Lecture Notes M6366; Optimization

UH Fall 2004.

Table of Contents.

Lecture 1; 8/22/2004
Please review material you have had in the past on metric spaces and the analysis of $\mathbb{R}^n$. Most of the following material is in Berkovitz chapter 1, sections 1-7.

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

$\mathbb{R}^n$ is the vector space of all n-tuples of real numbers. It 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;

(i): $p(x) \geq 0$ and $p(x) = 0$ implies $x = 0$,

(ii): (homogeniety) $p(cx) := |c|p(x)$,

(iii): (triangle inequality) $p(x + y) \leq p(x) + p(y)$.

Often norms are denoted $\|x\|$ instead of $p(x)$. Three examples of norms on $\mathbb{R}^n$ are the following;

Ex 1.1 The Euclidean norm is $\|x\|_2 := (\sum_{j=1}^n |x_j|^2)^{1/2}$.

Ex 1.2. The 1-norm on $\mathbb{R}^n$ is $\|x\|_1 := \sum_{j=1}^n |x_j|$.

Ex 1.3. The $\infty-$ (or max-) norm is $\|x\|_\infty := \max_{1 \leq i \leq n} |x_j|$.

For any $p \in [1, \infty)$, the expression $\|x\|_p := \{\sum_{j=1}^n |x_j|^p\}^{1/p}$ is a norm on $\mathbb{R}^n$. For all such p, it is easy to show that (i) & (ii) of the properties of a norm hold. The fact that $\|x + y\|_p \leq \|x\|_p + \|y\|_p$ holds is Minkowski's inequality.

The expression $\|x\|_p$ is defined for all $p > 0$ but when $0 < p < 1$, it is not a norm as it does not satisfy the triangle inequality (iii) and the unit ball with respect to $\|.\|_p$ is not a convex subset of $\mathbb{R}^n$.

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

$$\begin{aligned} B_r^p(a) &:= \{x \in \mathbb{R}^n : \|x - a\|_p < r\} \\ S_r^p(a) &:= \{x \in \mathbb{R}^n : \|x - a\|_p = r\}, \text{ and} \\ \overline{B}_n^p(a) &:= \{x \in \mathbb{R}^n : \|x - a\|_p \leq r\} = B_r^p(a) \cup S_n^n(a) \end{aligned}$$

When $p = 2$, we usually omit the superscript and often write $|x|$ in place of $\|x\|_2$. In this case the 2-norm is the norm induced by the inner product

$$\langle x, y \rangle \ := \ \sum_{j=1}^{n} \quad x_j y_j. \tag{1.1}$$

If I is a nontrivial subinterval of $\mathbb{R}$, we say that a function f is *decreasing* (or *increasing*) on I provided that $x_1 < x_2$ in I, implies $f(x_1) \geq (\leq) f(x_2)$. The function is *strictly decreasing (strictly increasing)* on I provided that $x_1 < x_2$ in I, implies $f(x_1) > (<) f(x_2)$.

Theorem 1.1 For any non-zero $x \in \mathbb{R}^n, p > 0$, $\|x\|_p$ is a decreasing function of p. It is strictly decreasing if x has more than 1 non-zero component.
Proof: Fix $x(\neq 0) \in \mathbb{R}^n$ and consider

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

Let $m = \|x|_\infty$ and $y \ := \ x/m$. Express $f(p)$ in terms of y and differentiate with respect to p, then a computation shows that $f'(p) \leq 0$ for all $p > 0$.

Corollary 1.2 For $0 < p_1 < p_2 \leq \infty, \quad B_r^{p_1}(a) \subset B_r^{p_2}(a)$.

A sequence $\Gamma := \{x^{(k)} : k \geq 1\} \subseteq \mathbb{R}^n$ converges to a point $\hat{x}$ in the p-norm provided

$$\lim_{k \to \infty} \|x^{(k)} - \hat{x}\|_p \ = \ 0.$$

That is for any $\epsilon > 0$ there is a $K(\epsilon)$ such that $k > K(\varepsilon)$ implies $\|x^{(k)} - \hat{x}\| < \epsilon$.
This convergence does not actually depend on p, as a consequence of the following results.

Theorem 1.3 If a sequence $\Gamma \subseteq \mathbb{R}^n$ converges to a point $\hat{x}$ with respect to $\|.\|_p$ for some $p \in [1, \infty)$, then it converges to $\hat{x}$ in every $\|.\|_p$, $1 \leq p \leq \infty$.

This theorem follows from the following general inequality.

Theorem 1.4 Suppose $1 \leq p < q \leq \infty$ and $x \in \mathbb{R}^n \backslash \{0\}$, then

$$0 < \ \|x\|_q \ \leq \ \|x\|_p \ \leq \ n^{1/p - 1/q} \ \|x\|_q \tag{1.2}$$

We'll prove this later. To prove convergence of a particular sequence, we'll usually choose the most convenient value of p.

In general if Γ is any subset of, (not necessarily a sequence in), $\mathbb{R}^n$, a point $a \in \mathbb{R}^n$ is said to be a limit point of Γ provided that for any $\epsilon > 0$, there are points in $\Gamma \cap (B_\epsilon(a) \backslash \{a\})$.

A subset U of $\mathbb{R}^n$ is said to be open if for each $x \in U$, there is an $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U$. 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 subset U of $\mathbb{R}^n$ is compact if it is closed and bounded. Suppose $K \subseteq \mathbb{R}^n$ is a compact set, then a basic result in metric topology says that every infinite subset of K has a limit point in K.

Lecture 2; 8/24/2004.

2. Metric spaces, infs and sups

This semester we will generally work with metric spaces. Here I'll collect the basic definitions in forms that will be used. If you've done more analysis, you may have seen more general definitions than the one's given here - and these definitions will be theorems in that case.

Let (X, d) be a metric space. 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

$$\text{For any } \epsilon > 0, \text{ there is an N such that } n \geq N \Rightarrow d(x_k, x) < \epsilon. \tag{2.1}$$

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

The metric space (X,d) is *compact* if every sequence in X has a subsequence which converges to an element of X.

Obviously a compact metric space is complete, but the converse doesn't hold.

The extended real line $\overline{\mathbb{R}}$.

The extended real line is the set

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\} = [-\infty, \infty]. \tag{2.2}$$

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

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

$$d_e(x,y) \;:=\; \left| \int_x^y \frac{dt}{1+t^2} \right| \;=\; |\arctan(y) - \arctan(x)|. \tag{2.3}$$

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

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

Theorem 2.1 $\overline{\mathbb{R}}$ with the metric (2.3) is a compact metric space. Let $\{x_n : n \geq 1\}$ be 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. The proof that the sequence converges wrt the metric d_e follows as

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

for all t.

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 $\mathbb{R}$.

Ex 2.1. 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_ϵ sufficiently large that $\pi/2 - \arctan(N_\epsilon) < \epsilon$. 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 $\mathbb{R}$.

Corollary 2.2. A subset K of $\overline{\mathbb{R}} \;:=\; [-\infty, \infty]$ is compact if and only if K is a closed subset of $\overline{\mathbb{R}}$.
Proof: This is a special case of the general theorem that subset of a compact metric space is compact if and only if it is closed.

Ex 2.2: $[1, \infty]$ is compact in $\overline{\mathbb{R}}$. It is not compact in $\mathbb{R}$ in view of the example in Ex 2.1.

The usual definition of scalar multiplication may be extended to products of $\pm\infty$ and a real number c by

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

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 convention when $c \in \mathbb{R}$.

(i.) $\quad c+\infty \;:=\; \infty, \qquad \infty+\infty \;:=\; \infty.$
(ii.) $\quad c+(-\infty) \;:=\; -\infty, \qquad -\infty+(-\infty) \;:=\; -\infty.$
(iii.) The expression $\infty+(-\infty)$ is not defined.

In view of (iii), $\overline{\mathbb{R}}$ is not a vector space.

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 $A \setminus B := \{x : \; x \in A, \& x \notin B\}$.

Infima, Suprema, Minima & Maxima.

Let E be a non-empty subset of $\overline{\mathbb{R}}$. $\gamma \in \overline{\mathbb{R}}$ is a *lower bound* for E provided $\gamma \le x$ for all $x \in E$. Similiarly, γ is an *upper bound* for E provided $\gamma \ge x$ 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 $\emptyset$. (Why?)

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 $\gamma \le \alpha$. 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 $\beta \le \; \gamma$.

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

Applying these definitions to the empty set, you observe (?) that $\alpha(\emptyset) = +\infty$ and $\beta(\emptyset) = -\infty$.

Ex 2.3: Suppose $E := (0,1)$. Then any $\gamma \le 0$ will be a lower bound for E and 0 is the infimum of E. Similarly, any $\delta \ge 1$ will be a upper bound for E and 1 is the supremum of E.

Theorem 2.3 If E is a non-empty subset of $\overline{\mathbb{R}}$, then E has a unique infimum $\alpha(E)$ and a unique supremum $\beta(E)$ in $\overline{\mathbb{R}}$. Moreover $\alpha(E) \leq \beta(E)$.

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 to ensure 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 $\alpha(E) \leq x \leq \beta(E)$. Note that when $E = \emptyset$, the last inequality in this theorem does not hold.

When E is non-empty we say that α 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.

Ex 2.4: (0,1) has no maximum or minimum. [0,1] has both a maximum and a minimum.

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

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

or $\alpha(-A) \ = \ -\beta(A)$ for any non-empty set $A \in \overline{\mathbb{R}}$.

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

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

3. Minimization of Continuous Functions.

Let K be a non-empty set in $\mathbb{R}^n$ and $f : K \to \overline{\mathbb{R}}$ be a continuous function. The *infimum* of f on K is

$$\alpha(f; K) \ := \ inf_{x\in K} \quad f(x) \tag{3.1}$$

and this number always exists in $\overline{\mathbb{R}}$. It is called the *value* of the problem. When there is a point $\hat{x} \in K$ such that $f(\hat{x}) = inf_{x \in K}\{f(x)\}$, then $\hat{x}$ is called a *minimizer* of f on K. When $K = \mathbb{R}^n$, K is often omitted and we write $\alpha(f)$ for $\alpha(f; \mathbb{R}^n)$.

Ex 3.1. 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 $\mathbb{R}$.

Ex 3.2. Let $K := [-\pi/2, \pi/2]$. Define $f : K \to [-\infty, \infty]$ by

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

Then 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.

Ex 3.3. Define $f : \mathbb{R} \to \overline{\mathbb{R}}$ by

$$f(x) = \begin{cases} \ln\ |x| & \text{if} \quad |x| \neq 0 \\ -\infty & \text{if} \quad x = 0 \end{cases}$$

This f 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.

Lecture 3; 8/29/2004

General Optimization Problems.

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

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

$$\alpha(f; X) := inf_{x \in X} f(x) \quad \text{and} \tag{3.2}$$
$$\mathcal{M}(f) := \{x \in X : f(x) = \alpha(f; X)\}. \tag{3.3}$$

The set $\mathcal{M}(f)$ is called the set of *minimizers of f on X*. $\alpha(f; X)$ is the *value* of this problem.

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

$$\beta(f; X) := sup_{x \in X} f(x) \quad \text{and} \tag{3.4}$$
$$\mathcal{M}(f) := \{x \in X : f(x) = \beta(f; X)\}. \tag{3.5}$$

One has $\beta(f;X) = -\alpha(-f;X)$ 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. We shall concentrate on minimization problems.

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

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

Existence of Minimizers of Optimization Problems.

This semester we will primarily consider problems where X is a closed subset of $\mathbb{R}^n$. To prove the existence of minimizers we usually use the following basic theorem from elementary real analysis.

Theorem 3.1 (Bolzano-Weierstrass) A subset K of $(\mathbb{R}^n, \|.\|_p)$ is compact if and only if K is closed and bounded.

The next result is called the *fundamental existence theorem* in Berkovitz, Chapter 1, section 9.

Theorem 3.2 (Weierstrass) Suppose K is a non-empty compact subset of a complete 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 as follows. Choose $x^{(1)}$ to be any point in K. If $x^{(1)}$ is a minimizer of f on K, put $x^{(k)} = x^{(1)}$ for all $k \geq 1$. If $x^{(1)}$ is not a minimizer of f on K, then choose a point $x^{(2)} \in K$ with $f(x^{(2)}) < f(x^{(1)})$.

Continue, either we find a minimizer in a finite number of steps, or we have a strict descent sequence. In the first case, $\alpha(f;K) = f(x^{(l)})$ for some $l \geq 1$ will be finite. In the second case choose the sequence so that $f(x^{(k)}) \to \alpha(f;K)$. 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, $f(\hat{x}) = \alpha(f;K)$ and this is finite. Hence (i) and (ii) hold for minimization.

The results for maximization hold by using the same arguments with $-f$ in place of f.

Comment: This is an existence theorem which is not "constructive". The argument does not provide an explicit way (ie a specific algorithm or construction) that produces $x^{(k+1)}$ when f and $x^{(k)}$ are given or known.

Corollary 3.3 Suppose K is a non-empty compact subset of a metric space (X, d) and $f : K \to \overline{\mathbb{R}}$ is continuous, then 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: $\alpha(f; K) := inf_{x \in K} f(x)$ and $\beta(f : K) := sup_{x \in K} f(x)$ exist as numbers in $\overline{\mathbb{R}}$. The construction of the theorem generates a minimizing sequence for f on K. Since K is compact, this sequence has a convergent subsequence and this subsequence has a limit $\hat{x}$ in X which is also in K. By continuity $\hat{x}$ is a minimizer of f on K. (Please make sure you understand why each of these statements holds.)

Ex 3.4 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 $\alpha(E; \overline{\mathbb{R}}) := inf_{x \in \overline{\mathbb{R}}} \quad e^x = 0$, 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 $\mathbb{R}$.

Let $f : X \to \overline{\mathbb{R}}$ be a given function and $c \in \mathbb{R}$. The *level sets (or contour sets)* of f are defined to be

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

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

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

The function f is said to be *proper* provided $f(x) \neq -\infty$ for any $x \in X$. That is the range of f $= f(X) \subseteq (-\infty, \infty]$.

Ex 3.5. Define $L : \mathbb{R} \to \overline{\mathbb{R}}$ by

$$L(x) = \begin{cases} \ln |x| & \text{if} \quad |x| \neq 0 \\ -\infty & \text{if} \quad x = 0 \end{cases}$$

This L is even and continuous as a map of $\mathbb{R}$ into the metric space $(\overline{\mathbb{R}}, d_e)$. This function is not proper - since $L(0) := -\infty$. The value $\alpha(L) = -\infty$ and 0 is the (unique) minimizer of L on $\mathbb{R}$. The supremum of L on $\mathbb{R}$ is ∞, but there is no maximizer.

Corollary 3.4. Suppose $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is continuous and the set $S_c(f)$ is non-empty and bounded for some real c, then there is an $\hat{x} \in \mathbb{R}^n$ such that $\hat{x}$ is a minimizer of f on $\mathbb{R}^n$. Furthermore, if f is proper then $\alpha(f)$ is finite.

Proof. Since f is continuous, $S_c(f)$ 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 3.3 with $K = S_c(f)$ then f has attains its value on $S_c(f)$ or there is an $\hat{x} \in S_c(f)$ such that

$$f(\hat{x}) = inf_{x \in S_c(f)} \ f(x) = \alpha(f). \tag{3.6}$$

When f is proper then $f(x) \neq -\infty$ for any x, so this implies that $\alpha(f)$ is finite as it is attained.

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

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

It is *coercive* if

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

Homework Pb 1.6 Prove that if $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and weakly coercive then $\alpha(f)$ is finite and $\mathcal{M}(f)$ is non-empty.

Ex 3.6. 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 $\mathbb{R}^n$.

Ex 3.7. 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 $\hat{x} = -b/2a$.

Ex 3.8. 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.

Ex 3.9. 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 of p on $\mathbb{R}$.

Lecture 4; 9/1/2004

4. 1-d Local Minimization

We shall first treat one-dimensional optimization in detail and will generalize these techniques to n-dimensional problems later.

