; = \; 2 \, \frac{\pi^{n/2}}{\Gamma(n/2)}$ where $\Gamma(s)$ is the Gamma function.

12. Convex Functions on $\mathbb{R}^n$.

Let C be a non-empty convex set in $\mathbb{R}^n$ and $f : C \to \overline{\mathbb{R}}$ be a given function. The *essential domain* of f is

$$\mathrm{dom}(f) \;:=\; \{\, x \in C \,:\, f(x) \in \mathbb{R}\}. \tag{12.1}$$

The *epigraph of f* is the set

$$\mathrm{epi}(f) \;:=\; \{(x, z) \in C \times \mathbb{R} \,:\, z \geq f(x)\}. \tag{12.2}$$

The function f is said to be

(i) *convex* provided epi(f) is a convex set in $\mathbb{R}^{n+1}$.

(ii) *concave* provided $-f$ is convex.

(iii) *real valued* on C if $|f(x)|$ is finite for every $x \in C$. Equivalently dom(f) = C.

(iv) The function f is *proper* if dom(f) is nonempty and $f(C) \subset (-\infty, \infty]$.

This definition of a convex function avoids the use of addition so it applies to extended real valued functions and generalizes the definition of section 8.

A function $f : C \to \overline{\mathbb{R}}$ may be extended to have domain $\mathbb{R}^n$ by $f_e(x) = +\infty$ for $x \notin C$. This extension has $\mathrm{epi}(f_e) = \mathrm{epi}(f)$ - so f is convex on C if and only if f_e is convex on $\mathbb{R}^n$. Thus we often assume that convex functions are defined on all of $\mathbb{R}^n$ and take values in $\mathbb{R}^{\#} := (-\infty, \infty]$.

When f is a proper convex function on C, then dom(f) will be a nonempty convex set in $\mathbb{R}^n$ and we can regard it as being a real valued function on this convex set.

Lemma 12.1. *Let C be a non-empty convex subset in $\mathbb{R}^n$. A proper function $f : C \to (-\infty, \infty]$ is convex if and only if*

$$f((1-t)x + ty) \;\leq\; (1-t)f(x) + tf(y) \qquad \text{for all} \quad 0 \leq t \leq 1, \quad x \neq y \in \mathit{dom}(f). \tag{12.3}$$

The function f is *strictly convex* on dom(f) provided inequality holds in (12.3) whenever $0 < t < 1$. The result in Lemma 12.1 can be extended to general finite sums by induction and yields

Theorem 12.2. *Suppose $f : C \to \mathbb{R}$ is convex and $\{a^{(j)} : 1 \leq j \leq J\}$ is a finite subset of C then*

$$f\left(\sum_{j=1}^{J} t_j\, a^{(j)}\right) \;\leq\; \sum_{j=1}^{J} t_j\, f(a^{(j)}) \tag{12.4}$$

when $(t_1, t_2, \ldots, t_J) \geq 0$ in $\mathbb{R}^J$ and $\sum_{j=1}^{J} t_j = 1$.

The following are some simple examples of convex functions defined on $\mathbb{R}^n$.

Example 12.1. $l(x) \;:=\; \langle\, a, x\,\rangle \;+\; c \;:=\; c \;+\; \sum_{j=1}^{n} a_j\, x_j$ is convex on $\mathbb{R}^n$.
Here $a \in \mathbb{R}^n \backslash \{0\}, c \in \mathbb{R}$. When $c \neq 0$, such functions are said to be *affine* functions .

Example 12.2. Let $p : \mathbb{R}^n \to [0, \infty)$ be a norm on $\mathbb{R}^n$, then p is convex The p-norms for $1 < p < \infty$ are strictly convex - but the 1-norm and the $\infty-$norm are not strictly convex.

Example 12.3 Given a non-empty convex set $C \subset \mathbb{R}^n$ and $b \in \overline{\mathbb{R}}$, the function defined by

$$f(x) \;:=\; \begin{cases} b & x \in C \\ \infty & x \notin C \end{cases}$$

is a convex function on $\mathbb{R}^n$. When $b = 0$, this function is called the indicator function of the set C and is denoted I_C. This function is not strictly convex whenever $b \in \mathbb{R}$ and C has at least two points.

12.1. **Operations on Convex Functions.** For the remainder of this section C will be a nontrivial convex subset of $\mathbb{R}^n$.

1. Addition and Positive multiplication.
Suppose $f_1, f_2 := C \to \mathbb{R}$ are convex functions, then $c_1 f_1 + c_2 f_2$ is convex whenever $c_1 \geq 0, c_2 \geq 0$. In general, if $\{f_1, f_2, \ldots, f_m\}$ is a finite set of convex functions on C, then

$$f(x) \;:=\; \sum_{j=1}^{m} c_j\; f_j(x) \quad \text{is convex whenever } c_1, c_2, \ldots, c_m \geq 0.$$

In this case we say that f is a positive linear combination (p.l.c.) of $f_1, f_2, \ldots, f_m$.

In general the product of two convex functions on a convex set need not be convex.

Example 12.4 Let $I := [-1, 1],\; f(x) := x^2,\; g(x) := x$ on I. Both f, g are convex on I, but their product is not. It is convex on $[0, 1]$.

2. Suprema of Convex Functions.
Suppose $\{f_k : k \in \mathcal{K}\}$ is a collection of proper convex functions on a convex subset $C \subset \mathbb{R}^n$. Define $F : C \to (-\infty, \infty]$ by

$$F(x) \;:=\; \sup_{k \in \mathcal{K}}\; f_k(x)$$

then F is a proper convex function on C. That is, the supremum of any family of convex functions is convex. This property will be used repeatedly to show that specific functions are convex. Note that if $\mathcal{K}$ is finite and each of the functions f_k is continuous on C so is F. When $\mathcal{K}$ is infinite, then the function F will be l.s.c. on C.

3. Composition of Convex Functions.
There are a number of composition laws for convex functions. If $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with $L(C) \subset C_1$ and $g : C_1 \to \overline{\mathbb{R}}$ is convex, then so is $f(x) \ := \ g(Lx)$. The translate of a convex function will also be convex. A useful result is the following

Proposition 12.3. *Let C be a nontrivial convex set in $\mathbb{R}^n$, I is an interval in $\mathbb{R}$ and $f : C \to \mathbb{R}$ be a convex function with $f(C) \subset I$. If $\varphi : I \to \mathbb{R}$ is convex and increasing, then $g(x) \ := \ \varphi(f(x))$ is convex on C.*

Proof. Choose $x, y \in C$ and define $x(t)$ as usual. Since f is convex on C, then (12.3) holds. Apply φ to both sides of this then

$$g(x(t))) \ \leq \ \varphi((1-t)f(x)) \ + \ tf(y))$$

as φ is increasing Since φ is convex on I, g satisfies (12.3) $\square$

Example 12.5. Suppose $f : C \to [0, \infty)$ is a convex function, then $f^m : C \to [0, \infty)$ defined by $f^m(x) \ := \ f(x)^m$ will be a convex function for any $m \in \mathbb{N}$. This need not hold if the range of f is not a subset of $[0, \infty)$

A convex function $g : C \to \overline{\mathbb{R}}$ is said to be a *convex minorant* of a function $f : C \to \overline{\mathbb{R}}$ provided $g(x) \leq f(x)$ for all $x \in C$. This is equivalent to requiring that $\text{epi}(g) \supset \text{epi}(f)$. Such a g is said to be an *affine minorant* of f when g is an affine function.

When f is bounded below on I, a constant function $g(x) \equiv C$ with $C \leq \alpha := \inf_{x \in I} f(x)$ is a convex minorant of f.

Given a function $f : C \to \overline{\mathbb{R}}$, let $\Gamma(f)$ be the set of all convex minorants of f and define $\overline{f}(x) : C \to \overline{\mathbb{R}}$ by

$$\overline{f}(x) \ = \ \sup \ \{g(x) : g \in \Gamma(f)\}. \tag{12.5}$$

$\overline{f}$ is called the convex hull, or convex envelope, of f on C and is a convex function on C - and has that $\alpha(\overline{f}, C) = \alpha(f, C)$. Also $\overline{f}(x) \leq f(x)$ for all $x \in C$ (since $g(x) \leq f(x)$ for each x). Moreover, $\overline{f}$ is the largest such convex function on I.