Suppose $f : \mathbb{R} \to \overline{\mathbb{R}}$ is a continuous function. Just as before, define

$$\alpha(f) \ := \ inf_{x\in\mathbb{R}} f(x) \quad \text{and} \quad \mathcal{M}(f) \ := \ \{x \in \mathbb{R} : f(x) = \alpha(f)\}.$$

Sometimes the elements of $\mathcal{M}(f)$ are called the *global minimizers* of f; remember that $\mathcal{M}(f)$ may be empty.

Definitions. A point $\tilde{x}$ in $\mathbb{R}$ is

(i) a *local minimizer* of f if there is a $\delta > 0$, such that

$$f(x) \geq f(\hat{x}) \text{ for } |x - \hat{x}| < \delta,$$

(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 $\tilde{x}$ in $|\tilde{x} - \hat{x}| < \delta_1$.

Comment. A local minimizer could be a local and/or a global maximizer of f. However, a strict (or isolated) local minimizer cannot be a local maximizer. Below, in result 4.1, it will be shown that $(iii) \Rightarrow (ii) \Rightarrow (i)$ above.

Ex 4.1: Consider $f(x) = \begin{cases} x^2(2 + cos(x^{-1})) & x \neq 0 \\ 0 & x = 0 \end{cases}$.

f is an even function and for all x, $\quad x^2 \leq f(x) \leq 3x^2$.

This function has infinitely many positive local minimizers $\{x_k : k \geq 1\} \subset (0, 2]\}$ which have a limit point at 0. Each of them is an isolated local minimizer. 0 is the unique global minimizer and it is a strict minimizer of f on $\mathbb{R}$, but 0 is not a isolated local minimizer of f.

Result 4.1: 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.

Limits, liminfs and limsups

Calculus is based on being able to take limits - and work with the resulting expressions - such as integrals and derivatives. For optimization theory, mathematicians have recently (since 1964) developed some better versions of differential calculus than what you learnt as an undergraduate - theories that were based on the work of Cauchy in the early 19th century.

An infinite sequence $\Gamma := \{x^{(k)} : k \geq 1\} \subset \mathbb{R}$ 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

$$\lim_{k\to\infty} x^{(k)} \;=\; \sup_{k\geq 1} x^{(k)} \quad (\text{or} = \inf_{k\geq 1} x^{(k)}.) \tag{4.1}$$

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.

$$\liminf_{k\to\infty} x^{(k)} \;:=\; \sup_{l\geq 1} \inf_{k\geq l} x^{(k)} \quad \text{and} \quad \limsup_{k\to\infty} x^{(k)} \;:=\; \inf_{l\geq 1} \sup_{k\geq l} x^{(k)} \tag{4.2}$$

These always exist and are finite when the sequence is bounded. More generally the definition (4.2) works for unbounded sequences, or for sequences of points in $\overline{\mathbb{R}}$. In this case the lim inf or lim sup could be $\pm\infty$. In particular every sequence Γ of points in $\overline{\mathbb{R}}$ has a liminf and a limsup and

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

The sequence has a limit $\gamma \in \overline{\mathbb{R}}$ if and only if these two numbers are equal to γ.

Ex 4.2. Define a sequence by $x_{2m+1} = 1 - (1/m)$ and $x_{2m} := -1 + (1/m)$. Then $-1 < x_m < 1$ for all m,

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

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

Comment: This definition of liminf and limsup does NOT involve a metric. There is no ϵ here, just infs and sups. More generally one can define the liminf and limsup of a countable family of sets; see the attached problem from page 3 of Frank Jones book on Lebesgue integration.

Suppose $I := (a, b)$ is an interval, $f : I \to \mathbb{R}$ is a function and $c \in [a, b]$. Define

$$f_k(c) := inf \ \{f(x) : x \in I \text{ and } |x - c| < 1/k\}$$

The sequence $\{f_k(c) : k \geq 1\}$ is an increasing function of k, so we define

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

Since the supremum of an increasing sequence always exists, this liminf exists as either a number or $\pm\infty$. When $c \in I$, then $f_k(c) \leq f(c)$ for all k, so that $\liminf_{x \to c} f(x) \leq f(c)$. In a similar manner you can define $\limsup_{x \to c} f(x)$. (Please try).

Ex 4.2: The function $f(x) \ := \ \sin(x^{-1})$ does not have a limit as $x \to 0$. However the set of all limit points of $f(x)$ as $x \to 0$ is [-1,1], so that $\liminf_{x \to 0} \ \sin(x^{-1}) \ = \ -1$ and $\limsup_{x \to 0} \ \sin(x^{-1}) \ = \ 1$.

Necessary Conditions for a Local Minimizer.

In the following theorem, we use the concept of left and right derivatives of f. These are defined as follows by using 1-sided difference quotients. Suppose $f : (a - \delta, a + \delta) \to \mathbb{R}$ is a function then

$$D_+ f(a) \quad := \quad \lim_{t \to 0^+} \ t^{-1} \ [f(a + t) - f(a)] \tag{4.3}$$

$$D_- f(a) \quad := \quad \lim_{t \to 0^+} [f(a) - f(a - t)]/|t| \tag{4.4}$$

These need not exist, but if they do and they are equal and finite, then f is said to be differentiable at a; The derivative of f at a will be denoted by either $f'(a)$ or $Df(a)$.

Theorem 4.2. Suppose $f : (a, b) \to \mathbb{R}$ is continuous and $\hat{x}$ is a local minimizer of f on (a, b), then

(i) $\liminf_{|h| \to 0} |h|^{-1} \ [f(\hat{x} + h) - f(\hat{x})] \geq 0$ (4.5)

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

$$D_- f(\hat{x}) \leq 0 \ \text{ and } \ D_+ f(\hat{x}) \geq 0. \tag{4.6}$$

(iii) (Fermat's rule) if f is differentiable at $\hat{x}$, then $f'(\hat{x}) = 0$. (F)

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

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

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 $D_+ f(\hat{x}) \geq 0$ - whenever this limit exists.

When h is negative then the inequality sign here is reversed so $D_- f(\hat{x}) \leq 0$. whenever this limit exists.
(iii) If f is differentiable at $\hat{x}$, then $D_- f(\hat{x}) = D_+ f(\hat{x})$, so part (ii) of this result implies that this common value must be 0 or (F) holds.

Ex 4.1 (continued). The function f described in example 4.1 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).
Ex 4.3: The function $f(x) = |x|$ has a unique local 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).

These are said to be sufficient conditions for a local minimizer because a point $\hat{x}$ in (a, b) could obey (F) or (ii) without being a local minimizer of f on (a,b). A simple example is $f(x) = x^3$ which has $f'(0) = 0$ but 0 is not a local minimizer.

Note that if $\hat{x}$ is a local maximizer of f on (a, b), then conditions (i) and (ii) in this last theorem change (to what?) - but condition (iii) remains true.

A point $\tilde{x} \in (a, b)$ is said to be a *critical point* of f if f is differentiable at $\tilde{x}$ and $f'(\tilde{x}) = 0$.

Second Derivative Conditions

If $\hat{x}$ is a local minimizer of f then the following second derivative-type conditions must hold at $\hat{x}$. It is called a *necessary* condition for $\hat{x}$ to be a local minimizer.

Theorem 4.3 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

$$\liminf_{h \to 0} \; h^{-1} f'(\hat{x}+h) \;=\; c \;\geq\; 0 (4.7) \tag{4.5}$$

(ii) If $f''(\hat{x})$ exists, then $f''(\hat{x}) \geq 0$. (4.8)

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 (4.3), then there is a h_1 such that

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

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 $f''(\hat{x}) = \lim_{h\to 0}\ h^{-1}\ f'(\hat{x}+h)\ =\ c$. From part (i), we must have $c \geq 0$.

Again what is the analog of this result if $\hat{x}$ is an isolated local maximizer?

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. It is a *sufficient* condition to be a local minimizer of f on (a,b).

Theorem 4.4 Suppose f is C^1 on (a,b), $\hat{x} \in (a,b)$ is a critical point of f and (i) of theorem 4.3 holds with $c > 0$. Then $\hat{x}$ is an isolated strict local minimizer of f on (a,b).
Proof: When (4.7) 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 $0 \leq |h| <\ \delta$. Hence $\hat{x}$ is an isolated strict local minimizer of f.

Comment: Just as in theorem 4.3, the conditions of this theorem hold provided $f''(\hat{x}) > 0$.

Lecture 5; 9/8/2004

5. Existence of Local Minimizers.

Under a variety of conditions, the existence of a local minimizer of a function on an interval can be guaranteed. Two useful criteria are the following; the first just involves the function f, the second involves the derivative f'. It usually helps to sketch some graphs to visualize the following results.

Theorem 5.1 Suppose $f : [a, b] \to \mathbb{R}$ is continuous and there exist

$$a \leq x_1 < x_3 < x_2 \leq b \quad \text{such that} \quad f(x_3) < \min\,(f(x_1), f(x_2)). \tag{5.1}$$

Then there is a local minimizer of f on $[a, b]$.
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 (5.1) 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 the following we say that a function f is *differentiable* on the closed interval [a,b], provided f is differentiable at each point in the open interval (a, b) and also $D_+f(a)$ and $D_-f(b)$ exist and are finite. In this case we usually write $f'(a), f'(b)$ for these derivatives. f is continuously differentiable (C^1) on [a,b] if the resulting function f' is continuous on [a,b].

Theorem 5.2 Suppose $f : [a, b] \to \mathbb{R}$ is C^1 and there are $x_1, x_2 \in [a, b]$ such that

$$a \leq x_1 < x_2 \leq b \quad \text{and} \quad f'(x_1) < 0 < f'(x_2). \tag{5.2}$$

Then there is at least one local minimizer $\hat{x}$ of f in (x_1, x_2) and $f'(\hat{x}) = 0$.
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 5.1, there is a local minimizer $\hat{x}$ of f in (x_1, x_2). Since f is C^1 on (a, b), the local minimizer will be a critical point of f.

You may find it interesting to work out what the corresponding criteria are for local maxima. What do you need to change in these results to have a corresponding result for local maximizers?

Error Estimates and Uniqueness for a Local Minimizer.

The following theorem is the usual result used to prove the uniqueness of a local minimizer.

Theorem 5.3 Suppose $f : [a, b] \to \mathbb{R}$ is C^1 and there is a $c_0 > 0$, such that

(i) $f'(x_2) - f'(x_1) \geq c_0(x_2 - x_1)$ for all $a \leq x_1 < x_2 \leq b$, and

(ii) $f'(a)f'(b) < 0$.

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

$$|x-\hat{x}| < {c_0}^{-1}\ |f'(x)| \qquad \text{for any } x \in (a,b). \tag{5.3}$$

Proof: From (i), $f'(b) \geq f'(a) + \ c_0(b-a)$, so (ii) implies that $f'(a) < 0$ and $f'(b) > 0$.
Now theorem 5.2 implies there is a local minimzer $\hat{x}$ of f in (a,b) and $f'(\hat{x}) = 0$.
When $a \leq x_1 < \hat{x}$, then (i) implies that $f'(\hat{x}) - f'(x_1) \ \geq \ c_0\ (\hat{x} - x_1) > \ 0$. Thus $f'(x_1) < 0$
Similarly, $\hat{x} < x_2 \leq b$, implies that $f'(x_2) > 0$. Thus $\hat{x}$ is the unique critical point of f in (a,b).
When $x \in [a, \hat{x})$, then from (i)

$$f'(\hat{x}) - f'(x) \geq \ c_0\ (\hat{x} - x) \quad \text{ or } \quad |\hat{x} - x| \leq {c_0}^{-1}\ |f'(x)|.$$

Similarly, when $x \in (\hat{x}, b]$, so (i) implies (5.3)

Comments 1. 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.
2. (5.3) is the basic *error estimate* in 1-d optimization. It is used as a stopping condition in most algorithms for minimization. One seeks points where $|f'(x)|$ is sufficiently small and then (5.3) provides an upper bound on the distance to the local minimizer in [a,b]. - provided we know c_0.

6. Convex Functions on an Interval

Let I be a closed nonempty interval in $\mathbb{R}$. That is $I = (-\infty, \infty)$, or $[a, \infty)$ or $(-\infty, b]$ or $[a,b]$ with $-\infty < a \leq b < \infty$. Usually we assume the interior of $I \neq \emptyset$ (that is I is not a single point or $b \neq a$).

A function $f : I \to (0, \infty)$ is *convex* provided

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

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

We first need some criteria for a function to be convex on a closed interval.

Theorem 6.1 Suppose $f : I \to \mathbb{R}$ is such that for each $x \in I$ there is a $\eta \in \mathbb{R}$ such that

$$f(y) \ \geq \ f(x) \ + \ \eta\ (y - x) \qquad \text{for all} \quad y \in I \tag{5.1}$$

then f is convex on I.
Proof. Choose $x < y \in I$ and put $z := (1-t)x + ty$ with $0 < t < 1$. Then (5.1) 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 first equation by $(1 - t)$, second by t and add to find

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

so f is convex on I.

Corollary 6.2 Suppose $f : I \to \mathbb{R}$ is such that for each $x \in I$ there is a $\eta \in \mathbb{R}$ such that

$$f(y) \ > \ f(x) \ + \ \eta \ (y - x) \qquad \text{for all} \quad y \in I, y \neq x, \tag{5.2}$$

then f is strictly convex on I.
Proof. Just as for the theorem - but with strict inequalities.

Corollary 6.3 Suppose $f : I \to \mathbb{R}$ is differentiable at each point $x \in I$. Then f is convex on I if and only if

$$f(y) \ \geq \ f(x) \ + \ f'(x) \ (y - x) \qquad \text{for all} \quad x, y \in I. \tag{5.3}$$

It is strictly convex on I iff (5.3) holds with $>$ in place of $\geq$ and $y \neq x$.
Proof. When (5.3) holds then f is convex on I from the theorem. Conversely if f is convex on I, $0 < t \leq 1$ then

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

for all $x, y \in I$. Let $t \to 0^+$ in this inequality then (5.3) 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.

Example 6.1. $f_p(x) := |x|^p$ is convex on $\mathbb{R}$ when $p \geq 1$. $-f_p(x)$ is convex on $(0, \infty)$ when $0 < p < 1$. (Sketch the graphs of these functions; they are even so it is enough to graph them on $[0, \infty)$.)

Example 6.2. Define $f_\infty(x) := \lim_{p \to \infty} f_p(x) \ := \begin{cases} 0 & |x| < 1 \\ 1 & |x| = 1 \\ \infty & |x| > 1 \end{cases}$
This is a convex function on $\mathbb{R}$ (with values in $[0, \infty]$).

Example 6.3 $f(x) := e^{\alpha x}$ is convex on $\mathbb{R}$ for any choice of $\alpha \in \mathbb{R}$.

Example 6.4. $f(x) := \begin{cases} x \ln x & x > 0 \\ 0 & x = 0 \end{cases}$ is convex and continuous on $[0, \infty)$.

Theorem 6.4 Suppose $f : I \to \mathbb{R}$ is convex and $x_1 < x_2 < x_3$ are in I, then

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

or $S_{12} \leq S_{13} \leq S_{23}$ 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, $x_3 - x_2 = (1 - t)(x_3 - x_1)$ 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 $S_{32} \geq S_{31}$.

Let $x_2 \to x_1^+$ or $x_2 \to x_3^-$, then the right derivative of f at x_1 and the left derivative of f at x_3 are defined by

$$D_+f(x_1) := \lim_{x_2 \to x_1^+} \frac{f(x_2) - f(x_1)}{x_2 - x_1} \quad \text{and} \quad D_-f(x_3) := \lim_{x_2 \to x_3^-} \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Taking careful limits in (CS) leads to the following result.

Theorem 6.5 Suppose I is an interval in $\mathbb{R}$ and $[x_1, x_1 + h) \subset I$ for some $h > 0$. If f is convex on I, then $D_+f(x_1)$ exists, it may be $-\infty$, it cannot be $+\infty$. If $(x_3 - h, x_3] \subseteq I$ for some $h > 0$, then $D_-f(x_3)$ exists, it cannot be $-\infty$, it may be $+\infty$. If $x \in (a, b)$, 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).

Comments: The theorem says that, when f is a convex function on $(a,b) \subset I$,
(i) f has left and right derivatives, $D_-f(x), D_+f(x)$ at each point x in (a,b),
(ii) these left and right derivatives are increasing on (a,b).
(iii) if f is differentiable on (a,b), then f' is an increasing function on (a,b).
(iv) Check how these results apply to the convex functions in examples 6.1 -6.4.

This implies that f is (Lipschitz) continuous on any closed subinterval of (a,b). When f is C^1 or C^2 there are some simpler criteria for convexity in terms of the derivatives of f.

Corollary 6.6 Suppose $f : I \to \mathbb{R}$ is differentiable at each point $x \in I$. 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. The above theorem shows that if 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 from the mean value theorem

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

Similarly if $y < x$. Hence corollary 6.3 implies that f is convex.

When $f'(x)$ is increasing and $f''(x)$ exists it must be ≥ 0.

The following result is the usual criteria for a function to be convex that is given in elementary calculus classes. I'll leave it as an exercise for you to prove it using the preceding results.

Corollary 6.8 Let $I := (a,b)$ with $a, b \in \overline{\mathbb{R}}$ and $f : I \to \mathbb{R}$ be twice-differentiable with $f''(x) \;\geq\; 0$ for every $x \in I$. Then f is convex on I.

Lecture 6; 9/13/2004

Optimization of Convex Functions.

Suppose $I = [a,b]$ is a closed subinterval of $\mathbb{R}$ with $-\infty < a < b < \infty$ and $f : I \to \mathbb{R}$ is a convex function. Then f will be continuous on (a,b) - but it could be discontinuous at the end-points a, b.

Lemma 6.9 Suppose f is continuous and convex on a closed interval I as above. Then the synoptic set $S_c(f) := \{x \in I : f(x) \leq c\}$ is either empty or else is a closed subinterval of I.

Proof. When $x_0, x_1 \in S_c(f)$, let $x_t := (1-t)x_0 + tx_1$. From the definition of convex function

$$f(x_t) \;\leq\; (1-t)f(x_0) + tf(x_1) \;\leq\; c \quad \text{for } 0 \leq t \leq 1.$$

Thus $x_t \in S_c(f)$ for all $0 \leq t \leq 1$ - so $S_c(f)$ is an interval. It is closed as it is the inverse image of the closed subset $(-\infty, c]$ of $\mathbb{R}$.

The essential results about minimizing or maximizing a continuous convex function f on I can be summarized as follows.

Theorem 6.10 Suppose f, I as in lemma and f is not constant on I. Then $\alpha(f;I)$ and $\beta(f;I)$ are finite and there are minimizers and maximizers of f on I. Moreover,
(i). the set of all minimizers of f on $[a,b]$ is a closed subinterval of $[a,b]$, and
(ii). one of the end-points of I will be a maximizer.

Proof: $\alpha := \alpha(f;I),\ \beta := \beta(f;I)$ are finite and there are minimizers and maximizers from Weierstrass' theorem since f is continuous.

(i) 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.

(ii) 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 f has a maximum is at b. Suppose $D_+f(a) < 0$, 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 a or b, we must compare $f(a)$ and $f(b)$ to see which one is $x^{\#}$.

Corollary 6.11 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 minimizers, then $f(\hat{x_1}) = f(\hat{x_2})$, so that

$$f(\frac{\hat{x_1} + \hat{x_2}}{2}) < \frac{1}{2}f(\hat{x_1}) + \frac{1}{2}f(\hat{x_2}) < \alpha \quad \text{as f is strictly convex.}$$

This is impossible unless $\hat{x_1} = \hat{x_2}$ or the minimizer is unique.

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

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

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

We often write $[x,y] := \{(1-t)x + ty :\ 0\ \leq\ t\ \leq\ 1\}$ 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.

The empty set and a singleton are 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 7.1. The only convex subsets of $\mathbb{R}$ are the empty set, singletons or intervals.

Example 7.2. If V is a subspace of $\mathbb{R}^n$ then V is a (closed) convex set.

Example 7.3. 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.

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

$$x = \sum_{j=1}^{J} t_j\, a^{(j)} \quad \text{with each} \quad t_j \ge 0 \;\&\; \sum_{j=1}^{J} t_j = 1. \tag{7.1}$$

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 $J = n + 1$, this is called a *simplex*.

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

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

will all be convex subsets of $\mathbb{R}^n$. Here $\lambda \in \mathbb{R}$.

The union of two convex sets need not be convex. When $\{C_k : k \in \mathcal{K}\}$ is any family of convex sets in $\mathbb{R}^n$, then the intersection $\cap_{k\in\mathcal{K}} C_k$ will again be convex.

The proofs of each of these statements is quite straightforward. Please verify them yourself.

Convex Functions on Convex Sets.

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

$$G(f) := \{(x, f(x)) : x \in C\}.$$

The *epigraph of* f is the set

$$epi(f) := \{(x, z) : x \in C \quad \& \quad z \ge f(x)\}.$$

The function f is said to be *convex* provided $epi(f)$ is a convex set in $\mathbb{R}^{n+1}$. f is said to be *concave* provided $-f$ is convex.

Example 7.4. Showed that a function which is convex in the sense of lecture 5 is convex in this sense.

Example 7.5. The function defined by $f(x) := \begin{cases} c & a \le x \le b \\ \infty & x < a \text{ or } x > b \end{cases}$ is a convex function on $\mathbb{R}$ (with values in $[c, \infty]$).

The function f is said to be *quasi-convex* provided the synoptic sets $S_c(f) := \{x \in C : f(x) \le c\}$ 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 an increasing function such as $f(x) := x^3$ is quasi-convex, but it is not convex.

Operations on Convex Functions

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

$$S(x) := \sum_{j=1}^{n} c_j \ f_j(x) \quad \text{is convex whenever } c_1, c_2, \ldots, c_n \ge 0.$$

In this case we say that f is a positive linear combination of $f_1, f_2, \ldots, f_n$. (P.L.C.)

2. Suppose $\{f_k : k \in \mathcal{K}\}$ is a collection of convex $(-\infty, \infty]$- valued functions on the 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 convex function on C which may take the value $+\infty$. That is the supremum of any family of convex functions is convex.

We will use this property repeatedly to show that various functions are convex.

A convex function $g : C \to \mathbb{R}$ is said to be a *convex minorant* of a function $f : C \to \mathbb{R}$ provided $\quad g(x) \le f(x)$ for all $x \in C$.

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

The function $g(x) := x^2$ is a convex minorant of the function $f(x)$ defined in example 4.1.

Let $\Gamma(f)$ be the set of all convex minorants of a function f on C, then the function $\overline{f}(x) : C \to \overline{\mathbb{R}}$ defined by

$$\overline{f}(x) \ = \ \sup \{g(x) : g \in \Gamma(f)\} \tag{7.2}$$

is called the convex hull of f on C.

Note that from property 2 above, $\overline{f}$ is a convex function on C. Also $\overline{f}(x) \le f(x)$ for all $x \in C$ (since $g(x) \le f(x)$ for each x). Moreover, $\bar{f}$ is the largest such convex function on I - because otherwise it would not be this supremum.

We will often use the convex hull of a given function. Find the convex hulls $\overline{f}$, and sketch the graphs of the functions $f : \mathbb{R} \to \mathbb{R}$ defined by
(i) $f(x) := (x^2 - 1)^2$
(ii) $f(x) := -x^2$
(iii) $f(x) := \cos x$.

Lecture 7; 9/15/2004
Lemma 7.1 Suppose $f : C \to (-\infty, \infty]$ is continuous and $\alpha(f, C)$ is finite. Then $\alpha(f, C) = \alpha(\overline{f}, C)$.
Proof: Good exercise.

Let C be a non-empty convex set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ be a given function. The preceding definition of a convex function is equivalent to the condition given in lecture 6. Namely we have

Lemma 7.2 A function $f : C \to \mathbb{R}$ is convex if and only if

$$f((1-t)x + ty) \le (1-t)f(x) + tf(y) \qquad \text{for all} \quad 0 \le t \le 1, \qquad x, y \in C. \tag{7.3}$$

Proof:: Done in class for the case C=I. The same proof works in general.

The function f is *strictly convex* on C provided inequality holds in (7.3) whenever $x \neq y$ and $0 < t < 1$.

Theorem 7.3 (Jensen's inequality) Suppose $f : C \to \mathbb{R}$ is convex and $\{a^{(j)} : 1 \le j \le J\}$ is a subset of C then

$$f(\sum_{j=1}^{J} t_j\, a^{(j)}) \;\le\; \sum_{j=1}^{J} t_j\, f(a^{(j)}) \tag{7.4}$$

when $(t_1, t_2, \ldots, t_J) \ge 0$ in $\mathbb{R}^J$ and $\sum_{j=1}^{J} t_j = 1$.
Proof: This is done by induction on J. It is lemma 7.2 when J=2.

Most examples of multivariate convex functions that arise in practice involve functions that are constructed from elementary operations including composition. The chain rule property that is usually used is the following.

Theorem 7.4 Let C be a nontrivial convex set in $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ be a convex function with $f(C) \subset I \subset \mathbb{R}$. Here I is an interval and assume $\varphi : I \to \mathbb{R}$ 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 (7.3) holds. Apply φ to both sides of this, then since φ is increasing

$$\varphi(f(x(t))) \;\le\; \varphi((1-t)f(x)) \;+\; tf(y)).$$

Now use the fact that φ is convex, then this RHS is $\le (1-t)\varphi(f(x)) + t\varphi(f(y))$. Thus g is convex as claimed.

The following examples all are real valued functions defined on $\mathbb{R}^n$.

Example 7.6. $f(x) := \langle a, x\rangle + c := c + \sum_{j=1}^n a_j x_j$ is convex on $\mathbb{R}^n$.
Here $a \in \mathbb{R}^n, c \in \mathbb{R}$. This is called an *affine* function when $a \neq 0$.

Example 7.7. Let $p : \mathbb{R}^n \to [0, \infty)$ be a norm as in lecture 1. Then p is convex on $\mathbb{R}^n$.
Proof: Let x, y be any two vectors in $\mathbb{R}^n$ and define $x(t) := (1-t)x + ty$. Then (7.3) holds as, when $0 \leq t \leq 1$,

$$\begin{aligned} p(x(t)) &\leq p((1-t)x) + p(ty) && \text{from the triangle inequality} && (7.5)\\ &= (1-t)p(x) + tp(y) && \text{as p is homogeneous.} && (7.6)\end{aligned}$$

Example 7.8. Let $A := (a_{jk})$ be an $n \times n$ real matrix. The quadratic form associated with A is the function $q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$q(x) := \langle Ax, x\rangle := \sum_{j,k=1}^n a_{jk} x_j x_k. \tag{7.7}$$

Without loss of generality, we can assume that A is a real symmetric matrix. A real symmetric matrix is defined to be *positive semi-definite* or *p.s.d.* if $q(x) \geq 0$ for all $x \in \mathbb{R}^n$. It is *positive definite* or *p.d.* if $q(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Later in section 9, the following theorem will be proved.

Theorem 7.5 Let A be a real symmetric $n \times n$ matrix. Then the function q defined by (7.7) 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.

The theory of eigenvalues of real matrices says that a real $n \times n$ symmetric matrix has n orthogonal eigenvectors corresponding to n real eigenvalues. (The eigenvalues need not be distinct). In section 16, we'll show, using constrained optimization theory, that the symmetric matrix A is *p.s.d.* if and only if all its eigenvalues are positive (that is ≥ 0.) It will be *p.d.* if and only if all its eigenvalues are strictly positive (that is > 0.

Lecture 8; 9/20/2004

8. Multivariate Differentiation

To describe further properties of multivariate functions we need a good theory of multivariate differentiation. The following are the essential definitions.

Let U be an open set in $\mathbb{R}^n$, $x^{(0)} \in U$ and $f : U \to \mathbb{R}$ be a continuous function. The *lower differential* of f at $x^{(0)}$ in the direction h is

$$d_l f(x^{(0)}, h) := \liminf_{t \to 0^+} t^{-1} [f(x^{(0)} + th) - f(x^{(0)})]. \tag{8.1}$$

This always exists and may be $\pm\infty$. 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

$$\lim_{t\to 0+} t^{-1}\ [f(x^{(0)}+th) - f(x^{(0)})]\ =\ d. \tag{8.2}$$

If f has a derivative d at $x^{(0)}$ in the direction h, then $d_l f(x^{(0)}, h) = d$. Often, we just require that $||h|| = 1$.

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

$$\lim_{t\to 0+} t^{-1}\ \left[\left(f(x^{(0)}+th) - f(x^{(0)})\right)\right]\ -\ v\cdot h\ =\ 0 \quad \text{for all } h \in \mathbb{R}^n \tag{8.3}$$

(or just all h with $||h||_2 = 1$). When this holds, the *gradient* of f at $x^{(0)}$ is defined to be $\nabla f(x^{(0)})\ :=\ v$. The jth component of $\nabla f(x^{(0)})$ is the usual partial derivative $\dfrac{\partial f}{\partial x_j}(x^{(0)})$.

Just as in one dimension, there are necessary conditions 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 differential of f or, when f is differentiable, in terms of the derivative of f.

Theorem 8.1 Suppose $x^{(0)}$ is a local minimizer of f on an open set U, then
(i) $d_l f(x^{(0)}, h) \geq 0$ for all $h \in \mathbb{R}^n$, and
(ii) if f is G-differentiable at $x^{(0)}$, then $\nabla f(x^{(0)}) = 0$. (F)

Proof: If $x^{(0)}$ is a local minimizer, then $f(x^{(0)}+th) \geq f(x^{(0)})$ for all $h \in \mathbb{R}^n$ and t small enough. Thus

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

Thus $\liminf_{t\to 0^+} t^{-1}[f(x^{(0)}+th) - f(x^{(0)})]\ \geq\ 0$ or (i) holds.
If f is G-differentiable at $x^{(0)}$, then part (i) implies that $d = \nabla f(x^{(0)})\cdot h \geq 0$ for all $h \in \mathbb{R}^n$. Take $h = \pm e^{(j)}$, then

$$\frac{\partial f}{\partial x_j}(x^{(0)})\ \geq\ 0 \quad \text{and} \quad \frac{\partial f}{\partial x_j}(x^{(0)})\ \leq\ 0,$$

so $\dfrac{\partial f}{\partial x_j}(x^{(0)}) = 0$ for each $j \in \{1, 2, , \ldots, n\}$.
Comments: 1. These are often called *first order conditions* for a local minimizer. Note that (ii) here does not require that f be differentiable at any point except $x^{(0)}$.

2. We often use the function $\varphi(t) := f(x^{(0)}+th)$ defined on an interval $(-\delta_1, \delta_2)$ which includes 0. When f is G-differentiable at $x^{(0)}$, then from (8.3), φ is differentiable at 0 and $\varphi'(0) = \nabla f(x^{(0)})\cdot h$.

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 D, such that

$$\lim_{t \to 0^+} \| \, t^{-1}[F(x^{(0)} + th) - F(x^{(0)})] - Dh \| \; = \; 0 \quad \text{for all } h \in \mathbb{R}^n \tag{8.4}$$

Here we can use any norm on $\mathbb{R}^m$. When this holds, we write $DF(x^{(0)}) = D$. Let

$$F(x) = \begin{pmatrix} F_1(x_1, x_2, ..., x_n) \\ F_2(x_1, x_2, ..., x_n) \\ \cdots \\ F_m(x_1, x_2, ..., x_n) \end{pmatrix} \tag{8.5}$$

then each $\dfrac{\partial F_j}{\partial x_k}(x^{(0)})$ exists and

$$DF(x^{(0)}) = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1}(x^{(0)}) & \dfrac{\partial F_1}{\partial x_2}(x^{(0)}) & \cdots & \dfrac{\partial F_1}{\partial x_n}(x^{(0)}) \\ \dfrac{\partial F_2}{\partial x_1}(x^{(0)}) & \dfrac{\partial F_2}{\partial x_2}(x^{(0)}) & \cdots & \dfrac{\partial F_2}{\partial x_n}(x^{(0)}) \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial F_m}{\partial x_1}(x^{(0)}) & \dfrac{\partial F_m}{\partial x_2}(x^{(0)}) & \cdots & \dfrac{\partial F_m}{\partial x_n}(x^{(0)}) \end{pmatrix}. \tag{8.6}$$

This is called the Jacobian matrix of F at $x^{(0)}$.