In optimization theory it is often useful to construct the convex hull of a given function. For example, the function $g(x) := x^2$ is a convex minorant of the function $f(x)$ defined in example 6.1 that has the same value and has the same minimizer.

A number of important results about convex functions extend to the class of quasi-convex functions on C. A function $f : C \to \overline{\mathbb{R}}$ is said to be *quasi-convex* provided the synoptic sets $S_c(f) := \{x \in C : f(x) \leq c\}$ are convex for every $c \in \mathbb{R}$.

Every convex function is quasi-convex but there are many examples of quasi-convex functions which are not convex. For example the function $\sqrt{|x|}$ of example 6.2 and increasing functions such as $f(x) := x^3$ on $\mathbb{R}$ are quasi-convex, but not convex.

Exercises. .

Exercise 12.1 Show that if $f : C \to \mathbb{R}$ is convex and also concave on C, then f is an affine function on C.

Exercise 12.2 Prove lemma 12.1.

Exercise 12.3 Prove theorem 12.2.

Exercise 12.4 What are the epigraphs of the functions in examples 12.1 and 12.3?

Exercise 12.5 Prove that if f is a convex function on C, then its essential domain is also convex. Show that if f is also continuous on C then its range is an interval in $\overline{\mathbb{R}}$.

Exercise 12.6 Prove the results about sums and suprema of convex functions described above.

Exercise 12.7 Sketch the graphs of the convex hulls $\overline{f}$, of the functions $f : \mathbb{R} \to \mathbb{R}$ defined by
(i) $f(x) \;:=\; (x^2 - 1)^2$
(ii) $f(x) \;:=\; -x^2$
(iii) $f(x) \;:=\; \cos x.$
Compare the minimizers of f and $\overline{f}$ for these examples.

Exercise 12.8 Let C_1, C_2 be non-trivial convex sets in $\mathbb{R}^m, \mathbb{R}^n$ respectively. Show that the Cartesian product $C := C_1 \times C_2$, is a convex subset of $\mathbb{R}^{m+n}$.
(b) Suppose $\psi_1 : C_1 \to \mathbb{R},\ \psi_2 : C_2 \to \mathbb{R}$ are continuous convex functions and define $\psi : C \to \mathbb{R}$ by $\psi(x, y) := \psi_1(x) + \psi_2(y)$ where $x \in C_1,\ y \in C_2$. Show that ψ is convex on C.
(c) Show that the map $F : C \to \mathbb{R}^2$ defined by $F(x, y) := (\psi_1(x), \psi_2(y))$ has a convex range.

13. Multivariate Differentiation

To describe optimization theory for multivariate functions we need a good theory of multivariate differentiation. A good description of the classical theory for undergraduates is given in Chapter 3 of Fleming [3] A more extensive and thorough treatment, including important counterexamples is given in Chapter 3 of Ortega and Rheinboldt [4]. Only the essential definitions and results needed for this course will be described here.

Throughout the next few sections, U is an open set in $\mathbb{R}^n,\ x^{(0)} \in U$ and $f : U \to \mathbb{R}$ is assumed to be a continuous function. A *direction* in $\mathbb{R}^n$ is a (Euclidean) unit vector in $\mathbb{R}^n$. The set of all directions is the unit sphere $S_1 := \{x \in \mathbb{R}^n : \|x\|_2 \;=\; 1\}$.

We say that f has a *derivative* at $x^{(0)} \in U$ in the direction h, provided there is a real number d such that

$$\lim_{t \to 0+} t^{-1} \left[f(x^{(0)} + th) - f(x^{(0)})\right] \;=\; d. \tag{13.1}$$

When $h = e^{(j)}$, this derivative is called the *j-th partial derivative* of f at $x^{(0)}$ and will be denoted $D_j f(x^{(0)})$. Note that, for our purposes here, derivatives are just "one-sided" derivatives. That is, they generalize D_+ and D_- from the one-variable case.

f is said to be *G(ateaux)-differentiable* at $x^{(0)}$ provided there is a vector $v \in \mathbb{R}^n$ such that

$$\lim_{t\to 0+} t^{-1}\ \left[\left(f(x^{(0)}+th) - f(x^{(0)})\right)\right]\ -\ v\cdot h\ =\ 0 \quad \text{for all } h \in S_1. \tag{13.2}$$

When this holds, $\nabla f(x^{(0)})\ :=\ v$ is called the *gradient* of f at $x^{(0)}$. The jth component of $\nabla f(x^{(0)})$ is the j-th partial derivative $D_j f(x^{(0)})$. Note that if this holds, then the derivative may be taken in "all-directions" as h is an arbitrary (direction)-vector in S_1.

Functions of the form $\varphi(t) := f(x^{(0)}+th)$ defined on an interval $(-\delta_1, \delta_2)$ which includes 0 will be used quite often. When f is G-differentiable at $x^{(0)}$ then, φ is differentiable at 0 for any $h \in S_1$ and from (13.2)

$$\varphi'(0) = \nabla f(x^{(0)}) \cdot h. \tag{13.3}$$

As usual the graph of the function f is the set $G(f) := \{(x, f(x)) : x \in U\}$. When f is G-differentiable at $x^{(0)}$, then the *tangent hyperplane* to this graph at $x^{(0)}$ is the graph of the function

$$z\ =\ f(x^{(0)})\ +\ \langle\, \nabla f(x^{(0)}), x\ -\ x^{(0)}\,\rangle. \tag{13.4}$$

The graph of $z\ =\ l(x)\ :=\ f(x^{(0)})\ +\ \langle\, a, x\ -\ x^{(0)}\,\rangle$ is said to be a *support hyperplane* for the (graph of the) function f at $x^{(0)}$ provided $f(x)\ \geq\ l(x)$ for all $x \in U$.

A vector $\hat{x}$ in U is said to be a *critical point* of f provided f is G-differentiable at $\hat{x}$ and $\nabla f(\hat{x}) = 0 \in \mathbb{R}^n$.

The properties of G-derivatives include linearity and a product rule. However a function may be G-differentiable at a point $x^{(0)}$ without being continuous there and the general chain rule does not hold. See page 61ff of [3] for discussions of this

Let $F : U \to \mathbb{R}^m$ be a continuous vector valued function on the open set U. F is said to be *G-differentiable* at $x^{(0)}$ provided there is a $m \times n$ matrix J, such that

$$\lim_{t\to 0^+}\ \|\ t^{-1}[F(x^{(0)}+th) - F(x^{(0)})]\ -\ Jh\|\ =\ 0 \quad \text{for all } h \in S_1. \tag{13.5}$$

Here any norm on $\mathbb{R}^m$ may be used. When this holds, we write $DF(x^{(0)})\ :=\ J$ and this matrix is called the Jacobian of F at $x^{(0)}$. Suppose

$$F(x) = \begin{pmatrix} F_1(x_1, x_2, ..., x_n) \\ F_2(x_1, x_2, ..., x_n) \\ \cdots \\ F_m(x_1, x_2, ..., x_n) \end{pmatrix} \tag{13.6}$$

When F is differentiable at $x^{(0)}$, then each $D_k F_j(x^{(0)})$ is finite and

$$DF(x^{(0)}) = \begin{pmatrix} D_1F_1(x^{(0)}) & D_2F_1(x^{(0)}) & \dots & D_nF_1(x^{(0)}) \\ D_1F_2(x^{(0)}) & D_2F_2(x^{(0)}) & \dots & D_nF_2(x^{(0)}) \\ \dots & \dots & \dots & \dots \\ D_1F_m(x^{(0)}) & D_2F_m(x^{(0)}) & \dots & D_nF_m(x^{(0)}) \end{pmatrix}. \tag{13.7}$$

For the particular case when $F(x) = \nabla f(x) = (D_1 f(x), \dots, D_n f(x))^T$ this derivative matrix is called the Hessian of f and

$$DF(x) \;=\; D^2 f(x) \;=\; \begin{pmatrix} D_{11}f(x) & D_{21}f(x) & \dots & D_{n1}f(x) \\ D_{12}f(x) & D_{22}f(x) & \dots & D_{n2}f(x) \\ \dots & \dots & \dots & \dots \\ D_{1n}f(x) & D_{2n}f(x) & \dots & D_{nn}f(x) \end{pmatrix} \tag{13.8}$$

is an $n \times n$ matrix. The entry $D_{jk}f(x) := D_j(D_k f)(x)$ is the j-th partial derivative of the function $D_k f$.