When $F(x) = \nabla f(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(x) \\ \dfrac{\partial f}{\partial x_2}(x) \\ \cdots \\ \dfrac{\partial f}{\partial x_n}(x) \end{pmatrix}$, then the derivative matrix is called the Hessian of f and

$$DF(x) \; = \; D^2 f(x) \; = \; \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(x) & \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \cdots & \dfrac{\partial f}{\partial x_n \partial x_1}(x) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dfrac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \dfrac{\partial f}{\partial x_n \partial x_1}(x) \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(x) & \cdots & \dfrac{\partial f}{\partial x_n^2}(x) \end{pmatrix} \tag{8.7}$$

When f is twice continuously differentiable on a neighborhood of x, then $D^2 f(x)$ will be a symmetric $n \times n$ matrix as

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j} \quad \text{for all} \quad 1 \le j, k \le n. \tag{8.8}$$

Theorem 8.2 (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

$$\langle D^2 f(\hat{x})h, h\rangle \ \geq \ 0 \qquad \text{for all } h \in \mathbb{R}^n. \tag{8.9}$$

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 4.3, if $\varphi''(0)$ exists, then it is ≥ 0 Moreover $\varphi''(0) \ = \ \langle D^2 f(\hat{x})h, h\rangle$ from the chain rule. Thus (8.9) holds.

Corollary 8.3: Under the above assumptions, when $\hat{x}$ is a local minimizer of f, and $D^2 f(\hat{x})$ exists then its eigenvalues are all greater than or equal to zero.
Proof: This is a result about quadratic forms - should be covered in an advanced linear algebra class.

Theorem 8.4 (Sufficient Condition) Suppose $\hat{x}$ is a critical point of f, f is C^1 on a neighborhood of $\hat{x}$, $D^2 f(\hat{x})$ exists and there is a $c_1 > 0$ such that

$$\langle D^2 f(\hat{x})h, h\rangle \ \geq \ c_1 ||h||^2 \quad \text{for all} \quad h \in \mathbb{R}^n. \tag{8.10}$$

Then $\hat{x}$ is an isolated, strict local minimizer of f.
Proof: Choose φ as above, then (8.10) implies

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

Assume $||h|| = 1$, then since $\hat{x}$ is a critical point of f, as in the proof of theorem 4.4,

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

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

Similarly when $t < 0$, so $f(\hat{x}+th) - f(\hat{x}) \ \geq \ \frac{c_1}{4} t^2$ for each direction $h \in \mathbb{R}^n$ and $0 < |t| < \delta$. Thus $\hat{x}$ is a strict local minimizer as claimed.
Since $\nabla f(\hat{x}) = 0$ and $\langle D^2 f(\hat{x})h, h\rangle \ \geq \ c_1 ||h||_1^2$, then $\hat{x}$ is an isolated critical point.

Comments 1. The necessary and sufficient conditions for local maximizers have the reverse inequality in (8.9) or (8.10).
2. Condition (8.10) holds iff all the eigenvalues of $D^2 f(\hat{x})$ are greater than c_1.

The preceding argument is based on the following result about differentiation - which is a good exercise in ϵ, δ analysis:

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

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

Proof: The condition on $\psi'(0)$ implies that there is a $\delta > 0$ such that

$$|t| < \delta \quad \text{implies that} \quad t^{-1}\psi(t) \ \geq \ a/2.$$

Rearranging gives the stated result.

We used this with $\psi(t) := \varphi'(t)$, thus $\varphi''(0) \ = \ a > 0$.

When $\hat{x}$ is a critical point of f then we say that $\hat{x}$ is a *degenerate critical point* if either (i) $D^2 f(\hat{x})$ does not exist, or else (ii) the matrix $D^2 f(\hat{x})$ is singular (or 0 is an eigenvalue of $D^2 f(\hat{x})$).

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.*

Lecture 9; 9/22/2004

9. 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 a differentiable function. Then $\nabla f : C \to \mathbb{R}^n$ is well-defined. The following results provide differential criteria for the function to be convex. See also Chapter III, Section 3 of Berkowitz for more information and compare these results with the 1-dimensional results in lecture 6.

Theorem 9.1 Suppose f, C as above, then f is convex on C if and only if

$$f(y) \ \geq \ f(x) \ + \ \langle \nabla f(x), y - x \rangle \quad \text{for all } x, y \in C. \tag{9.1}$$

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 (9.1) holds.

Conversely when (9.1) holds, $x, y \in C, x \neq y$, let $z := (1-t)x + ty$ with $0 < t < 1$. Substitute z for x in (9.1), 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 (7.1) holds.

Corollary 9.2 Suppose f, C as above, then f is strictly convex on C if and only if strict inequality holds in (9.1) when $x \neq y$.
Proof. Straightforward.

In lecture 6, the basic results about 1-dimensional minimization and maximization of convex functions on a convex set C were described. Here we shall describe the results about the unconstrained minimization of a nontrivial convex function f on C.

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 when f is convex. To prove these results we use an elementary result about inner products - which is basic in the theory of equations.

Lemma Suppose $a \in \mathbb{R}^n$ and $\langle a, x\rangle \ \geq \ 0$ for all $x \in \mathbb{R}^n$, then $a \ = \ 0$.
Proof. Suppose $a \neq 0$ and take $x := -a$. Then $-|a|^2 \ \geq \ 0$. The rules for inner products say that this implies $a = 0$.

Theorem 9.3 Suppose f, C as above with f convex. Then $\hat{x}$ minimizes f on C if and only if $\hat{x}$ is a critical point of f. That is iff $\nabla f(\hat{x}) \ = \ 0$.
Proof. When $\hat{x}$ is a critical point of f, then (9.1) 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 $t > 0, d \in \mathbb{R}^n$. Thus $\langle \nabla f(\hat{x}), d\rangle \ \geq \ 0$ for all $d \in \mathbb{R}^n$. This implies $\hat{x}$ is a critical point of f.

Corollary 9.4 Suppose f, C as above with f is convex. The set of all minimizers of f on C is a closed convex set. If f is strictly convex, then this set contains at most one point.
Proof. Just as in 1-d.

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

Theorem 9.5 Suppose f, C as above, then f is convex on C if and only if

$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \ \geq \ 0 \quad \text{for all } x, y, \in C. \tag{9.2}$$

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 1d theory, $\varphi'(1) \geq \varphi'(0)$ so (9.2) holds.

Conversely, when (9.2) 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 (9.1) holds, so f is convex on C. The strictness part is similar.

When (9.2) holds, then $\nabla f(x)$ is said to be a *monotone* mapping of C into $\mathbb{R}^n$.

Example 9.1 Define $f(x) \ := \ (1/2) \ \langle Ax, x \rangle$ for $x \in \mathbb{R}^n$. where A is a symmetric $n \times n$ matrix. Then $\nabla f(x) \ = \ Ax$. Theorem 9.1 says that this function will be convex on $\mathbb{R}^n$ if and only if

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

Theorem 9.5 says that this function will be convex on $\mathbb{R}^n$ if and only if

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

This last statement is the same as saying that $\langle Az, z \rangle \ \geq \ 0 \qquad$ for all $z \in \mathbb{R}^n$. or that the matrix A is p.s.d. - which is what was claimed in Theorem 7.3.

From the last part of theorem 9.5, the function f will be strictly convex on $\mathbb{R}^n$ provided $\langle Az, z \rangle > 0 \qquad$ for all $z \in \mathbb{R}^n \setminus \{0\}$.

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 9.5. 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 that

$$f(y) \ = \ f(x) \ + \ \langle \nabla f(x), y - x \rangle \ + \ (1/2) \langle D^2 f(x(\tau))(y - x), (y - x) \rangle \qquad (9.3)$$

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 9.6 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.

Lecture 10; 9/27/2004

10. Unconstrained Quadratic Optimization Problems.

Many important problems can be written as convex optimization problems. In each case the questions are

(i) does the optimization problem have a global minimum on $\mathbb{R}^n$?

(ii) what equations do the minimzers satisfy? and
(iii) how can we find the minimizers / critical points.

To answer (i) we usually show that f is convex and coercive - or that some synoptic set is non-empty and bounded. To answer (ii) just compute the G-derivatives. (iii) is a matter of developing algorithms for finding the minimizers.

This lecture will treat the two optimization problems usually associated with solving a linear equation of the form

$$Ax \;=\; b \qquad (LE)$$

Here A is an $m \times n$ matrix and b is a given vector in $\mathbb{R}^m$.

Energy methods for linear equations

Example 10.1 Let A be a real symmetric $n \times n$ matrix, $b \in \mathbb{R}^n$ and define $\mathcal{E} : \mathbb{R}^n \to \mathbb{R}$ by

$$\mathcal{E}(x) \;:=\; \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \;=\; \frac{1}{2}\sum_{j,k=1}^{n} a_{jk}x_j x_k - \sum_{j=1}^{n} b_j x_j \qquad (10.1)$$

Consider the problem of finding

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

and the minimizers $\mathcal{M}(\mathcal{E})$ of $\mathcal{E}$ on $\mathbb{R}^n$.

Note that $\nabla\mathcal{E}(x) = Ax - b$, so $\hat{x}$ is a critical point of $\mathcal{E}$ if and only if it is a solution of (LE). Hence finding the critical points of the function $\mathcal{E}$ is equivalent to solving this linear equation.

The first question is about the minimization of $\mathcal{E}$ on $\mathbb{R}^n$. First remember the following definitions from lecture 7. A real symmetric matrix is defined to be *positive semi-definite* or *p.s.d.* if $q(x) \;\geq\; 0$ for all $x \in \mathbb{R}^n$. It is *positive definite* or *p.d.* if $q(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$. Here $q(x) := \langle Ax, x\rangle$. There are three cases.
(a) If A is positive definite, then $\mathcal{E}$ is coercive on $\mathbb{R}^n$ and $\alpha(\mathcal{E})$ is finite, for any $b \in \mathbb{R}^n$.
(b) If A has a negative eigenvalue, then $\alpha(\mathcal{E}) = -\infty$ for any $b \in \mathbb{R}^n$.
(c) If A is p.s.d. with A singular, then the answer depends on b.

The following describes each of these cases.

When (a) holds, then the matrix A is non singular (ie $Ax = 0$ implies $x = 0$) so A^{-1} exists and the unique critical point is $\tilde{x} = A^{-1}b$ so

$$\mathcal{E}(\tilde{x}) \;=\; -\frac{1}{2}\langle A^{-1}b, b\rangle \;=\; -\frac{1}{2} < \tilde{x}, b > .$$

This is the value $\alpha(\mathcal{E})$ when A is positive definite.

When (b) holds, and A is nonsingular, there will be a unique critical point of $\mathcal{E}$ - just as in (a) - but it will not be a minimizer of $\mathcal{E}$ and $\mathcal{E}$ will not be a convex function on $\mathbb{R}^n$. The value $\alpha(\mathcal{E}) = -\infty$. If (b) holds and A is singular, then we again have $\alpha(\mathcal{E}) = -\infty$ and we may, or may not have a solution of (LE).

In case (c), the function $\mathcal{E}$ is convex on $\mathbb{R}^n$, but the infimum may be either $-\infty$ or be a finite number. Specifically let $N(A)$ be the null space (or kernel) of A. Then $\alpha(\mathcal{E})$ is finite if and only if $\langle b, e\rangle = 0$ for all $e \in N(A)$. (Try to prove this; its not hard from ordinary linear algebra).

When $\alpha(\mathcal{E}) = -\infty$, then there is no solution of the equation (LE) as if there is a solution, it would be a critical point of $\mathcal{E}$. When $\mathcal{E}$ has a critical point, then theorem 9.3 says that it is a minimizer of $\mathcal{E}$ as $\mathcal{E}$ is convex. Since the value is $-\infty$, there is no minimizer of $\mathcal{E}$ so this contradiction implies that there is no solution of (LE).

When $\alpha(\mathcal{E})$ is finite, then the following theorem holds - but we need more linear analysis to prove it. We will give the proof later.

Theorem 10.1. Suppose A is an $n \times n$ symmetric matrix and $\mathcal{E}$ is defined by (10.1). If $\mathcal{E}$ is convex and bounded below on $\mathbb{R}^n$, then there is a minimizer of $\mathcal{E}$ on $\mathbb{R}^n$. If $\mathcal{E}$ is strictly convex then this minimizer is unique.

Least Squares Solutions of Linear Equations.

Example 10.2: Let A be a $m \times n$ matrix, $b \in \mathbb{R}^m$ and define $\mathcal{F} : \mathbb{R}^n \to [0, \infty)$ by

$$\mathcal{F}(x) \ := \ \frac{1}{2}||Ax - b||_2^2 \tag{10.2}$$

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

$$\mathcal{F}(x) \ = \ \frac{1}{2}\langle Ax - b, Ax - b\rangle \ = \ \frac{1}{2}\langle A^T Ax, x\rangle - \ \langle Ax, b\rangle \ + \ \frac{1}{2} \ ||b||_2^2 \tag{10.3}$$

Thus $$\mathcal{F}(x) \ = \ \mathcal{E}(x) \ + \ \frac{1}{2} \ ||b||_2^2.$$

Here $\mathcal{E}$ as in Example 10.1, but with $A^T A$ in place of A and $A^T b$ in place of b. This function has derivatives

$$\nabla\mathcal{F}(x) = A^T(Ax - b) \qquad \text{and} \quad D^2\mathcal{F}(x) = A^T A. \tag{10.4}$$

and $\mathcal{F}$ is convex on $\mathbb{R}^n$ from example 9.1 where theorem 7.3 is proved. The critical points of $\mathcal{F}$ will be the solutions of

$$A^T Ax = A^T b. \tag{LS}$$

From theorem 9.3. a vector $x_{LS} \in \mathbb{R}^n$ is a solution of this equation if and only if it minimizes $\mathcal{F}$ on $\mathbb{R}^n$ - since $\mathcal{F}$ is convex.

Since this function is a special case of the function defined in Ex 10.1 with the extra property that it is always positive, the properties of this optimization problem are special cases of the optimization problem for $\mathcal{E}$.

The important result about this least squares problem is that for any matrix A and any $b \in \mathbb{R}^m$, there is a minimizer x_{LS} of $\mathcal{F}$ on $\mathbb{R}^n$. That is, there is a *least squares solution* x_{LS} of (LS). This follows from theorem 10.1 above. Moreover the least squares solution x_{LS} will be a solution of (LE) if and only if $\mathcal{F}(x_{LS}) = 0$.

So if we try minimizing $\mathcal{F}$ on $\mathbb{R}^n$ and find that the minimal value is positive - then there is no solution of (LE). Sometimes this least squares solution is called a generalized solution of (LE).

Suppose $N(A) = \{0\}$. In this case the symmetric matrix $A^T A$ is non-singular and positive definite, as $||Ax||_2 > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$ so that we are in case (a) of the preceding example. Then for any $b \in \mathbb{R}^m$, there will be a unique minimizer x_{LS} of $\mathcal{F}$ on $\mathbb{R}^n$ and it will be the unique solution of (LS).

x_{LS} need not be a solution of (LE) unless we also have that $N(A^T) = \{0\}$. When $dim\ N(A^T) \geq 1$, then a least squares solution of (LS) need not be a solution of (LE).

If $dim\ N(A) \geq 1$, 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

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

Another very interesting unconstrained minimization problem is the following

Example 10.3: Let A, M be real , symmetric $n \times n$ matrices and define

$$\mathcal{G} : \mathbb{R}^n \to \mathbb{R} \quad \text{by} \quad \mathcal{G}(x) := \frac{1}{4}\langle Mx, x\rangle^2 - \frac{1}{2}\langle Ax, x\rangle \tag{10.5}$$

Consider the problem of minimizing $\mathcal{G}$ on $\mathbb{R}^n$. When is this function bounded below? You may want to determine the G-derivatives of this and describe what the possible critical points are.

We'll treat this problem later.

Lecture 11; 9/29/2004

Preconditioned Least Squares Problems.

Example 10.4: 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 $\mathcal{F}_M : \mathbb{R}^n \to [0, \infty)$ by

$$\mathcal{F}_M(x) \ := \ \langle M(Ax - b), Ax - b\rangle/2 \tag{10.6}$$

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

$$\mathcal{F}_M(x) = \frac{1}{2}\langle A^T MAx, x\rangle - \langle MAx, b\rangle + \frac{1}{2}\langle Mb, b\rangle. \tag{10.7}$$

Thus $$\mathcal{F}_M(x) = \mathcal{E}(x) + \frac{1}{2}\langle Mb, b\rangle.$$

Here $\mathcal{E}$ as in Example 10.1, but with $A^T MA$ in place of A and $A^T Mb$ in place of b. When $M := I$, this reduces to the least squares function of example 10.2. If $m = n$ and A is p.d symmetric, then so is A^{-1}, and choosing $M = A^{-1}$ yields the least squares function of example 10.1. Thus this function includes both the previous examples as special cases. M is called a preconditioner for solving the equation (LE).

This function has derivatives

$$\nabla\mathcal{F}_M(x) = A^T M(Ax - b) \qquad \text{and} \quad D^2\mathcal{F}_M(x) = A^T MA. \tag{10.8}$$

and $\mathcal{F}_M$ is convex on $\mathbb{R}^n$ from example 9.1 or theorem 7.3.

11. Algorithms for quadratic minimization

Here we shall describe and analyze some methods for finding the minimizers of quadratic strictly convex functions on $\mathbb{R}^n$. That is we shall look at the problem of example 10.1 with A positive definite and symmetric, $b \in \mathbb{R}^n$. We would like to minimize

$$\mathcal{E}(x) := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle.$$

In this case $\nabla\mathcal{E}(x) = Ax - b$ and $D^2\mathcal{E}(x) = A$ for each $x \in \mathbb{R}^n$.

To develop an iterative algorithm for minimizing $\mathcal{E}$, we first must decide on a stopping criterion. Usually we preselect an $\epsilon > 0$ and stop when $||\nabla\mathcal{E}|| < \epsilon$. The general outline of an algorithm has the following general form. A descent direction for $\mathcal{E}$ at a point x is a vector d such that $\langle \nabla\mathcal{E}(x), d\rangle < 0$. Often we require that $||d|| = 1$.

1. Choose an initial point x^0, and a descent direction d^0.
2. For $k \geq 0$, let
   $\varphi_k(t) := \mathcal{E}(x^k + td^k)$ and choose t^k so that $\varphi_k(t^k) < \varphi_k(0)$.
3. Put $x^{k+1} := x^k + t^k d^k$ and evaluate $||\nabla\mathcal{E}(x^{k+1})||$.

4. If $||\nabla\mathcal{E}(x^{k+1})|| \leq \epsilon$, stop. Else choose a new descent direction d^{k+1} and go to (2).

There are two sets of choices here. The step size (t^k) choice in step 2, and the new direction (d^{k+1}) choice in step 4.

When $\mathcal{E}$ is quadratic as above then the general form of $\varphi_k(t)$ is

$$\varphi(t) := t^2 \langle Ad, d\rangle/2 + t \langle Ax - b, d\rangle + c \tag{11.1}$$

where $c := \mathcal{E}(x)$ is a constant. Often we call $r = r(x) := Ax - b = \nabla\mathcal{E}(x)$ the *residual* of this problem at x.

A direction $d \in \mathbb{R}^n$ is a descent direction for $\mathcal{E}$ at x if and only if $\langle r, d\rangle < 0$. The function $\varphi(t)$ is quadratic in t and its minimum occurs at

$$\tau := -\langle r, d\rangle/\langle Ad, d\rangle > 0.$$

For quadratic problems, t^k in step 2 is usually chosen using this formula.

The simplest choice of the directions in step 1 and 4 is to choose $d := -r(x) = \nabla\mathcal{E}(x)$. This is called the *steepest descent method* and numerically it has poor convergence properties.

Around 1952, Hestenes and Stieffel introduced the conjugate gradient (CG) algorithm for this problem. If we write $r^k := Ax^k - b$ for the k-th residual then the successive directions in the CG method are given by

$$d^0 := -r^0 \quad \text{and for k} \geq 0 \quad d^{k+1} := -r^{k+1} + \beta_k\, d^k \tag{11.2}$$

where

$$\beta_k := \langle Ad^k, r^{k+1}\rangle / \langle Ad^k, d^k\rangle. \tag{11.3}$$

The surprising result about the conjugate gradient algorithm is the following.

Theorem 11.1. Suppose A is an $n \times n$ symmetric p.d. matrix and $\mathcal{E}$ is defined by (10.1). Then the conjugate gradient algorithm will find the unique solution $\hat{x}$ of (LE) in at most n steps.

In practice this is not quite the case as any actual computation involves some round-off error. Nevertheless, the CG method generates good approximate solutions in a reasonable number of steps - independently of the starting point. It is a very effective way to obtain good approximate solutions of very large systems in which you do not want to do all the algebra involved in finding an exact solution.

For more general, convex minimizations, one can look for algorithms that have similar properties to the CG algorithm. They have two parts, the line search algorithm for

a t^k and then a search direction choice. These will be described next semester - together with analyses of convergence of various algorithms.

Lecture 12; 10/4/2004

12. Important Inequalities on $\mathbb{R}^n$

In n-dimensional analysis, we repeatedly use a number of standard inequalities. They include Cauchy's, Holder's, Minkowski's and Young's inequalities. Generally these inequalities are written in terms of norms on $\mathbb{R}^n$ and usually they are proved by using special formulae. Here we shall provide proofs based on optimization methods.

When $1 < p < \infty$ define $p^* := p/(p-1)$ then p^* is called the conjugate, or dual, index to p. The following inequalities hold for all $x, y \in \mathbb{R}^n$:

$$\text{Young} \qquad |\langle x, y\rangle| \leq \frac{\|x\|_p^p}{p} + \frac{\|y\|_{p^*}^{p^*}}{p^*} \tag{12.1}$$

$$\text{Holder} \qquad |\langle x, y\rangle| \leq \|x\|_p \, \|y\|_{p^*} \tag{12.2}$$

$$\text{Minkowski} \qquad \|x + y\|_p \leq \|x\|_p + \|y\|_p \tag{12.3}$$

Cauchy's inequality is the special case of Holders inequality with $p = p^* = 2$. Minkowski's inequality also holds for $p = 1, \infty$. When $p = 1$, $p^* = \infty$, then Holder's inequality holds. These cases, as well as the cases when either x or y is zero, are straightforward.

Theorem 12.1 For any non-zero $x, y \in \mathbb{R}^n, p > 1$, then (12.1) holds. Moreover equality holds in Young's inequality if and only if $y_j = |x_j|^{p-2} x_j$ for all j.

Proof: In your last homework set you showed that when p, p^* as above,

$$|x\ y| \leq \frac{|x|^p}{p} + \frac{|y|^{p^*}}{p^*} \quad \text{for } x, y \in \mathbb{R}$$

and equality holds here if and only if $y = |x|^{p-2} x$. Thus, for vectors,

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

which yields (12.1).

Theorem 12.2 For any non-zero $x, y \in \mathbb{R}^n, p \geq 1$, then (12.2) holds. Equality holds here if and only if $y_j = c|x_j|^{p-2} x_j$ for some $c \in \mathbb{R}$ and all j.

Proof: For $x, y \in \mathbb{R}^n \setminus \{0\}$, let $u := x/\|x\|_p$ and $v := y/\|y\|_{p^*}$. Apply Young's inequality to $\langle u, v\rangle$, then

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

This implies (12.2). The equality condition here follows from that of theorem 12.1.

We are now in a position to prove theorem 1.4 of the first lecture. It said that when $1 \leq p < q \leq \infty$, then

$$\|x\|_q\ \leq\ \|x\|_p\ \leq\ n^{1/p-1/q}\ \|x\|_q \quad \text{for all} \quad x \in \mathbb{R}^n. \tag{12.4}$$

The result is straightforward if $q = \infty$, and the first inequality follows from the montonicity described in theorem 1.1. To prove the second inequality, let $z_j := |x_j|^p$ and $r := q/p > 1$. Then

$$\|x\|_p^p\ =\ \|z\|_1\ =\ \langle z, e\rangle \quad \text{where} \quad e := (1, 1, \ldots, 1). \tag{12.5}$$

Now

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

Apply Holder's inequality to (12.5), then

$$\|x\|_p^p\ \leq\ \|z\|_r\ \|e\|_{r^*}\ =\ \ n^{1-1/r}\ \|x\|_q^p \tag{12.6}$$

Take p-th roots of both sides of this, then (12.4) follows.

It remains to prove Minkowski's inequality. This is usually done by using some identities and Holder's inequality. Here, I will show how certain convex functions can define norms on $\mathbb{R}^n$ in a way that generalizes the usual p-norms. The convexity of these defining functions implies that the triangle inequality holds for the associated norms.

The *positive orthant* in $\mathbb{R}^n$ is the set of all vectors x all of whose components satisfy $x_j \geq 0$. The set of all strictly positive vectors in $\mathbb{R}^n$ is denoted $\mathbb{R}^n_+ :=\ (0, \infty)^n$. The *unit simplex* $\Delta_n \subset \mathbb{R}^n$ is the closed convex hull of the set of points $\{0, e^{(1)}, \ldots, e^{(n)}\}$. Here $e^{(j)}$ is the j-th unit vector in $\mathbb{R}^n$. Δ_n is the set of vectors x which satisfy

$$\sum_{j=1}^{n}\ x_j\ \leq\ 1 \quad \text{and} \quad x_j\ \geq\ 0 \quad \text{for all } j. \tag{12.7}$$

This simplex has $n + 1$ vertices and also $n + 1$ faces (or sides). Let

$$\Delta_n'\ :=\ \{x \in \mathbb{R}^n : x_j\ \geq\ 0 \text{ for all j and } \sum_{j=1}^{n}\ x_j\ =\ 1\} \tag{12.8}$$

Then Δ'_n is the closed convex hull of $\{e^{(1)}, \ldots, e^{(n)}\}$. It also is the intersection of the positive orthant with the unit sphere in $\mathbb{R}^n$ with respect to the 1-norm and is the only face of Δ_n that is not a subset of one of the coordinate hyperplanes $H_j := \{x \in \mathbb{R}^n : x_j = 0\}$. In applications, it is the set of all probability vectors in $\mathbb{R}^n$.

When $x \in \mathbb{R}^n$, we shall now write $|x| := (|x_1|, |x_2|, \ldots, |x_n|)$. Then $|x| = \|x\|_1 \; d$ for some $d \in \Delta'_n$.

Suppose $\psi : [0, \infty)^n \to [0, \infty)$ is a continuous convex function which satisfies

(P1): $\psi(0) = 0, \; \psi(x) > 0$ for $x \neq 0$, and

(P2): $\lim_{t \to \infty} \; \psi(td) \; = \; \infty \quad$ for each $d \in \Delta'_n$.

When ψ obeys (P1) and (P2), define

$$\|x\|_\psi \; := \; inf \; \{\lambda > 0 : \; \psi(|x|/\lambda) \; \leq \; 1\}. \tag{12.9}$$

The p-norms on $\mathbb{R}^n$ are special cases of norms induced by functions of this type.

Example 12.1. For $1 \; \leq \; p < \infty$, define

$$\psi_p(x) \; := \; \sum_{j=1}^{n} \; |x_j|^p. \tag{12.10}$$

Then ψ_p satisfies (P1)-(P2) and the $\psi-$norm associated with this function is the usual p-norm. (Verify this!).

There are many other useful norms on $\mathbb{R}^n$ that arise in applications and statistics. Some of them involve exponentials and logarithms - not just powers of a variable. Many have the following form

Example 12.2. Let $\varphi : [0, \infty) \to [0, \infty)$ be a continuous convex function that satisfies

(P3): $\varphi(0) = 0, \quad \varphi(t) > 0$ for $t > 0 \quad$ and $\quad \lim_{t \to \infty} \; \varphi(t) \; = \; \infty$.

Then the function ψ defined by

$$\psi(x) \; := \; \sum_{j=1}^{n} \; \varphi(|x_j|)$$

obeys (P1)-(P2).

Examples of functions that satisfy (P3) include $t \ln(1 + t)$ and $e^t - 1$ - so these can be used to define norms on $\mathbb{R}^n$.

The following theorem says that the "$\psi-$ norm" induced by a convex function which satisfies (P1)-(P2) is actually a norm. When ψ is the function of example 12.1 it yields (12.3) for $1 \; \leq \; p < \infty$.

Theorem 12.3 (generalized Minkowski inequality) When ψ satisfies (P1)-(P2), then $\|.\|_\psi$ defined by (12.9) is a norm on $\mathbb{R}^n$. Equality holds in the triangle inequality whenever

$y = c\,x$ for some $c \ \geq \ 0$.

Proof: The properties (i) and (ii) of a norm (see lecture 1) are easily verified. Suppose $x, y \in \mathbb{R}^n \setminus \{0\}$ with $\|x\|_\psi = c_1$ and $\|y\|_\psi = c_2$. Let $u := x/c_1,\ v := y/c_2,\ z := (u+v)/2$. Then

$$\frac{x+y}{2\lambda} \ = \ \frac{c_1}{c_1+c_2}\ u \ + \ \frac{c_2}{c_1+c_2}\ v$$

when $\lambda := (c_1+c_2)/2$. Therefore

$$\psi((x+y)/2\lambda) \ \leq \ \frac{c_1}{c_1+c_2}\ \psi(u) \ + \ \frac{c_2}{c_1+c_2}\ \psi(v) \ \leq \ 1$$

as ψ is convex. Thus

$$\|(x+y)/2\|_\psi \ \leq \ \lambda \ = \ (c_1+c_2)/2 \ \leq \ (\|x\|_\psi \ + \ \|y\|_\psi)/2$$

so the triangle inequality holds for this norm on $\mathbb{R}^n$. The criterion for equality is easy to verify.

Lecture 13; 10/6/2004

13. Optimization on a convex set

Let U be an open set in $\mathbb{R}^n$, C be a convex subset of U and $f : U \to \mathbb{R}$ be a continuous function. Consider the problem of characterizing the local minimizers of f on C.

The analog of theorem 8.1 is the following necessary condition for minimization on a convex subset.

Theorem 13.1 Suppose f, U, C as above and $\tilde{x}$ is a local minimizer of f on C. If f is differentiable at $\tilde{x}$, then $\tilde{x}$ satisfies

$$\langle \nabla f(x), y - x \rangle \ \geq \ 0 \qquad \text{for all } y \in C \qquad\qquad (VI)$$

Proof: If $\tilde{x}$ is a local minimizer, then $f((1-t)\tilde{x}+ty) \geq f(\tilde{x})$ for all $y \in C$ and $0 \leq t \leq 1$. Thus $\tilde{x}$ satisfies

$$t^{-1}[f(x+t(y-x)) - f(x)] \ \geq \ 0 \quad \text{for all } y \in C,\ 0 < t < 1.$$

Take the limit as $t \to 0^+$, then (VI) follows.

(VI) is called a variational inequality. When f is convex on C, then this condition is also sufficient.

Corollary 13.2 Suppose f, U, C as above with f convex and differentiable on C. If $\tilde{x}$ satisfies (VI) then it minimizes f on C.

Proof: Since f is convex and differentiable on C then, from theorem 9.1, we have that

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

If $\tilde{x}$ satisfies (VI), this implies $f(y) \ \geq \ f(\tilde{x})$ for all y in C so the corollary holds.

If we seek local maximizers instead of minimizers, the sign in (VI) is reversed and one obtains the following result - whose proof is similar to that of the theorem.