In this case equation (13.5) becomes

$$\lim_{t \to 0^+} \; \| \, t^{-1}[\nabla f(x^{(0)} + th) - \nabla f(x^{(0)})] \; - \; D^2 f(x^{(0)})\, h \| \;=\; 0 \quad \text{for all } h \in S_1. \tag{13.9}$$

Take inner products with h here, then

$$t^{-1}\,[\nabla f(x^{(0)} + th) - \nabla f(x^{(0)})] \cdot h \to \langle\, D^2 f(x^{(0)})\, h, h \,\rangle$$

as $t \to 0$ since (13.9) holds for both $h, -h$,

When $\varphi(t) := f(x^{(0)} + th)$ is substituted here, this and (13.3) yield that

$$\varphi''(0) = \langle\, D^2 f(x^{(0)})\, h, h \,\rangle. \tag{13.10}$$

Another result from multivariate calculus is that when f is twice continuously differentiable on a neighborhood of x, then $D^2 f(x)$ will be a symmetric $n \times n$ matrix and

$$D_{jk}f(x) \;=\; D_{kj}f(x) \quad \text{for all} \quad j, k \in I_n \tag{13.11}$$

See theorem 3.3 of Fleming [3] for a proof. Sections 3.5 and 3.6 of [3] also provide a different version of the following material.

Exercises. .

Exercise 13.1 Suppose $r : \mathbb{R}^3 \to [0, \infty)$ is the Euclidean radius function $r(x) := \|x\|_2$. Let $\psi : [0, \infty) \to \mathbb{R}$ be a continuous function that is C^2 on $(0, \infty)$ and let $f(x) := \psi(r(x))$.
(a) Evaluate $\nabla r(x)$ and $D^2 r(x)$ for $x \in \mathbb{R}^3 \setminus \{0\}$.
(b) Find formulae for $\nabla f(x),\ D^2 f(x)$ in terms of derivatives of ψ for $x \in \mathbb{R}^3 \setminus \{0\}$.
(c) Prove that $f(x) := r(x)^p$ is convex on $\mathbb{R}^3$ for $1 \le p < \infty$ and evaluate this $\nabla f(x)$.

Exercise 13.2 (a) Suppose $f(x) := \|x\|_p^p$ with $p \in (1, \infty)$. Find a formula for the gradient of this function. When is this function C^1- on $\mathbb{R}^n$?
(b) Suppose $f(x) := \|x\|_1$ with $x \in \mathbb{R}^3$. Describe the set of points where this function is not G-differentiable. Find an explicit expression for the gradient of this function at points where it is differentiable.

Exercise 13.3 A real valued function on $\mathbb{R}^n$ is said to be homogeneous of degree p provided $f(cx) \;=\; |c|^p\, f(x)$ for all $x \in \mathbb{R}^n, c \in \mathbb{R}$. If f is also G-differentiable on $\mathbb{R}^n$, prove Euler's rule that

$$\langle\, \nabla f(x), x \,\rangle \;=\; p\, f(x) \qquad \text{for all } x \in \mathbb{R}^n.$$

14. Multivariate Minimization

Just as for one dimensional optimization, there are necessary, and also sufficient, conditions for a point in U to be a local minimizer of f on U. The necessary conditions may be expressed in terms or either the lower variation of f or, when f is differentiable, in terms of the derivative of f.

The *lower variation* of f at $x^{(0)}$ in the direction $h \in S_1$ is

$$\delta_l f(x^{(0)}, h) \ := \ \liminf_{t \to 0^+} \ t^{-1} \ [f(x^{(0)} + th) - f(x^{(0)})]. \tag{14.1}$$

This always exists and may be $\pm\infty$. If f has a derivative d at $x^{(0)}$ in the direction h, then $\delta_l f(x^{(0)}, h) = d$.

Theorem 14.1. *(First Order Necessary conditions) Suppose $f : U \to \mathbb{R}$ is continuous $\hat{x}$ is a local minimizer of f on U, then*
(i) $\delta_l f(\hat{x}, h) \geq 0$ for all $h \in S_1$, and
(ii) if f is G-differentiable at $\hat{x}$, then $\nabla f(\hat{x}) = 0$. *(F)*

Proof. (i) If $\hat{x}$ is a local minimizer, then $f(\hat{x} + th) \geq f(\hat{x})$ for all $h \in S_1$ and t small enough. Thus

$$t^{-1} \ [f(\hat{x} + th) - f(\hat{x})] \geq 0 \qquad \text{for} \quad 0 < t < \delta(h)$$

so $\liminf_{t \to 0^+} \ t^{-1}[f(\hat{x} + th) - f(\hat{x})] \ \geq \ 0$ or (i) holds.

(ii) If f is G-differentiable at $\hat{x}$, then part (i) implies that $\nabla f(\hat{x}) \cdot h \geq 0$ for each $h \in S_1$. Take $h = \pm e^{(j)}$, then

$$D_j f(\hat{x}) \ \geq \ 0 \quad \text{and} \quad D_j f(\hat{x}) \ \leq \ 0, \quad \text{so} \quad D_j F(\hat{x}) = 0$$

for each $j \in I_n$. □

Note that (ii) here does not require that f be differentiable at any point except $\hat{x}$.

In view of theorem 14.1, the only possible minimizers of a G-differentiable function f on an open set U occur at critical points of f on U. For $n \geq 2$ the level sets of f near a critical point $\hat{x}$ may be very complicated in general. A critical point $\tilde{x}$ of f on U may be either a local minimizer, a saddle point or a local maximizer of f on U - just as for univariate functions.

Also we say that a critical point $\hat{x}$ of f is a *degenerate critical point* provided either
(i) $D^2 f(\hat{x})$ does not exist, or else
(ii) the matrix $D^2 f(\hat{x})$ is singular.

When $D^2 f(\hat{x})$ exists and is a non-singular matrix then $\hat{x}$ is said to be a *non-degenerate critical point.* of f.

When f is twice differentiable at the critical point with a non-singular Jacobian, however, the Hessian matrix provides some effective criteria for local minimizers.

Theorem 14.2. *(2nd order Necessary condition) Suppose $\hat{x}$ is a local minimizer of f on U and f is C^1 on an open neighborhood of $\hat{x}$ in U. If $D^2 f(\hat{x})$ exists, then*

$$\langle \, D^2 f(\hat{x}) h, h \, \rangle \ \geq \ 0 \qquad \textit{for all } h \in \mathbb{R}^n. \tag{14.2}$$

Proof. Put $\varphi(t) = f(\hat{x} + th)$ for some $h \in S_1$. $\hat{x}$ is a local minimizer of f implies that 0 is a local minimizer of φ. From theorem 7.2, if $\varphi''(0)$ exists, then it is ≥ 0. The analysis of the second G-derivative above yields that $\varphi''(0) \; = \; \langle\, D^2 f(\hat{x})h, h \,\rangle$ so (14.2) follows. □

Theorem 14.3. *(Sufficient Condition) Suppose f is C^1 on a neighborhood of a critical point $\hat{x}$, $D^2 f(\hat{x})$ exists and there is a $c_1 > 0$ such that*

$$\langle\, D^2 f(\hat{x})h, h \,\rangle \;\geq\; c_1 ||h||^2 \quad \text{for all} \quad h \in \mathbb{R}^n. \tag{14.3}$$

Then $\hat{x}$ is an isolated, strict local minimizer of f on U.

Proof. Choose φ as above, then (14.3) implies

$$\varphi''(0) \;\geq\; c_1 ||h||^2 \;>\; 0$$

Just as in the proof of theorem 7.3, when $\hat{x}$ is a critical point of f,

$$\varphi(t) - \varphi(0) \;\geq\; c_1 t^2/4 \quad \text{since} \quad |\varphi'(t) - \varphi'(0)| \;\geq\; \frac{c_1}{2}|t| \quad \text{for } 0 < |t| < \delta.$$

$$\text{Thus} \qquad \varphi(t) - \varphi(0) \;=\; \int_0^t \varphi'(\tau) d\tau \;\geq\; \frac{c_1}{4} t^2 \quad \text{for } t > 0.$$

Similarly when $t < 0$, so $f(\hat{x} + th) - f(\hat{x}) \;\geq\; c_1 t^2/4$ for each direction $h \in S_1$ and $0 < |t| < \delta$. Thus $\hat{x}$ is a strict local minimizer as claimed. Since $\nabla f(\hat{x}) = 0$ and $\langle\, D^2 f(\hat{x})h, h \,\rangle \;\geq\; c_1 ||h||^2$, then $\hat{x}$ is an isolated critical point. □

Comments 1. The necessary and sufficient conditions for local maximizers require the reverse inequalities in (14.2) or (14.3).

Exercises. .

Exercise 14.1 Define a function $G : \mathbb{R}^n \to \mathbb{R}$ by

$$G(x) \;:=\; \|\, x \,\|_2{}^4 \;-\; 2\,\langle\, Ax, x \,\rangle$$

where A is an $n \times n$ symmetric matrix. Prove
(i) Prove that this function is bounded below and has minimizers on $\mathbb{R}^n$.
(ii) Find the equation satisfied by the critical points of G on $\mathbb{R}^n$.
(iii) What mathematical properties can you say about the critical points and/or minimizers of G? What can you say about the value of this problem?

15. Convex Functions and Convex Minimization.

Let C be a non-empty, open convex set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ be G-differentiable at each point in C. Then $\nabla f : C \to \mathbb{R}^n$ is defined. The following results provide differential criteria for the function to be convex and extend the 1-dimensional results described in section 8.

Theorem 15.1. *Suppose f, C as above, then f is convex on C if and only if*

$$f(y) \;\geq\; f(x) \;+\; \langle\, \nabla f(x), y - x \,\rangle \quad \text{for all } x, y \in C. \tag{15.1}$$

Proof. Suppose f is convex, $x, y \in C, x \neq y$ and $x(t) := (1-t)x + ty$ with $0 < t < 1$. Then

$$f(x(t)) \;\leq\; (1-t)f(x) + tf(y)$$

Rearranging this with $0 < t \leq 1$ leads to

$$t^{-1}\,[f(x(t)) - f(x)] \;\leq\; f(y) - f(x).$$

Take limits as $t \to 0^+$, then

$$\langle\, \nabla f(x), y - x \,\rangle \;\leq\; f(y) - f(x)$$

so (15.1) holds.

Conversely when (15.1) holds, $x, y \in C, x \neq y$, let $z := (1-t)x + ty$ with $0 < t < 1$. Substitute z for x in (15.1), then

$$f(x) \;\geq\; f(z) \;+\; \langle\, \nabla f(z), x - z \,\rangle \quad \text{and } f(y) \;\geq\; f(z) \;+\; \langle\, \nabla f(z), y - z \,\rangle.$$

Multiply first equation by $(1-t)$, second by t and add to find that the convexity inequality (9.3) holds. □

Corollary 15.2. *Suppose $z = l(x) \;=\; f(x^{(0)}) + \langle\, a, x \;-\; x^{(0)} \,\rangle$ is a support hyperplane for the graph of f at $x^{(0)} \in C$. If f is convex on C, then $f(y) \;\geq\; l(y)$ for all $y \in C$.*

Proof. □

The following only requires a simple modification of this proof.

Corollary 15.3. *Suppose f, C as above, then f is strictly convex on C if and only if strict inequality holds in (15.1) when $x \neq y$.*

The next two theorems give the conditions that are usually used to check whether a particular differentiable function is convex on C.

Theorem 15.4. *Suppose f, C as above, then f is convex on C if and only if*

$$\langle\, \nabla f(y) - \nabla f(x), y - x \,\rangle \;\geq\; 0 \quad \textit{for all } x, y, \in C. \tag{15.2}$$

It is strictly convex on C if strict inequality holds here when $y \neq x$.

Proof. Suppose f is convex, $x, y \in C, x \neq y$ and let $\varphi(t) := f(x(t))$. Then φ is a convex differentiable function of t and $\varphi'(t) \;=\; \langle\, \nabla f(x(t)), y - x \,\rangle$ from (13.3). From 1-d theory, $\varphi'(1) \geq \varphi'(0)$ so (15.2) holds.

Conversely, when (15.2) holds, then $\varphi'(t) \geq \varphi'(0)$ for all $t > 0$. Then

$$\varphi(t) \;=\; \varphi(0) \;+\; \int_0^t \varphi'(s)\, ds \;\geq\; \varphi(0) \;+\; \varphi'(0)t.$$

That is

$$f(x(t)) \;\geq\; f(x) \;+\; t\,\langle\, \nabla f(x), y - x \,\rangle.$$

This implies that (15.1) holds, so f is convex on C. The strictness part is similar. □

When (15.2) holds, then $\nabla f(x)$ is said to be a *monotone* mapping of C into $\mathbb{R}^n$.

The basic results about 1-dimensional minimization and maximization of convex functions on a interval were summarized previously in theorem 9.1. Here the corresponding results for multivariate convex functions f on C will be given.

A general existence result is the following.

Theorem 15.5. *Suppose C is a nonempty compact convex set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ is lower semi-continuous and quasi-convex. Then*
(i) the set of all minimizers of f on C is a nonempty closed convex subset of C.
(ii) If f is strictly convex on C, then this set consists of exactly one point.

Proof. The existence follows from Weierstrass' theorem 5.1. The fact that each synoptic set is convex and closed means that the set $S_c(f)$ is convex and closed so the set of all minimizers of f on C is this set with $c = \alpha(f, C)$ so (i) holds. (ii) is the standard argument. □

A very useful version of this is the following "unconstrained" version. Its proof just uses the same argument as in the 1 variable case.

Theorem 15.6. *Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is proper, lower semi-continuous, quasi-convex and weakly coercive. Then $\alpha(f) := \inf_{x \in \mathbb{R}^n} f(x)$ is finite and the set of minimizers of f on $\mathbb{R}^n$ is non-empty, bounded, closed and convex.*

The following theorem says that, when f is differentiable and convex on an open convex set C then the only critical points of f on C are the minimizers of f on C. Alternatively being a critical point of f is a necessary and sufficient condition to be a minimizer - on an open set - when f is convex and differentiable everywhere.

Theorem 15.7. *Suppose C is a non-empty open convex set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ is convex and G-differentiable on C. A vector $\hat{x} \in C$ minimizes f on C if and only if $\hat{x}$ is a critical point of f.*

Proof. When $\hat{x}$ is a critical point of f, then (15.1) implies that $f(y) \geq f(\hat{x})$ for all $y \in C$. Thus $\hat{x}$ minimizes f on C.

Conversely, if $\hat{x}$ minimizes f on C, then $f(\hat{x} + td) \geq f(\hat{x})$ for all $t > 0, d \in S_1$. Thus $\langle \nabla f(\hat{x}), d \rangle \geq 0$ for all $d \in S_1$. This implies $\hat{x}$ is a critical point of f. □

A number of important mathematical problems can be written as convex optimization problems. The important questions about each of them usually include
(i) does the optimization problem have a finite value and a minimizer?
(ii) what equations do the (local) minimizers satisfy? and
(iii) how can we find the minimizers and/or critical points?

To answer (i) we usually show that f is continuous (or l.s.c.) and coercive - or that some synoptic set is non-empty, closed and bounded. To answer (ii) just find the G-derivatives while (iii) often involves developing algorithms for finding the minimizers.

Another treatment of this material is in section 3.6 of Fleming [3]. Chapter III, Section 3 of Berkowitz [1] has much more information about many of these issues.

Exercises. .