Corollary 13.3 Suppose f, U, C as above and $x^{(1)}$ is a local maximizer of f on C. If f is differentiable at $x^{(1)}$, then $x^{(1)}$ satisfies

$$\langle \nabla f(x), y - x \rangle \ \leq \ 0 \qquad \text{for all } y \in C \qquad (VIX)$$

Lecture 14; 10/11/2004

14. Optimization with Linear Equality Constraints.

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a given differentiable function. Let A be a $m \times n$ matrix and $b \in \mathbb{R}^m$ be in the range of A. Assume $rankA = m < n$ and define K to be the set of all solutions of the equation

$$Ax \ = \ b \qquad (14.1)$$

Write $A^T \ := \ [a^{(1)}, a^{(2)}, \ldots, a^{(m)}]$ so that the $a^{(j)}$ are the row vectors of A. Then (14.1) is equivalent to requiring that

$$\langle a^{(j)}, x \rangle \ = \ b_j \quad \text{for all} \quad 1 \leq j \leq m.$$

This is a system of m *linear equality constraints*. The rank condition implies that the vectors $a^{(j)}$ are linearly independent. Let $\tilde{x}$ be a solution of (14.1), then $K \ = \ \{\tilde{x} \ + \ z : z \in N(A)\}$. The null space of A will have dimension $n - m$. Consider the problem $(\mathcal{P})$ of minimizing f on K and finding

$$\alpha(f, K) \ := \ inf_{x \in K} f(x) \qquad (14.2)$$

The basic existence results about this may be summarized as follows.
(i) If f is weakly coercive on $\mathbb{R}^n$, then $\alpha(f, K)$ is finite and there is an $\hat{x} \in K$ which minimizes f on K.
(ii) If f is (strictly) convex on $\mathbb{R}^n$ and there is a minimizer of f on K, then the set of all minimizers is convex (a singleton).

This problem can be converted into an unconstrained problem on $N(A)$. Namely define $g : N(A) \to \mathbb{R}$ by

$$g(z) \ := \ f(\tilde{x} \ + \ z) \quad \text{where } \tilde{x} \text{ as above.} \tag{14.3}$$

Then

$$\alpha(f, K) \ = \ \alpha(g, N(A)) \ := \ inf_{z \in N(A)} g(z).$$

When there are constraints, then the conditions obeyed at a local minimizer, or maximizer, will be different. The simplest criterion is the following:

Theorem 14.1 Suppose f, K as above and $\tilde{x}$ is a local minimizer of f on K, then $\tilde{x}$ satisfies

$$\langle \nabla f(x), z \rangle \ = \ 0 \qquad \text{for all } z \in N(A). \tag{14.4}$$

Moreover if f is convex on K and $\tilde{x}$ satisfies (14.4), then $\tilde{x}$ minimizes f on K.

Proof: If $\tilde{x}$ is a local minimizer, then it satisfies (VI) for all $y \in K$. Thus (LVI) holds as $\tilde{x} \pm z \in K$ for all $z \in N(A)$. The last sentence follows from corollary 13.2 applied to this case.

An equivalent form of this result is the following.

Theorem 14.2 (Linear Lagrange multiplier rule) Suppose f, K as above and $\tilde{x}$ is a local minimizer of f on K, then there is a $\lambda \in \mathbb{R}^m$ such that $\tilde{x}$ satisfies

$$\nabla f(x) \ = \ A^T \lambda \ = \ \sum_{j=1}^{m} \ \lambda_j \ a^{(j)} \qquad\qquad (LMR)$$

Moreover if f is convex on K and $\tilde{x}$ satisfies (LMR), then $\tilde{x}$ minimizes f on K.

Proof: Theorem 14.1 says that if $\tilde{x}$ is a local minimizer of f on K then $\nabla f(\tilde{x})$ is orthogonal to N(A). Thus there is a $\lambda \in \mathbb{R}^m$ such that $\nabla f(\tilde{x}) = A^T \lambda$ from the fundamental theorem of linear algebra. (This says that $N(A) \oplus R(A^T) \ = \ \mathbb{R}^n$.) Thus (LMR) holds and the second part holds just as in theorem 13.4.

Example 14.1 Take $n \ = \ 2, \ f(x) \ := \ x_1^2 + 2x_2^2$ and minimize this subject to the constraint $2x_1 \ + \ x_2 \ = \ 2$. Do this both by elimination and by the Lagrange multiplier rule and verify that you obtain the same minimizer and value for this problem.

Least Norm Solutions of Linear Equations.

An important example of minimization on an affine subspace is the problem of minimizing $f(x) \ := \ \|x\|_2^2$ on the set K defined by (14.1). If x_{LN} minimizes f on K, then it will be the solution of (14.1) of least 2-norm - or the point closest to the origin in the Euclidean metric on $\mathbb{R}^n$.

The G-derivative of f is $\nabla f(x) = 2x$, so x_{LN} minimizes f on K if and only if it is in K and satisfies $2x \;=\; A^T\,\lambda$ for some $\lambda \in \mathbb{R}^m$. That is λ satisfies

$$A\,A^T\,\lambda \;=\; 2\,b \tag{14.5}$$

Equivalently λ minimizes the functional $\mathcal{E}$ on $\mathbb{R}^m$ where

$$\mathcal{E}(\lambda) \;:=\; \|A^T\,\lambda\|^2 - 4\langle b, \lambda\rangle \tag{14.6}$$

Since A^T has rank m, the equation (14.5) has a unique solution $\hat{\lambda}$ and then $x_{LN} := A^T\,\hat{\lambda}$ is the least norm solution of (14.1). Moreover, from (14.4), x_{LN} is orthogonal to N(A), so we can write

$$K \;=\; \{x_{LN}\} \oplus N(A) \tag{14.7}$$

where $\oplus$ means that we have an orthogonal sum here.

Example 14.2 Take $d(x) \;:=\; \|x\|_2^2$ and minimize this subject to the constraint $\sum_{j=1}^n\; x_j \;=\; 1$. That is find the point on this hyperplane closest to the origin. We have just one linear constraint equation in n variables.

From theorem 14.2, there is a real number λ such that the minimizer satisfies

$$\nabla d(x) \;=\; 2x \;=\; \lambda(1, 1, \ldots, 1).$$

Summing over the entries yields $\lambda = 2/n$ so the minimizer is $\tilde{x} \;=\; n^{-1}\;(1, 1, \ldots, 1)$ and $d(\tilde{x}) \;=\; 1/\sqrt{n}$.

Lecture 15; 10/13/2004

15. Tangent and Normal Cones of a Convex Set

The description of the systems satisfied by solutions of optimization problems on convex sets depends on the theory of convex cones in $\mathbb{R}^n$. A convex subset K of $\mathbb{R}^n$ is said to be a *convex cone* provided that whenever $x \in K$, then $tx \in K$ for all $t > 0$.

A convex cone is said to be *pointed* if $0 \in K$. When K is a convex cone so is $K \cup \{0\}$. The set $\{0\}$ is a cone - the trivial cone. Any non-trivial cone is an unbounded convex subset of $\mathbb{R}^n$.

When K is a pointed cone in $\mathbb{R}^n$, then $V := K \cap (-K)$ is a subspace of K - and it is the maximal subspace of K. A cone is said to be *strict* if $K \cap (-K)$ is empty or $\{0\}$. The closed positive orthant $\mathbb{R}^n_+ := [0, \infty)^n$ is a strict cone.

When K is a cone, so is cK for any real number c. Note that $cK = K$ for $c > 0$, and $cK = -K$ for $c < 0$. When K_1, K_2 are cones, so also are $K_1 \cap K_2$ and $K_1 \;+\; K_2$. In fact the intersection of any family of convex cones is again a convex cone.

Let $S := \{d^{(1)}, d^{(2)}, \ldots, d^{(m)}\}$ be a finite subset of $\mathbb{R}^n$. A vector x is said to be a *positive linear combination (p.l.c.)* of the elements of S provided

$$x \;=\; \sum_{j=1}^{m} \mu_j \, d^{(j)} \qquad \text{with all the} \quad \mu_j \;\geq\; 0. \tag{15.1}$$

This is a *strictly positive linear combination (s.p.l.c.)* if all the $\mu_j > 0$.

When S is any nonempty subset of $\mathbb{R}^n$, then $\mathcal{K}(S)$ will be the smallest convex cone that contains S. It is always well-defined and non-empty and contains the set of all strictly positive linear combinations of finite subsets of S. The closure of $\mathcal{K}(S)$ is called the closed convex cone generated by S.

A convex cone is said to be *polyhedral* if it is the convex cone generated by a finite set. That is there is a finite set of vectors such that every vector in K has the form (15.1). When $n \;\geq\; 3$, there are cones which are not polyhedral.

Henceforth let C be a convex set in $\mathbb{R}^n$ which contains at least 2 points. For each $x \in C$, define the
tangent cone $T_x(C)$ of C at x to be the closure of the convex cone generated by $C - \{x\}$.
normal cone $N_x(C)$ of C at x to be the set of all vectors $d \in \mathbb{R}^n$ which satisfy

$$\langle d, y - x \rangle \;\leq\; 0 \quad \text{for all} \;\; y \in C. \tag{15.2}$$

Both of these are closed convex cones.

Example 15.1 Let Δ_2 be the unit simplex in $\mathbb{R}^2$. It is the triangle whose vertices are $(0,0), (1,0), (1,0)$. Equivalently it is the set in $\mathbb{R}^2$ defined by the three inequalities

$$x_1 \;\geq\; 0, \quad x_2 \;\geq\; 0, \quad \text{and} \quad x_1 \;+\; x_2 \;\leq\; 1$$

If x is an interior point of the triangle, then $T_x(C) \;=\; \mathbb{R}^2$ and the normal cone is trivial. At the origin the tangent cone is $T_0(C) \;=\; \mathbb{R}^2_+$ and the normal cone is $-\mathbb{R}^2_+$. At the other two vertices of the triangle the tangent cone is the cone of all directions that point into the triangle. (Write down the algebraic formulae for these sets.) The normal cones are not what you might guess - they are not the negative of the tangent cones.

In general suppose that a convex set $C \subset \mathbb{R}^n$ is defined to be the set of all points which satisfy the inequalities

$$\langle a^{(j)}, x \rangle \;\leq\; b_j \qquad \text{for} \quad 1 \;\leq\; j \;\leq\; m. \tag{15.3}$$

This can be written as the matrix inequality

$$Ax \;\leq\; b \tag{15.4}$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. Here m may be (much) larger than n. The set C is called the *feasible set* associated with the inequalities (15.3). For example a box B in $\mathbb{R}^n$ is defined by the 2n inequalities

$$c_j \ \leq \ x_j \ \leq \ d_j \qquad \text{for} \quad 1 \ \leq \ j \ \leq \ n.$$

(Exercise: Write this system of inequalities in the forms (15.3) and also (15.4) - in a way that produces simple vectors $a^{(j)}$ and a nice matrix A).

When $\tilde{x} \in C$ we say that the j-th inequality in (15.3) is *active at* $\tilde{x}$ provided equality holds in the inequality. Otherwise it is inactive. Let $\mathcal{J}(\tilde{x})$ be the set of indices of the active inequalities for C at $\tilde{x}$. Then a vector z is in $T_{\tilde{x}}(C)$ if and only if it satisfies

$$\langle a^{(j)}, z \rangle \ \leq \ 0 \qquad \text{for each} \quad j \in \mathcal{J}(\tilde{x}). \tag{15.5}$$

It often requires a lot of algebra to determine this tangent cone from the specification of C via the inequalities (15.3). However there is a simple description of the normal cone to C at $\tilde{x}$. It is the following

Theorem 15.1 Suppose C as above and x is a point in C. Then the normal cone $N_x(C)$ of C at x is the closed convex cone generated by the active constraints at x.

The proof of this is purely algebraic and quite straightforward - assuming you have worked with cones and linear inequalities. I will not give it here. It may be found in any text that treats the theory of linear inequalities. It says that a vector z is in $N_x(C)$ if and only if it is a p.l.c. of the vectors $\{a^{(j)} : j \in \mathcal{J}(x)\}$, or that

$$z \ = \ \sum_{j \in \mathcal{J}(x)} \mu_j \, a^{(j)} \qquad \text{with all the} \quad \mu_j \ \geq \ 0.$$

The importance of the this result is that it provides a different form for the necessary optimization conditions of theorem 13.1 and corollary 13.3. This result is a form of what is often called the Karush-Kuhn-Tucker (KKT) theorem. The equations in the next theorem are called the *extremality conditions* satisified by local minimizers or maximizers of f on K. Note that when there are inequality constraints, the condition may be different for a local maximizer than for a local minimizer.

Theorem 15.2 (KKT) Suppose f, U as in section 13, C is a nonempty convex subset of U, $\tilde{x}$ is a local minimizer of f on C and f is differentiable at $\tilde{x}$, then $\tilde{x}$ satisfies

$$\nabla f(x) \ + \ z \ = \ 0 \qquad \text{for some } z \in N_x(C). \tag{15.6}$$

If $\tilde{x}$ is a local maximizer, then it satisfies

$$\nabla f(x) \ = \ z \qquad \text{for some } z \in N_x(C). \tag{15.7}$$

When the tangent cone $T_x(C)$ to C at x is defined by (15.5), then $\tilde{x}$ is a local maximizer (minimizer) of f on C provided

$$(-)\nabla f(x) \;=\; \sum_{j\in\mathcal{J}(x)} \mu_j\, a^{(j)} \qquad \text{with all the} \quad \mu_j \;\geq\; 0. \tag{15.8}$$

Proof: If $\tilde{x}$ is a local minimizer, then it satisfies (VI). This implies that $-\nabla f(x) \in N_x(C)$, so (15.6) follows. Then theorem 14.1 gives the last part of this theorem.

The coefficients μ_j are called the Kuhn-Tucker, or inequality, multipliers associated with the active constraints at x. Quite often the condition (15.8) for a minimizer of f on C is stated in the form that there is a $\mu \in {\mathbb{R}_+}^m$, such that $\tilde{x}$ satisfies

$$\nabla f(x) \;+\; \sum_{j=1}^{m} \mu_j\, a^{(j)} \;=\; 0 \tag{15.9}$$

with $\mu_j = 0$ when the j-th constraint is inactive at $\tilde{x}$.

One consequence of this result is that somewhat different conditions hold at a local minimizer of f on C depending on which constraints are "active" at the point. Inactive constraints are not "observed" by the extremal equations.

Lecture 16; 10/20/2004

16. Optimization subject to a Single Convex Inequality Constraint.

Let $g : \mathbb{R}^n \to \mathbb{R}$ be a differentiable, convex function and $C = S_c(g) \;:=\; \{x \in \mathbb{R}^n : g(x) \;\leq\; c\}$ be a synoptic set with non-empty interior. Then C is a closed convex set. Assume that the boundary of C is

$$\partial C \;=\; L_c(g) \;=\; \{x \in \mathbb{R}^n : g(x) \;=\; c\}. \tag{16.1}$$

(This is an extra assumption; in general LHS $\subset$ RHS here.)

If $g(x) < c$, then $T_x(C) = \mathbb{R}^n$ as g is continuous and x is an interior point of C

Suppose $x \in \partial C$ so $g(x) = c$. If v satisfies $\langle \nabla\, g(x), v\rangle < 0$, we have $g(x + tv) < c$ for t small enough and positive as g is differentiable at x. Thus $v \in T_x(C)$. Assume $\nabla\, g(x) \neq 0$, this implies that

$$N_x(C) \;=\; \{\mu \nabla\, g(x) : \mu \;\geq\; 0\}. \tag{16.2}$$

This leads to the following special case of the KKT theorem.

Theorem 16.1 Suppose f, g are G-differentiable real valued functions on $\mathbb{R}^n$ with g convex. Define C as above and assume it has nonempty interior and (16.1) holds. If $\tilde{x}$ is a local minimizer of f on C and $\nabla g(\tilde{x}) \neq 0$, then $\tilde{x}$ satisfies

$$\nabla f(x) + \mu \nabla g(x) = 0 \quad \text{for some } \mu \geq 0, \quad \text{and} \tag{16.3}$$

$$\mu(g(x) - c) = 0. \tag{16.4}$$

When $\hat{x}$ is a local maximizer of f on C, then this system still holds except that now $\mu \leq 0$.

Proof: Under these assumptions, theorem 13.1 says that $\tilde{x}$ satisfies (VI). That is $-\nabla f(x) \in N_x(C)$, so (16.3) follows from (16.2) when $g(\tilde{x}) = c$. If $g(\tilde{x}) < c$, then (16.3) holds with $\mu = 0$. These results may be combined as (16.3) and (16.4).

We shall now use this theorem to describe the characterization of eigenvalues and eigenvectors of a real symmetric $n \times n$ matrix A by optimization methods.