Exercise 15.1 Suppose $S := \{a^{(1)}, \dots, a^{(m)}\}$ is a finite set of distinct points in space and $F : \mathbb{R}^3 \to [0, \infty)$ is defined by

$$F(x) \ := \ \sum_{j=1}^{m} c_j \ \| x - a^{(j)} \|_2^2$$

with each $c_j > 0$.
(a) Show that F is convex and coercive and that there is a unique minimizer of this function.
(b) What equations hold at the minimizer of this problem?
(c) Suppose $S = \{0, e^{(1)}, e^{(2)}, e^{(3)}\}$ and each $c_j = 1$. Find the solution of this problem.

Exercise 15.2 Suppose $\psi : [0, \infty)^n \to \mathbb{R}_+$ is a continuous, convex function with $\psi(0) = 0$ and $\psi(x) > 0$ for $x \neq 0$.
(a) Show that there are constants $c_2 \ \geq \ c_1 > 0$ such that

$$\psi(x) \ \geq \ c_1 \, \|x\|_1 \qquad \text{and} \quad \psi(x) \ \geq \ c_2 \, \|x\|_\infty \quad \text{when} \quad \|x\|_\infty \ \geq \ 1. \tag{15.3}$$

and that ψ is weakly coercive on $[0, \infty)^n$.
(b) Let $A : \mathbb{R}^n \to [0, \infty)^n$ be the map defined by $A(x) := (|x_1|, |x_2|, \dots, |x_n|)$. Define $p : \mathbb{R}^n \to [0, \infty)$ by

$$p(x) \ := \ \inf \{ s > 0 : \psi(A(x)/s) \ \leq \ 1 \}$$

Show that p is a norm on $\mathbb{R}^n$.
(c) Take $n = 2$ and $\psi(x_1, x_2) := e^{x_1} + e^{x_2} - 2$. Find the equation satisfied by the set B of vectors $x \in \mathbb{R}^2$ that obey $p(x) \ \leq \ 1$ with p as in (b). Is B a subset of the unit ball with respect to the $\infty-$norm? Is the unit ball with respect to the $\infty-$norm a subset of B?

16. Matrices, Norms and Quadratic Forms.

Let $A := (a_{jk})$ be an $m \times n$ real matrix. Its transpose is the $n \times m$ real matrix with $A^T := (a_{kj})$. That is the rows and columns of A are interchanged. The set of all $m \times n$ real matrices will be denoted by M_{mn} and is a real vector space. For $p \in [1, \infty]$, the p-norm of a matrix A is defined by

$$\| \, A \, \|_p \ := \ \sup_{\| \, x \, \|_p = 1} \ \| \, Ax \, \|_p. \tag{16.1}$$

It is a good exercise (see Ex 16.1 below) to prove that this is a norm on M_{mn}. In general there is no explicit formula for this norm in terms of the entries in A - even when $p = 2$. It is a number defined as the maximum of a convex function on the (non-convex) compact set S_{1p} in $\mathbb{R}^n$. This can be changed to a convex domain by modifying the constraint set to the convex hull, namely $B_{1p} := \{x : \| \, x \, \|_p \leq 1\}$.

When $m = n$, the set of all $n \times n$ real matrices will be denoted M_n and is a (non-commutative) algebra. An $n \times n$ matrix A is said to be *symmetric* if $A^T = A$, it is skew-symmetric if $A^T = -A$. The quadratic form associated with a square matrix A is the function $q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$q(x) \;:=\; \langle\, Ax, x \,\rangle \;:=\; \sum_{j,k=1}^{n} \; a_{jk}\, x_j\, x_k. \tag{16.2}$$

It is straightforward to evaluate the partial derivatives of q and find that

$$\nabla q(x) \;=\; (A \;+\; A^T)\, x \qquad \text{and also} \qquad D^2 q(x) \equiv A \;+\; A^T \tag{16.3}$$

with A^T being the transpose matrix of A. When A is a skew-symmetric matrix then $\nabla q(x) \equiv 0$ on $\mathbb{R}^n$ so $q(x) \equiv 0$ on $\mathbb{R}^n$.

The symmetric part of A is the matrix $A_s \;:=\; (A \;+\; A^T)/2$, so (16.3) becomes

$$\nabla q(x) \;=\; 2A_s\, x, \qquad D^2 q(x) \equiv 2A_s.$$

Observe that the quadratic forms associated with A and A_S are identical. That is $q(x) \equiv q_s(x)$ where $q_s(x) \;:=\; \langle\, A_s\, x, x \,\rangle$. So, without loss of generality, A can be assumed to be real symmetric.

A real symmetric $n \times n$ matrix A is defined to be *positive semi-definite* (p.s.d.) if $q(x) \;\geq\; 0$ for all $x \in \mathbb{R}^n$. It is *positive definite* (p.d.) if $q(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

The following theorem says that the convexity of q is determined by the positivity, or otherwise, of q on $\mathbb{R}^n$.

Theorem 16.1. *Let A be a real symmetric $n \times n$ matrix. The function q defined by (16.2) is convex if and only if A is positive semi-definite. q is strictly convex if and only if A is a positive definite matrix.*

Proof. From theorem 15.1 q will be convex on $\mathbb{R}^n$ if and only if

$$q(y) \;\geq\; q(x) \;+\; 2\,\langle\, Ax, y - x \,\rangle \qquad \text{for all } x, y, \in \mathbb{R}^n.$$

Substitute for q, then this holds if and only if

$$\langle\, A(y - x), y - x \,\rangle \;\geq\; 0 \qquad \text{for all } x, y, \in \mathbb{R}^n.$$

Write $z = y - x$, then this becomes $\langle\, Az, z \,\rangle \;\geq\; 0 \qquad$ for all $z \in \mathbb{R}^n$ - or the matrix A is p.s.d.

From the last part of theorem 15.4, the function q will be strictly convex on $\mathbb{R}^n$ provided $\langle\, Az, z \,\rangle > 0 \qquad$ for all $z \in \mathbb{R}^n \setminus \{0\}$. $\square$

Consider the bilinear form $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$B(x, y) \;:=\; [x, y]_A \;:=\; \langle\, Ax, y \,\rangle. \tag{16.4}$$

If A is positive definite then this bilinear form will actually be an inner product on $\mathbb{R}^n$ and the following holds.

Corollary 16.2. *Suppose A be a positive definite symmetric $n \times n$ matrix. Then (16.4) defines an inner product on $\mathbb{R}^n$ and there are constants $0 < c_1 \leq c_2$ such that*

$$c_1 \,\|\, x \,\|_2^2 \;\leq\; [x, x]_A = q(x) \;\leq\; c_2 \,\|\, x \,\|_2^2 \qquad \textit{for all } x \in \mathbb{R}^n. \tag{16.5}$$

In particular q is strictly convex and coercive on $\mathbb{R}^n$.

Proof. To show that (16.4) defines an inner product we need only verify that $[x,x]_A \geq 0$ and equality here holds only when $x = 0$. This is just the requirement that A be p.d. The quadratic form q is a continuous function on $\mathbb{R}^n$. Consider the problem of minimizing, or maximizing, q on the unit sphere S_1. These minimum and maximum values exist and are finite from Weierstrass' theorem as S_1 is compact. Let them be c_1, c_2 respectively. One sees that $c_1 > 0$ as otherwise the matrix A would not be p.d.

(16.5) holds for $x = 0$. Suppose $x \neq 0$ and define $e := x/|x|$. Then $c_1 \leq [e,e]_A \leq c_2$ from the definitions of c_1, c_2. Multiply both sides by $|x|^2$ and (16.5) follows. □

Essentially this says that when A is a positive definite symmetric matrix, this inner product (16.4) and the corresponding norm are "equivalent" to the Euclidean inner product and norm.

Suppose now that $f : C \to \mathbb{R}$ is twice continuously differentiable (C^2-) on C and define $\varphi(t) := f(x(t))$ - as in the proof of theorem 15.4. Then

$$\varphi'(t) \;=\; \langle \nabla f(x(t)), y - x \rangle \quad \text{and} \quad \varphi''(t) \;=\; \langle D^2 f(x(t))(y-x), (y-x) \rangle.$$

From the 1-d Taylor's theorem for φ, one has

$$f(y) \;=\; f(x) \;+\; \langle \nabla f(x), y - x \rangle \;+\; (1/2)\langle D^2 f(x(\tau))(y-x), (y-x) \rangle \tag{16.6}$$

for some $\tau \in (0,1)$.