Let $q : \mathbb{R}^n \to \mathbb{R}$ be the quadratic form associated with A as defined in example 7.7. Take $g : \mathbb{R}^n \to \mathbb{R}$ to be the function defined by

$$g(x) := \|x\|_2^2 := \sum_{j=1}^{n} |x_j|^2. \tag{16.5}$$

Let B_1 be the unit ball in $\mathbb{R}^n$ defined by $B_1 := \{x \in \mathbb{R}^n : \|x\|_2^2 \leq 1\}$. Consider the problem of minimizing, and maximizing, q on B_1 and define

$$\alpha(q) := inf_{x \in B_1}\, q(x), \qquad \beta(q) := sup_{x \in B_1}\, q(x). \tag{16.6}$$

Obviously we have $\alpha(q) \leq 0 \leq \beta(q)$ as $0 \in B_1$. These maxima and minima turn out to be related to the eigenvalues of A. A eigenvector e of A is said to be normalized if $\|e\|_2 = 1$. When $e^{(j)}$ is a normalized eigenvector of A corresponding to the eigenvalue λ_j, then $q(te^{(j)}) = t^2\lambda_j$. Thus

$$\alpha(q) \leq \lambda_1 \leq \lambda_n \leq \beta(q)$$

where λ_1, λ_n are the least, respectively largest, eigenvalues of A.

In fact the first and last inequalities will often be equalities - in particular when $\alpha(q) < 0$ or $\beta(q) > 0$ respectively. The following result gives the conditions for this. It also provides a proof that there are real eigenvalues and eigenvectors of a real symmetric matrix, as well as a characterization of the smallest and the largest eigenvalues.

Theorem 16.2 Suppose A is a real symmetric $n \times n$ matrix, q and B_1 are defined as above. Then $\alpha(q), \beta(q)$ defined above are finite and there are vectors $\hat{x}, \tilde{x} \in B_1$ such that $\alpha(q) = q(\tilde{x})$ and $\beta(q) = q(\hat{x})$. Moreover,
(a) if $\tilde{x} \neq 0$, then it is an eigenvector of A corresponding to the eigenvalue $\alpha(q)$, and

this is the least eigenvalue of A, and
(b) if $\hat{x} \neq 0$, then it is an eigenvector of A corresponding to the eigenvalue $\beta(q)$, and this is the largest eigenvalue of A.

Proof: The unit ball B_1 is compact in $\mathbb{R}^n$, so these minimization and maximization problems have finite values which are attained from Weierstrass' theorem.

First consider the case of minimization of q on B_1. If the infimum is attained at an interior point of B_1, then from theorem 15.1, the minimizer satisfies $Ax = 0$. Thus either $\tilde{x} = 0$ or it is an eigenvector of A corresponding to the eigenvalue 0. In either case there cannot be any negative eigenvalues λ of A, as if u is a corresponding iegenvector of norm 1 then $q(u) = \lambda < 0$ which contradicts the assumption that $\tilde{x}$ minimizes q on B_1. If the infimum is attained at a vector $\tilde{x}$ on the boundary ∂B_1, then theorem 15.1 says that $\tilde{x}$ satisfies

$$Ax \ + \ \mu x \ = \ 0 \qquad \text{for some} \ \ \mu \geq 0. \tag{16.7}$$

Take inner products with $\tilde{x}$ then $q(\tilde{x}) := \alpha(q) = -\mu$, so $-\mu$ is an eigenvalue of A and it is less that or equal to 0. It must be the least eigenvalue of A, as if there were a more negative eigenvalue, that would yield the minimium value of q on B_1. Hence (a) holds. Similar arguments prove (b).

This result did not use the fundamental theorem of algebra, or related results, to show that there are eigenvalues or eigenvectors. It just depends on differential calculus and optimzation theory. The following corollary is often used and its generalizations are very important in analysis.

Corollary 16.3 Suppose A is a real symmetric $n \times n$ matrix and the least eigenvalue of A is λ_1. Then

$$\langle Ax, x \rangle \ \geq \ \lambda_1 \|x\|_2^2 \qquad \text{for all } x \in \mathbb{R}^n. \tag{16.8}$$

Proof: If A has λ_j as an eigenvalue then $A - kI$ has $\lambda_j - k$ as an eigenvalue. Thus the least eigenvalue of $B \ := \ A - \lambda_1 I$ will be zero. B is again symmetric. From the theorem $\alpha(q, B) = 0$, or $\langle Bx, x \rangle \geq 0$ for all $x \in B_1$. That is (16.8) holds for all $x \in B_1$. If $\|z\|_2 > 1$, then normalize z and one sees that (16.8) holds on $\mathbb{R}^n$.

(16.8) shows that the quadratic form q will be p.d. on $\mathbb{R}^n$ if and only if the least eigenvalue of A is strictly positive. It will be p.s.d when the least eigenvalue of A is 0. This result was stated earlier without proof. The analysis above shows why the eigenvalues of A are connected with the quadratic form q. This was known for centuries before linear algebra was invented and has interpretations in both geoemtry and mechanics - where it is associated with the theory of *principal axes*.

Lecture 17; 10/25/2004

17. Penalty Methods for problems with inequality constraints.

Today we'll describe penalty methods for modifying inequality constrained problems and use them to prove a result about the conditions satisfied by their local minimizers.

Suppose $\{g_1, \ldots, g_J\}$ are continuously differentiable real valued functions on $\mathbb{R}^n$ and define

$$S \;:=\; \{x \in \mathbb{R}^n : g_j(x) \;\leq\; 0 \quad \text{for } 1 \leq j \leq J\} \tag{16.1}$$

S is a closed set. For each point $x \in S$, let $\mathcal{J}(x)$ be the indices of the active constraints at x. That is, $\mathcal{J}(x) \;:=\; \{j \;:\; g_j(x) \;=\; 0\}$. A point $x \in S$ is said to be a *regular point* for the constraints defining S providing the set $\{\nabla g_j(x) \;:\; j \in \mathcal{J}(x)\}$ is linearly independent. Any interior point of S is a regular point.

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and we want to find the *extremality conditions* obeyed by the local minimizers of f on S. When S is not a convex set, we usually need some extra conditions on the functions g_j to obtain such equations. There are many different such conditions that provide different results. See section 5 of chapter 4 of Berkovitz for a description of some different results and, in particular, of the theory of the Fritz John condition. Here we'll prove a general version of the KKT condition. The result is the following

Theorem 17.1 (KKT2) Suppose f, S as above and $\tilde{x}$ is a local minimizer of f on S. If $\tilde{x}$ is a regular point for the inequality constraints, then there is a unique positive vector μ such that $\tilde{x}$ satisfies

$$\nabla f(x) \;+\; \sum_{j \in \mathcal{J}(x)} \mu_j \, \nabla \, g_j(x) \;=\; 0 \qquad \text{with all the} \quad \mu_j \;\geq\; 0. \tag{16.2}$$

The condition that the only nonzero μ_j correspond to $j \in \mathcal{J}(x)$ is often written

$$\sum_{j=1}^{J} \mu_j \, g_j(x) \;=\; 0 \tag{16.3}$$

This is a system of $n+1$ equations and J inequalities that must hold at the local minimizer. Usually there are $n + J_1$ unknowns, namely the n entries in $\tilde{x}$ and J_1 possible positive values of the μ_j. Here J_1 is the number of indices in $\mathcal{J}(\tilde{x})$.

We shall prove this result by using a penalty function formulation. Given a local minimizer $\tilde{x}$ of f on S and an $\epsilon > 0$, consider the function

$$F_k(x) \;:=\; f(x) \;+\; \frac{k}{2} \sum_{j=1}^{J} g_{j+}(x)^2 \;+\; \frac{\epsilon}{2} \, \|x - \tilde{x}\|_2^2 \tag{16.4}$$

Here $g_{j+}(x) := max(0, g_j(x))$. Note that $g_{j+}(x)^2$ will be G-differentiable on $\mathbb{R}^n$ with

$$\nabla(g_{j+}(x)^2) = 2\, g_{j+}(x)\nabla g_j(x) \quad \text{whenever } g_j(x) > 0. \tag{16.5}$$

Let $B_1 := \{x \in \mathbb{R}^n : \|x - \tilde{x}\|_2 \leq \delta\}$ be a closed ball in $\mathbb{R}^n$ of radius δ and center $\tilde{x}$.

Consider the *penalized problem* of minimizing F_k on B_1. This is a problem of minimizing a continuously differentiable function on a closed ball in $\mathbb{R}^n$; so the constraints are easy to work with. This problem always has a solution from Weierstrass' theorem. Let $x^{(k)}$ be a minimizer of this problem and $\Gamma := \{x^{(k)} : k \geq 1\}$ be a corresponding sequence of minimizers.

Lemma 17.2 (Convergence) Suppose f, S, F_k, B_1 as above and $\tilde{x}$ is a local minimizer of f on S. If Γ is a sequence of minimizers of F_k on B_1, then $x^{(k)}$ converges to $\tilde{x}$ as $k \to \infty$.

Proof: Note that $F_k(\tilde{x}) = f(\tilde{x})$ for all k, so

$$F_k(x^{(k)}) \leq f(\tilde{x}) \quad \text{for all } k \geq 1 \tag{16.6}$$

Define $\alpha_1 := inf_{x \in B_1}\, f(x)$. Then the last inequality and the definition of F_k show that

$$\alpha_1 + \frac{k}{2} \sum_{j=1}^{J} g_{j+}(x^{(k)})^2 \leq f(\tilde{x}) \quad \text{for all } k \geq 1$$

Since $k \nearrow \infty$, this implies that

$$\sum_{j=1}^{J} g_{j+}(x^{(k)})^2 \to 0^+ \quad \text{as } k \to \infty.$$

Thus each $g_{j+}(x^{(k)}) \to 0$ as $k \to \infty$. Let $\hat{x}$ be a limit point of the sequence Γ. Then since each g_j is continuous, we have $g_{j+}(x^{(k)}) \to g_{j+}(\hat{x})$. Thus $\hat{x} \in S$. Take limits in (16.6), then

$$f(\hat{x}) + \frac{\epsilon}{2}\|\hat{x} - \tilde{x}\|_2^2 \leq f(\tilde{x}) = \alpha(f, S).$$

This can only happen if $\hat{x} = \tilde{x}$, so the result holds.

Proof of theorem 17.1:
A consequence of this result is that for sufficiently large $k, x^{(k)}$ will be in the interior of B_1. When this holds, then $x^{(k)}$ will satisfy

$$\nabla f(x) + k \sum_{j=1}^{J} g_{j+}(x)\nabla g_j(x) + \epsilon(x - \tilde{x}) = 0 \tag{16.7}$$

If $x^{(k)} \in S$ then we must have $x^{(k)} = \tilde{x}$ and this formula implies that $\nabla f(\tilde{x}) = 0$ which has the form (16.2).

Otherwise, define $\mu_j^k := kg_{j+}(x^{(k)})$ and observe that for k sufficiently large $\mathcal{J}(x^{(k)}) = \mathcal{J}(\tilde{x})$. Then (16.7) implies that

$$\nabla f(x^{(k)}) \; + \; \sum_{j=1}^{J} \; \mu_j^k \; \nabla g_j(x^{(k)}) \; = \; \epsilon \; (\tilde{x} - x^{(k)}) \tag{16.8}$$

This is a linear equation of the form $D_k \; \mu \; = \; c^k$ where D_k is an $n \times J_1$ matrix of rank J_1 whose columns are the active constraints at $x^{(k)}$. μ is a column vector with J_1 entries and $c^k := \epsilon(\tilde{x} - x^{(k)}) - \nabla f(x^{(k)})$. Premultiply this equation by D_k^T, to obtain

$$D_k^T D_k \; \mu \; = \; D_k^T c^k$$

The matrix on this left hand side is a $J_1 \times J_1$ matrix that has rank J_1 as $\tilde{x}$ is a regular point of the constraints. This equation has a unique solution for the vector μ^k. Let $k \to \infty$, then since the functions are continuously differentiable, D_k converges to $\tilde{D}$ and $c^k \to \tilde{c}$. Thus the multipliers μ^k converge to

$$\tilde{\mu} \; := \; (\tilde{D}^T \tilde{D})^{-1} \tilde{D}^T \tilde{c}$$

Now take limits as $k \to \infty$ in (16.8), each term converges to a limit and the RHS goes to zero, so (16.2) holds at the local minimizer $\tilde{x}$ of f on S and μ is unique.

This choice of a penalized function is good for showing that a local minimizer of f on S which is regular, must satisfy the KKT equations. Sometimes penalized functions like this are used to develop numerical algorithms for approximating the local minimizers of f on S.

A function $\psi : \mathbb{R} \to [0, \infty)$ is called an *exterior penalty function* provided ψ is continuous and increasing with
(i) $\psi(s) \equiv 0$ for $s \leq 0$, and $\psi(s) > 0$ for $s > 0$, with
(ii) $\psi(s) \to \infty$ as $s \to \infty$.

When the set S is described by inequality constraints as above, define

$$\Psi(x) \; := \; \sum_{j=1}^{J} \; \psi(g_j(x)).$$

Then $\Psi(x) \equiv 0$ on S and $\Psi(x) > 0$ on $\mathbb{R}^n \setminus S$. Consider the penalized function $P_\sigma f(x)$ defined by

$$P_\sigma f(x) \; := \; f(x) \; + \; \sigma \Psi(x) \; = \; f(x) \; + \; \sigma \; \sum_{j=1}^{J} \; \psi(g_j(x))$$

The associated penalized problem is to minimize $P_\sigma f(x)$ either on all of $\mathbb{R}^n$, or else on some simple subset such as a ball or a box that is of interest. As $\sigma \nearrow \infty$, one might

expect the solutions of this problem will converge to solutions of the original problem of minimizing f on S. This often is true and leads to good algorithms for some problems.

Lecture 18; 10/27/2004

Example 17.1 Let S be the region in the plane defined by the 5 inequalities

$$0 \le x_1 \le 2, \quad 0 \le x_2 \le 2, \quad \text{and} \qquad x_1^2 + x_2^2 \ge 1.$$

Suppose that a differentiable function f attains its minimum on S at the point $P = (1, 0)$, then what can be said about $\nabla f(1, 0)$? Repeat this problem when the minimum is at $Q = (1, 2)$.

Answer. First note that S is not convex so we must use the theory of section 16. Rewrite the inequalities in the standard form (16.1). Then

$$-x_1 \le 0, \ -x_2 \le 0, \ x_1 - 2 \le 0, \ x_2 - 2 \le 0, \ 1 - x_1^2 - x_2^2 \le 0.$$

Minimization at P. At P, the two active constraints are $x_2 = 0$ and $x_1^2 + x_2^2 = 1$. The gradients of these constraints at P are

$$d^{(1)} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad d^{(2)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

Then the KKT2 theorem says that, since $d^{(1)}, d^{(2)}$ are l.i. when f is minimized at P, there are positive μ_1, μ_2 such that

$$\nabla f(1, 0) + \mu_1 d^{(1)} + \mu_2 d^{(2)} = 0$$

Thus

$$\frac{\partial f}{\partial x_1}(1, 0) = 2\mu_2 \ge 0,$$
$$\frac{\partial f}{\partial x_2}(1, 0) = \mu_1 \ge 0,$$

. That is the two partial derivatives of f at (1,0) must both be ≥ 0. So these two inequalities must hold when f has a minimum at P.

Minimization at Q. At Q only 1 constraint is active, namely $x_2 - 2 = 0$. The gradient of this constraint is $d := \binom{0}{1}$. Theorem 16.1 applies so there will be a positive μ_3 such that

$$\nabla f(1, 2) + \mu_3 d = 0$$

In terms of components, this becomes

$$\frac{\partial f}{\partial x_1}(1, 2) = 0 \quad \text{and} \quad \frac{\partial f}{\partial x_2}(1, 2) = -\mu_3 \le 0,$$

This time the extremality conditions say that 1 equation and 1 inequality hold when f has a minimum at (1,2).

There are other conditions besides those given in Theorem 17.1 which allow one to conclude that if a point $\tilde{x}$ is a local minimizer of f on S, then (16.2) holds. These are called *constraint qualification conditions*. There are, however, points $\tilde{x}$ and C^1-functions g_j such that (16.2) may not hold when $\tilde{x}$ is a local minimizer of f on S.