This leads to the following *second derivative criterion* for convexity of a function. It is theorem 3.3 of Berkowitz chapter III - and a proof is given there.

Theorem 16.3. *Suppose f, C as above with f of class C^2- on C. Then f is convex on C if and only if $D^2 f(x)$ is p.s.d on C. If $D^2 f(x)$ is p.d. on C, then f is strictly convex on C.*

16. Exercises.

.

Exercise 16.1: Given $p \in [1,\infty]$, prove that the function $f_p : M_{mn} \to [0,\infty)$ defined by $f_p(A) \;:=\; \| A \|_p$ is both a norm and a convex function on M_{mn}.

Exercise 16.2: Suppose f is a real valued function on $\mathbb{R}^n$ that is homogeneous of degree $r > 0$. If there is an $e \in S_1$ such that $f(e) > 0$, prove that

$$\sup_{\| x \|_p \,\leq\, 1} f(x) \;=\; \sup_{\| x \|_p \,=\, 1} f(x).$$

Show that either $f(x) \;\geq\; 0$ for all $x \in \mathbb{R}^n$ or else

$$\inf_{\| x \|_p \,\leq\, 1} f(x) \;=\; \inf_{\| x \|_p \,=\, 1} f(x).$$

Exercise 16.3: (a) Show that if $A \in M_{mn}$, then

$$\| A \|_1 \;=\; \sup_{\|x\|_1 \leq 1} \; \sup_{\|y\|_\infty \leq 1} \; \langle Ax, y \rangle.$$

(b) Find an explicit formula for the 1-norm of the matrix A.
(c) Find conditions on the entries in A for $\| A^T \|_1 = \| A \|_1$.

Exercise 16.4: Given $A \in M_{mn}$ prove that

$$\| A \|_2 = \sup_{\|x\|_2 \le 1} \sup_{\|y\|_2 \le 1} \langle Ax, y \rangle, \qquad \text{and} \qquad \| A^T \|_2 = \| A \|_2.$$

Exercise 16.5: Given that A, B are $n \times n$ matrices and $p \in [1, \infty]$, prove that

$$\| A B \|_p \le \| A \|_p \, \| B \|_p.$$

17. Energy Principles for Linear Equations

This section describes an optimization problem that yields existence, uniqueness and solvability criteria for linear equations involving a symmetric $n \times n$ matrices. Suppose we want to find solutions of the system

$$Ax = b \tag{LEq}$$

where A is an $n \times n$ symmetric matrix. Define $\mathcal{E} : \mathbb{R}^n \to \mathbb{R}$ by

$$\mathcal{E}(x) := \langle Ax, x \rangle - 2 \langle b, x \rangle = \sum_{j,k=1}^{n} a_{jk} x_j x_k - 2 \sum_{j=1}^{n} b_j \, x_j. \tag{17.1}$$

The *energy principle* for (LEq) is the problem of finding the minimizers of $\mathcal{E}$ on $\mathbb{R}^n$ and

$$\alpha(\mathcal{E}) := \inf_{x \in \mathbb{R}^n} \mathcal{E}(x).$$

This function is continuous and has gradient $\nabla \mathcal{E}(x) = 2(Ax - b)$, so $\hat{x}$ is a critical point of $\mathcal{E}$ if and only if it is a solution of (LEq). The following result justifies the study of this optimization problem.

Lemma 17.1. *Assume A is a positive semi-definite symmetric matrix and $\mathcal{E}$ is defined by (17.1). A vector $\hat{x} \in \mathbb{R}^n$ minimizes $\mathcal{E}$ on $\mathbb{R}^n$ if and only if $\hat{x}$ is a solution of (LEq).*

Proof. When A is p.s.d then the functional $\mathcal{E}$ will be convex and continuously differentiable on $\mathbb{R}^n$ with $\nabla \mathcal{E}(x) = 2(Ax - b)$. Then theorem 15.7 yields that $\hat{x}$ minimizes $\mathcal{E}$ if and only if (LEq) holds. □

It is reasonable to ask what happens when A is not positive definite? In the case when A is not p.s.d, there is a $d \in \mathbb{R}^n$ such that $\langle Ad, d \rangle < 0$. Then $\mathcal{E}(td) \to -\infty$ as $t \to \infty$. Thus the value of the problem is $-\infty$ and there is no minimizer of $\mathcal{E}$ The function $\mathcal{E}$ is non-convex as the quadratic part is non-convex from theorem 16.1.

When A is a positive definite symmetric matrix, the results we have obtained lead to the following existence-uniqueness theorem for the linear equation (LEq).

Theorem 17.2. *Let A be a positive definite symmetric matrix, then there is a unique minimizer $\hat{x}$ of $\mathcal{E}$ on $\mathbb{R}^n$. This $\hat{x}$ is the unique solution of (LEq) and the matrix A has an inverse A^{-1} that is positive definite and symmetric with*

$$\| A^{-1} b \|_2 \ \leq \ c_1^{-1} \| b \|_2 \quad \textit{and} \qquad \alpha(\mathcal{E}) \ = \ \langle A^{-1}b, b \rangle. \tag{17.2}$$

Proof. When A is positive definite then Corollary 16.2 says that there is a $c_1 > 0$ such that

$$\begin{aligned} \langle Ax, x \rangle \ &\geq \ c_1 \| x \|_2^2 \qquad \text{for all} \quad x \in \mathbb{R}^n. \qquad \text{Thus} \\ \mathcal{E}(x) \ &\geq \ c_1 \|x\|_2^2 \ - \ \|b\| \, \|x\|_2 \qquad \text{for all} \quad x \in \mathbb{R}^n, \end{aligned}$$

so $\mathcal{E}$ is coercive on $\mathbb{R}^n$. From corollary 5.4, there is an $\hat{x} \in \mathbb{R}^n$ that minimizes $\mathcal{E}$ on $\mathbb{R}^n$. When A is p.d., then $\mathcal{E}$ is strictly convex from theorem 16.1, and theorem 15.5 implies that this minimizer is unique. The previous lemma 17.1 yields that $\hat{x}$ will be the unique solution of (LEq).

Choose $b = e^{(j)}$, and denote the unique minimizer of the associated $\mathcal{E}$ to be $\xi^{(j)}$ for $j \ = \ 1, 2, \ldots n$ respectively. Let A^{-1} be the matrix whose column vectors are these $\xi^{(j)}$. It is straightforward to verify that the solution of (LEq) will be $\hat{x} = A^{-1}b$ with this specification of A^{-1}. Suppose (LEq) holds and $A\tilde{x} = d$, then

$$\langle A^{-1}b, d \rangle \ = \ \langle \hat{x}, d \rangle \ = \ \langle \hat{x}, A\tilde{x} \rangle \ = \ \langle A\hat{x}, \tilde{x} \rangle$$

using the definitions and the fact that A is symmetric. The last inner product equals $\langle b, A^{-1}d \rangle$ so A^{-1} is symmetric.

Take inner products of (LEq) with $\hat{x}$, then $\langle A\hat{x}, \hat{x} \rangle \ = \ \langle b, \hat{x} \rangle$. Use the first inequality in (16.5) and Cauchy's inequality to obtain

$$c_1 \| x \|_2^2 \ \leq \ \| b \|_2 \, \| \hat{x} \|_2$$

This yields the first inequality in (17.2). Also observe that

$$\mathcal{E}(\hat{x}) \ = \ - \langle b, \hat{x} \rangle \ = \ - \langle A^{-1}b, b \rangle$$

so the value of the problem is as claimed. When $b \neq 0$, then this value must be negative as if the value is zero then 0 would be a minimizing vector. Hence A^{-1} is positive definite. $\square$

It is worth noting that this is a proof that (LEq) has a solution for any $b \in \mathbb{R}^n$ based on optimization theory and analysis. No non-trivial results from linear algebra have been used - just formulae for derivatives and properties of convexity.

When A is p.s.d but not p.d., then $\mathcal{E}$ will be convex and the a more careful analysis is needed. Let $q(x) := \langle Ax, x \rangle$ be the quadratic form defined as in (16.2) and consider the problem of minimizing q on $\mathbb{R}^n$. The following lemma holds.

Lemma 17.3. *Suppose A is an $n \times n$ real symmetric matrix that is p.s.d. A vector $\tilde{x}$ minimizes q on $\mathbb{R}^n$ if and only if it obeys $Ax \ = \ 0$. The set of all such minimizers is a subspace of $\mathbb{R}^n$.*

Proof. Since A is p.s.d. then $\alpha(q) = 0$ and it is attained. From theorem 16.1, q is convex so theorem 15.7 says that the minimizers of q on $\mathbb{R}^n$ are precisely the solutions of $Ax = 0$. The set of all solutions of this equation is a closed subspace $N(A)$ of $\mathbb{R}^n$. $\square$

This theorem implies that A is positive definite if and only if it is p.s.d. and 0 is not an eigenvalue of A. The subspace $N(A)$ is called the null space of A. Suppose A is p.s.d and $dimN(A) = m \geq 1$. Then $N(A)$ has an orthogonal complement W in $\mathbb{R}^n$ which will be a subspace of dimension $n - m$. Each vector $x \in \mathbb{R}^n$ will have a unique decomposition

$$x \;:=\; y \;+\; z \qquad \text{where } y \in \text{ N(A) and } z \in W \tag{17.3}$$

Substitute this in the definition of $\mathcal{E}$, then

$$\mathcal{E}(x) \;=\; \mathcal{E}(z) \;-\; 2\,\langle\, b, y\,\rangle. \tag{17.4}$$

This formulation leads to both non-existence and existence results for (LEq). First a solvability (or non-existence) result for the linear equation.

Theorem 17.4. *Suppose A is a p.s.d. symmetric matrix with $dimN(A) = m \geq 1$. If there is a $y \in N(A)$ such that $\langle\, b, y\,\rangle \neq 0$, then there is no minimizer of $\mathcal{E}$ on $\mathbb{R}^n$ and no solution of (LEq).*

Proof. If there is such a $y \in N(A)$, then $\mathcal{E}(ty) = -t\langle\, b, y\,\rangle$ for all real t. Thus $\alpha(\mathcal{E}) = -\infty$ and there is no minimizer of $\mathcal{E}$ on $\mathbb{R}^n$. The energy principle lemma 17.1 then implies there is no solution of (LEq). □

The existence result is the following

Theorem 17.5. *(Solvability Condition)* *Suppose A is a p.s.d. symmetric matrix with $dimN(A) = m \geq 1$. If $\langle\, b, y\,\rangle = 0$ for all $y \in N(A)$ then*
(i) $\alpha(\mathcal{E})$ is finite and there are minimizers $\tilde{x}$ of $\mathcal{E}$ on $\mathbb{R}^n$, and
(ii) the set of all solutions of (LEq) is $\{\tilde{x} + y : y \in N(A)\}$, where $\tilde{x}$ is a minimizer of $\mathcal{E}$.

Proof. Under these assumptions, (17.4) shows that minimizing $\mathcal{E}$ on $\mathbb{R}^n$ is equivalent to minimizing $\mathcal{E}$ on W. Given a nonzero $z \in W$, $q(z)$ is non-zero, so q is positive definite on W. Since q is a quadratic form, then Corollary 16.2 shows that q, and thus $\mathcal{E}$, is coercive on W. Thus $\mathcal{E}$ is bounded below and attains its minimum value on W from theorem 17.2. Thus (i) holds while (ii) holds by linearity. □

The last two results constitute a special case (namely when A is p.s.d and symmetric) of the Fredholm alternative for solving (LEq). They say that, for this case, (LEq) has a solution if and only if

$$\langle\, b, u\,\rangle = 0, \qquad \text{for all} \quad u \in N(A). \tag{17.5}$$

The preceding results may be combined to prove the following theorem which says that whenever an energy function $\mathcal{E}$ is bounded below, then it has minimizers - a result that is quite special to this class of problems.

Theorem 17.6. *(Existence)* *Suppose A is an $n \times n$ symmetric matrix and $\mathcal{E}$ is defined by (17.1). If $\mathcal{E}$ is convex and bounded below on $\mathbb{R}^n$, then there is at least one minimizer of $\mathcal{E}$ on $\mathbb{R}^n$. If $\mathcal{E}$ is strictly convex this minimizer is unique.*

Proof. When $\mathcal{E}$ is convex then $q(x) = \langle Ax, x \rangle$ is convex as it only differs from $\mathcal{E}$ by a linear function. Theorem 16.1 implies that A is p.s.d. When A is positive definite this theorem follows from theorem 17.2. When A is only p.s.d., the assumption that $\mathcal{E}$ is bounded below implies $\langle b, y \rangle = 0$, for all $y \in N(A)$ from theorem 17.4. Then theorem 17.5 implies that there are minimizers of $\mathcal{E}$ on $\mathbb{R}^n$. $\square$

18. Least Squares Optimization and Linear Equations.

You might think that equations involving the class of symmetric p.s.d square matrices are quite a special class of equations. Here we shall show that any system of the form (LEq) with A an $m \times n$ matrix may be formulated as such problems. Suppose A is an $m \times n$ matrix, $b \in \mathbb{R}^m$ and we are interested in solving (LEq). When $m < n$, the system (LEq) is said to be an *underdetermined* system of linear equations. When $m > n$, it is an *overdetermined* linear system.

Define the *least squares* function $\mathcal{F} : \mathbb{R}^n \to [0, \infty)$ by

$$\mathcal{F}(x) \;:=\; \| Ax - b \|_2 \tag{18.1}$$

The problem of minimizing $\mathcal{F}$ on $\mathbb{R}^n$ is called the *least squares* optimization problem for (LEq).

Obviously a solution $\hat{x}$ of (LEq) is a minimizer of $\mathcal{F}$ on $\mathbb{R}^n$ and then $\mathcal{F}(\hat{x}) = 0$. In general the minimizers of $\mathcal{F}$ may not be solutions of (LEq) - but they can provide important information about the equation. Least squares methods are almost always used for studying overdetermined systems with $m \geq n$ and are very important in numerical and statistical applications. Much of the fundamental work on least squares methods was done by C.F. Gauss in connection with his work in geodesy and astronomy. Apparently it was the primary topic that he lectured on when students paid him to give university lectures at Goettingen.

From the definition,

$$\mathcal{F}(x) \;=\; \langle Ax - b, Ax - b \rangle \;=\; \langle A^T A x, x \rangle \;-\; 2 \langle Ax, b \rangle \;+\; \| b \|_2^2 \tag{18.2}$$

So
$$\mathcal{F}(x) \;=\; \mathcal{E}_2(x) \;+\; \| b \|_2^2$$

where $\mathcal{E}_2$ is an energy function with $A^T A, A^T b$ replacing A and b for the function $\mathcal{E}$ used in the previous section. Thus this $\mathcal{F}$ is an example of a convex energy function $\mathcal{E}$ that is bounded below.

Some basic results about this optimization problem may be summarized as follows

Theorem 18.1. *Assume A is an $m \times n$ matrix, $b \in \mathbb{R}^m$ and $\mathcal{F}$ is defined by (18.1), then there are minimizers of $\mathcal{F}$ on $\mathbb{R}^n$. A vector $\tilde{x}$ minimizes $\mathcal{F}$ on $\mathbb{R}^n$ if and only if it satisfies the equation*

$$A^T A\, x \;=\; A^T b. \tag{LSEq}$$

A minimizer of $\mathcal{F}$ on $\mathbb{R}^n$ will be a solution of (LEq) if $N(A^T) = \{0\}$.

Proof. From (18.2), $\mathcal{E}_2$ is a convex energy function associated with the p.s.d. matrix $A^T A$. Since $\mathcal{F}$ is positive, $\mathcal{E}$ is bounded below on $\mathbb{R}^n$ by $-\|b\|_2^2$. Thus theorem 17.6 implies there are minimizers of $\mathcal{F}$ on $\mathbb{R}^n$

The derivatives of $\mathcal{F}$ are given by

$$\nabla\mathcal{F}(x) = 2\,A^T(Ax-b) \qquad \text{and} \quad D^2\mathcal{F}(x) = 2A^T\,A. \tag{18.3}$$

Since $\mathcal{F}$ is convex, a vector $\tilde{x}$ minimizes $\mathcal{F}$ if and only if (LSeq) holds. This equation may be written as $A^T(Ax-b) = 0$, so the last sentence holds. □

A vector $x_{LS} \in \mathbb{R}^n$ is a said to be a *least squares solution* of (LEq) provided it satisfies (LSEq). The theorem says that such vectors are exactly the minimizers of $\mathcal{F}$ on $\mathbb{R}^n$. When the minimal value is positive, there is no solution of (LEq). Sometimes least squares solutions are called *generalized solutions* of (LEq).

Suppose $N(A) = \{0\}$. Then the symmetric matrix $A^T A$ is non-singular and positive definite, as $||Ax||_2 > 0$ for non-zero x, so theorem 17.2 applies to $\mathcal{E}_2$. For each $b \in \mathbb{R}^m$, there is a unique minimizer x_{LS} of $\mathcal{F}$ on $\mathbb{R}^n$ and then x_{LS} is the unique solution of (LSq).

If dim N(A) ≥ 1, and x_{LS} is a minimizer of $\mathcal{F}$ on $\mathbb{R}^n$, then $x_{LS}+y$ is again a minimizer for any $y \in N(A)$ as $\mathcal{F}(x_{LS}+y) = \mathcal{F}(x_{LS})$. Thus there is an *affine subspace* of minimizers of $\mathcal{F}$ on $\mathbb{R}^n$ - or the set of all minimizers is

$$\mathcal{M}(\mathcal{F}) \;=\; \{x_{LS}\} \;+\; N(A) \;=\; \{x_{LS} \;+\; y : y \in N(A).\}$$

When *dim* $N(A^T) \;\geq\; 1$, then a least squares solution of (LSq) need not be a solution of (LEq). For a general $m \times n$ matrix A, the fundamental theorem of linear algebra is that

$$\mathbb{R}^m \;=\; N(A^T) \oplus R(A) \qquad \text{and} \qquad \mathbb{R}^n \;=\; N(A) \oplus R(A^T).$$

See Strang [7], chapter 2, section 4. Moreover the ranks of A and A^T are the same, or $dim\,R(A) \;=\; dim\,R(A^T) \;=\; r$. If m > n, then

$$\text{dim N}(A^T) \;=\; m \;-\; r \;>\; \text{dim N}(A) \;=\; n \;-\; r.$$

Thus, for overdetermined systems, the last part of theorem 18.1 shows that the minimizers of $\mathcal{F}$ need not be solutions of (LEq).

18.1. **Preconditioned Least Squares Problems.** .

The previous optimization problem could be regarded as a special case of the following problem for the solutions of (LEq). Again this analysis is primarily used when $m \;\geq\; n$.

Let A be a $m \times n$ matrix, $b \in \mathbb{R}^n$ and M be a p.d. $m \times m$ symmetric matrix. Define the *preconditioned least squares* function $\mathcal{F}_M : \mathbb{R}^n \to [0,\infty)$ by

$$\mathcal{F}_M(x) \;:=\; \langle\, M(Ax-b), Ax-b\,\rangle. \tag{18.4}$$

The *preconditioned least squares optimization problem* is to minimize $\mathcal{F}$ on $\mathbb{R}^n$. When $M :=\; I$, this reduces to the least squares problem described above.

Note that $\mathcal{F}_M(x) \;\geq\; 0$ so $\alpha(\mathcal{F}_M) \geq 0$ and

$$\mathcal{F}_M(x) \;=\; \langle\, A^T M A x, x\,\rangle -\; 2\,\langle\, MAx, b\,\rangle \;+\; \langle\, Mb, b\,\rangle. \tag{18.5}$$

Thus $$\mathcal{F}_M(x) \;=\; \mathcal{E}(x) \;+\; \langle\, Mb, b\,\rangle.$$
Here $\mathcal{E}$ as in the energy method but with $A^T MA, A^T Mb$ replacing A, b.

If $m = n$ and A is p.d symmetric, then so is A^{-1}, and choosing $M = A^{-1}$ yields the energy function of the previous section. Thus functions $\mathcal{F}$ of this form include both energy and least squares functions as special cases. M is called a *preconditioner* for (LEq).

This function has derivatives

$$\nabla\mathcal{F}_M(x) \;=\; 2\,A^T M(Ax - b) \qquad \text{and} \quad D^2\mathcal{F}_M(x) = 2\,A^T MA. \tag{18.6}$$

so $\mathcal{F}_M$ is convex on $\mathbb{R}^n$ from the analysis of quadratic forms in section 16.

The introduction of the matrix M does not change the existence and uniqueness results but may significantly change the convergence properties of algorithms for finding the minimizers. The proof of Theorem 18.1 is easily modified to prove the following result about this optimization problem.

Theorem 18.2. *Assume A is an $m \times n$ matrix, M is a p.d. $m \times m$ symmetric matrix, $b \in \mathbb{R}^m$ and $\mathcal{F}_M$ is defined by (18.4). There are minimizers of $\mathcal{F}_M$ on $\mathbb{R}^n$. A vector $\tilde{x}$ minimizes $\mathcal{F}_M$ on $\mathbb{R}^n$ if and only if it satisfies the equation*

$$A^T\, M\, A\, x \;=\; A^T M\, b. \tag{18.7}$$

If $N(A^T) = \{0\}$, then a minimizer of $\mathcal{F}_M$ on $\mathbb{R}^n$ is a solution of (LEq).

Exercises. .

Exercise 18.1 Let A be the 3×3 matrix

$$A \;:=\; \begin{pmatrix} 1 & 0 & 1 \\ 0 & a_1 & 0 \\ 1 & 0 & a_2 \end{pmatrix} \tag{18.8}$$

Evaluate the associated quadratic form on $\mathbb{R}^3$ and find conditions on a_1, a_2 for A to be positive semi-definite. When is it positive definite? When A is p.s.d under what conditions on $b \in \mathbb{R}^3$ are there solutions of $Ax = b$? Find all solutions of the equation in this case.

Exercise 18.2 Let $S \;:=\; \{a^{(j)} : 1 \leq j \leq m\}$ be a finite set of points in $\mathbb{R}^3$ with $m \;\geq\; 4$. Consider the problem of finding a plane in space that provides the "best approximation" to this data. Suppose this plane Σ has the equation

$$\langle\, c, x\,\rangle \;=\; \sum_{k=1}^{3} c_k\, x_k \;=\; \gamma.$$

Find the expression for the Euclidean distance d_j of the point $a^{(j)}$ from a point in this plane. This distance is a function of c, γ. Define the function $F : \mathbb{R}^4 \to \mathbb{R}$ by

$$F(c, \gamma) \;:=\; \sum_{j=1}^{m} d_j^2 \;+\; \|c\|_2^2 \;-\; 1.$$

This is a quadratic function of c, γ. The above analysis shows that there are minimizers of this function. Find the linear equations satisfied by the minimizers. How many equations are there? When does your matrix have a non-trivial null space?

References

[1] Berkowitz, L.D. "Convexity and Optimization in $\mathbb{R}^n$", Wiley-Interscience (2002).

[2] Boyd, Stephen and L Vandenberghe, "Convex Optimization", Cambridge U Press. (This book is available for free download at http://www.stanford.edu/~boyd/cvxbook/).

[3] Fleming, Wendell. "Functions of Several Variables, 2nd edition", Springer Undergraduate Texts in Mathematics, (1977).

[4] Lay, Steven R. "Convex Sets and their Applications", Dover (2007).

[5] Ortega, J.M. and W.C. Rheinboldt. "Iterative Solution of Nonlinear Equations in Several Variables", Academic Press (1970) (recently reprinted by SIAM). (2001).

[6] Salas, S, E. Hille and G.J. Etgen, "Calculus, One and Several Variables", (8th ed), J. Wiley (1999).

[7] Strang, G. "Linear Algebra and its Applications", Academic Press (1980)