18. Optimization with Nonlinear Equality Constraints.

In theorem 13.5 the linear Lagrange multiplier rule was stated for multivariate optimization subject to linear equality constraints. We would like to have a similar rule for minimizing a continuously differentiable function f on a set S defined by L equality constraints. Let $\{h_1, \ldots, h_L\}$ be continuously differentiable real valued functions on $\mathbb{R}^n$ and define

$$S \ := \ \{x \in \mathbb{R}^n : h_l(x) \ = \ 0 \quad \text{for } 1 \leq l \leq L\} \tag{18.1}$$

A point $x \in S$ is said to be a *regular point* for these equality constraints provided the set $\{\nabla\, h_l(x) \ : \ 1 \leq l \leq L\}$ is a linearly independent set in $\mathbb{R}^n$.

Suppose $f \ : \ \mathbb{R}^n \ \rightarrow \ \mathbb{R}$ is continuously differentiable and we want to find the *extremality conditions* obeyed by the local minimizers of f on S.

Theorem 18.1 (Lagrange Mulltiplier) Suppose f, S as above and $\tilde{x}$ is a local minimizer of f on S. If $\tilde{x}$ is a regular point for the equality constraints, then there is a unique vector $\lambda \in \mathbb{R}^L$ such that $\tilde{x}$ satisfies

$$\nabla f(x) \ + \ \sum_{l=1}^{L} \ \lambda_l \ \nabla\, h_l(x) \ = \ 0 \tag{18.2}$$

This is a system of $n + L$ equations for $n + L$ unknowns; namely the n entries in $\tilde{x}$ and the L Lagrange multipliers λ_l.

This result will be proved by using a penalty function formulation similar to that for the inequality case described in the preceding section. Given a local minimizer $\tilde{x}$ of f on S and an $\epsilon > 0$, consider the function

$$F_k(x) \ := \ f(x) \ + \ \frac{k}{2} \ \sum_{l=1}^{L} \ h_l(x)^2 \ + \ \frac{\epsilon}{2} \ \|x - \tilde{x}\|_2^2 \tag{18.3}$$

Let $B_1 \ := \ \{x \in \mathbb{R}^n \ : \ \|x - \tilde{x}\|_2 \leq \delta\}$ be a closed ball in $\mathbb{R}^n$ of radius δ and center $\tilde{x}$.

Consider the *penalized problem* of minimizing F_k on B_1. This is a problem of minimizing a continuously differentiable function on a closed ball in $\mathbb{R}^n$; so the constraints

are easy to work with. This problem always has a solution from Weierstrass' theorem. Let $x^{(k)}$ be a minimizer of this problem and $\Gamma := \{x^{(k)} : k \geq 1\}$ be a corresponding sequence of minimizers.

Lemma 18.2 (Convergence) Suppose f, S, F_k, B_1 as above and $\tilde{x}$ is a local minimizer of f on S. If Γ is a sequence of minimizers of F_k on B_1, then $x^{(k)}$ converges to $\tilde{x}$ as $k \to \infty$.

Proof: Note that $F_k(\tilde{x}) = f(\tilde{x})$ for all k, so

$$F_k(x^{(k)}) \leq f(\tilde{x}) \quad \text{for all } k \geq 1. \tag{18.4}$$

Define $\alpha_1 := inf_{x \in B_1} f(x)$. Then the last inequality and the definition of F_k show that

$$\alpha_1 + \frac{k}{2} \sum_{l=1}^{L} h_l(x^{(k)})^2 \leq f(\tilde{x}) \quad \text{for all } k \geq 1$$

Since $k \nearrow \infty$, this implies that

$$\sum_{l=1}^{L} h_l(x^{(k)})^2 \to 0^+ \quad \text{as } k \to \infty.$$

Thus each $h_l(x^{(k)}) \to 0$ as $k \to \infty$. Let $\hat{x}$ be a limit point of the sequence Γ. Then since each h_l is continuous, we have $h_l(x^{(k)}) \to h_l(\hat{x})$. Thus $\hat{x} \in S$. Take limits in (18.4), then

$$f(\hat{x}) + \frac{\epsilon}{2}\|\hat{x} - \tilde{x}\|_2^2 \leq f(\tilde{x}) = \alpha(f, S).$$

This can only happen if $\hat{x} = \tilde{x}$, so the result holds.

Proof of theorem 18.1:
A consequence of the preceding lemma is that, for sufficiently large $k, x^{(k)}$ will be in the interior of B_1. When this holds, then $x^{(k)}$ will satisfy

$$\nabla f(x) + k \sum_{l=1}^{L} h_l(x) \nabla h_l(x) + \epsilon(x - \tilde{x}) = 0 \tag{18.5}$$

If $x^{(k)} \in S$ then we must have $x^{(k)} = \tilde{x}$ and this formula implies that $\nabla f(\tilde{x}) = 0$ which has the form (18.2).

Otherwise, define $\lambda_l^k := k h_l(x^{(k)})$ then (18.5) implies that

$$\nabla f(x^{(k)}) + \sum_{l=1}^{L} \lambda_l^k \nabla h_l(x^{(k)}) = \epsilon (\tilde{x} - x^{(k)}) \tag{18.6}$$

This is a linear equation of the form $D_k \lambda^k = c^k$ where $D_k := DH(x^{(k)}) := [\nabla h_1((x^{(k)}), \ldots, \nabla h_l((x^{(k)})]$ is an $n \times L$ matrix. The matrix $\tilde{D} := DH(\tilde{x})$ has rank L by

assumption. Thus for large k, D_k has rank L as the columns are continuous functions of x on B_1. λ^k is a column vector with L components and $c^k := \epsilon(\tilde{x} - x^{(k)}) - \nabla f(x^{(k)})$. Premultiply this equation by D_k^T, to obtain

$$D_k^T D_k \ \lambda^k \ = \ D_k^T c^k$$

When k is large enough, the matrix on this left hand side is an $L \times L$ matrix of rank L. This equation has a unique solution for the vector λ^k. Let $k \to \infty$, then since the functions are continuously differentiable, D_k converges to $\tilde{D}$ and $c^k \to \tilde{c}$. Thus the λ^k converge to

$$\tilde{\lambda} \ := \ (\tilde{D}^T \tilde{D})^{-1} \tilde{D}^T \tilde{c}$$

Now take limits as $k \to \infty$ in (18.6), each term converges to a limit and the RHS goes to zero, so the Lagrange multiplier rule (18.2) holds at the local minimizer $\tilde{x}$ of f on S and λ is unique.

This theorem is Corollary 2 of theorem 5.2 in Chapter 4 of Berkovitz, page 156. He gives a number of examples of its use.

Note that the constraints could just as well have been that

$$h_l(x) \ = \ c_l \qquad \text{for } 1 \le l \le L. \tag{18.7}$$

Just define some new functions $m_l(x) := h_l(x) - c_l$ and the set S will have the form (18.1). Moreover $\nabla h_l(x) = \nabla m_l(x)$ so the Lagrange multiplier rule (18.2) is the same.

Again there are some other conditions under which a Lagrange multiplier rule holds at a local minimizer of an equality constrained C^1-problem like this one. There also are examples where this rule does not hold. The criterion of this theorem is usually the easiest to verify in practice.

lecture 19; 11/1/04.

19. Optimal Portfolio Problems.

Optimal portfolio problems have the following mathematical formulation.

$$\text{Minimize} \qquad V(x) \ := \ \langle Cx, x \rangle \qquad \text{with } x \in \mathbb{R}^n \tag{19.1}$$

$$\text{subject to} \qquad \sum_{j=1}^{n} x_j \ = \ A \qquad \text{and} \tag{19.2}$$

$$\sum_{j=1}^{n} r_j \ x_j \ = \ R. \tag{19.3}$$

Here x_j is the amount invested in the j-th asset, A is the initial amount invested and R is the expected return - both are assumed to be strictly positive. $C := (c_{jk})$ is

a positive definite symmetric $n \times n$ matrix which is the variance-covariance matrix of the asset prices. This problem says that you wish to find the allocation of the amount available which will produce the given return R while minimizing the variance.

This is a problem with two linear equality constraints, so the theory of section 13 applies. This function $V(x)$ is strictly convex from theorem 7.3 and it is also coercive, so V will attain its infimum on the affine subspace of all vectors which satisfy (19.2)-(19.3).

Let $e := (1, 1, \ldots, 1)^T$ and $r := (r_1, \ldots, r_n)^T$ be the vectors in the constraints. Without loss of generality (wlog), assume that e, r are linearly independent. (If not then each asset has the same rate of return - so they are indistinguishable from the point of view of return and you just put all your resources into the security with least variance). Henceforth we will also assume that all the $r_j > 0$. Why invest in a security where the expected return is negative or zero? Also note that if $\hat{x}$ is the solution of this problem with $A = 1, R = R_1$, then Ax is the solution of this problem subject to (19.2) and $R = AR_1$. So wlog we often take $A = 1$ and then adjust R. This is equivalent to just determining the proportion that should be invested in each security.

There are a number of ways to minimize V subject to these constraints. First we could use the two constraints to eliminate two variables - say x_n, x_{n+1} and then substitute back into $V(x)$. This will be a quadratic function of $n - 2$ variables and one seeks an unconstrained minimizer of this function. This is fine when n is small and you do not change the vector r.

There are many *direct* minimization algorithms that are specially designed to find minima of problems such as this. Here I'll primarily look at the equations and inequalities satisified at the optimal allocation.

From theorem 14.2, $\hat{x}$ will be a minimizer of this problem if and only if there is a $\lambda \in \mathbb{R}^2$ which satisfies

$$Cx \;=\; \lambda_1\, e \;+\; \lambda_2\, r \tag{19.4}$$

Since C is positive definite, it is invertible and its inverse C^{-1} will also be symmmetric and positive definite. Thus if we know the Lagrange multipliers λ_1, λ_2, we only have to solve this system (19.4) to find the optimal $\hat{x}$. Now the solution of this system is given by

$$\hat{x} \;=\; C^{-1}(\lambda_1\, e \;+\; \lambda_2\, r) \tag{19.5}$$

Substitute this back into (19.2)- (19.3), to obtain 2 linear equations for λ. Namely

(19.6)

When e, r are linearly independent this system is non-singular (Why?), and there is a unique solution given by

where L is the inverse of the matrix in (19.6). Substitute this into (19.5), to find the optimal allocation. That is to find the optimal allocation, we need to know the (Cholesky decomposition of the) inverse matrix C^{-1} and solve (19.6) to find the Lagrange multipliers. Then use the formula (19.5).

In practice most investors, or funds, impose further constraints. Many investors are uncomfortable with, or not allowed to, short sell stocks. This means that all the x_j are required to be positive. Then the allowable allocations are

$$S \;:=\; \{x \in \mathbb{R}^n_+ :\; (19.2) \text{ and } (19.3) \text{ hold}\} \tag{19.7}$$

This is a bounded closed convex set, with

$$0 \;\le\; x_j \;\le\; \min\left(A, R/r_j\right) \qquad \text{for each } j.$$

provided the vector $r > 0$. (What happens when $r_j \le 0$?) The optimization problem now is to minimize a strictly convex function on a compact convex set, so there is a unique minimizer of V on S.

The theory of section 15 applies here. To find the conditions obeyed at a local minimizer, we need to describe the normal cone to S at a general point in S. This can be done, as in today's homework problem set. The KKT theorem 15.2 says that the minimizer $\hat{x}$ of V on S satisfies

$$Cx \;+\; z \;=\; 0 \qquad \text{for some } \; z \in N_x(S) \tag{19.8}$$

A somewhat simpler condition may be obtained using Lagrange and KKT multipliers. Namely there is a $\lambda \in \mathbb{R}^2$ such that the minimizer satisfies

$$\sum_{k=1}^{n} c_{jk}\, x_k - \;\lambda_1 \;-\; \lambda_2\, r_j \quad\ge\quad 0 \qquad \text{for all } j \ge 1, \text{ and} \tag{19.9}$$

$$\sum_{k=1}^{n} c_{jk}\, x_k - \;\lambda_1 \;-\; \lambda_2\, r_j \quad=\quad 0 \qquad \text{whenever } x_j > 0. \tag{19.10}$$

If we multiply each equation here by x_j and add, then we find the minimizer $\hat{x}$ satisfies

$$\langle Cx, x\rangle \;=\; \lambda_1 A \;+\; \lambda_2\, R \tag{19.11}$$

These conditions are not usually sufficient to determine the solution $\hat{x}$ and λ directly. We usually have less than $n+2$ equations for the $n+2$ unknowns. Nevertheless solutions of this optimization problem may be found very efficiently using standard algorithms for minimizing convex functions on compact convex sets.

Another common requirement is to say that no single security can be more than a fixed percentage of the portfolio. This enforces diversification. In this case, define $I := [0, b]$ where b is the maximum investment in a given security. Then the allowable portfolios lie in

$$S \;:=\; \{x \in \mathbb{R}^n :\; (19.2) \text{ and } (19.3) \text{ hold and each } x_j \in I\} \tag{19.12}$$

This S is a compact convex set. For it to be non-empty, we must have $nb \geq A$ and $b(\sum_{j=1}^n\; r_j) \;\geq\; R$ - assuming $r \geq 0$. That is

$$b \;>\; \min\left(A/n, R/|r|\right) \qquad \text{where } |r| := \sum_{j=1}^{n}\; r_j$$

In the preceding case this holds with $b = A$. As b decreases, this set S becomes smaller and will be empty when b is too small. As long as $S \neq \emptyset$, there will be a unique minimizer of V on S and the algorithms for minimizing convex functions on compact convex sets will find the minimizer.

This time the conditions satisfied by a local minimizer of V on S will be that there is a $\lambda \in \mathbb{R}^2$ such that the minimizer $\hat{x}$ satisfies

$$\sum_{k=1}^{n} c_{jk}\, x_k - \;\lambda_1\; -\; \lambda_2\, r_j \quad \geq \quad 0 \qquad \text{whenever } x_j = 0, \tag{19.13}$$

$$\sum_{k=1}^{n} c_{jk}\, x_k - \;\lambda_1\; -\; \lambda_2\, r_j \quad = \quad 0 \qquad \text{whenever } 0 < x_j < b, \tag{19.14}$$

$$\sum_{k=1}^{n} c_{jk}\, x_k - \;\lambda_1\; -\; \lambda_2\, r_j \quad \leq \quad 0 \qquad \text{whenever } x_j = b, \tag{19.15}$$

This is a system of n linear inequalities and 2 equations, for the $n+2$ unknowns. Now we do not even have a simple analog of (19.11).

Still for each of these problems, we have
(i) the existence of a unique minimizer,
(ii) the criteria that must hold at a local minimizer of V on the set, and
(iii) good algorithms for finding this minimizer.

In each case the criteria that hold at the minimizer has a different form. It is
(i) a system of $n+2$ linear equations for $n+2$ unknowns when there are no inequality

constraints,
(ii) a system of at least 3 equations and at most $n-1$ inequalities when we require the components to be positive
(iii) a system of at least two equations and up to n inequalities when the components must lie in a fixed interval.

There are other variations on these problems. For example, one could allow the components x_j to lie in different intervals $I_j \subset \mathbb{R}$, or the intervals could have the form $I = [-b_1, b_2]$ where b_1, b_2 are both positive. For each such problem, the theory we have described allows us to conclude that there is a unique minimizer of V on S and to find the extremality conditions that such a minimizer must satisfy.

lecture 20; 11/3/04.

20. Lagrangians and Dual problems.

Associated with a convex optimization problem, there often are *dual optimization problems* which will provide extra information about the original or *primal* problem. Often these are variational problems for the multipliers that are present in the optimality conditions; but sometimes they involve new *dual variables*. There may be a number of very different dual problems to the same original problem.

The primal problem $(\mathcal{P})$ is assumed to have the form

$$\text{Find} \qquad \alpha(f,C) \;:=\; \inf_{x\in C}\; f(x) \quad \text{and} \quad \mathcal{M}(f) \;:=\; \{x \in C : f(x) \;=\; \alpha(f,C)\}. \tag{20.1}$$

with $f : C \to (-\infty, \infty]$ assumed to be continuous and C a nonempty closed convex subset of $\mathbb{R}^n$.

Let Λ be a nonempty, closed subset of $\mathbb{R}^m$, then a function $\mathcal{L} : C \times \Lambda \to \mathbb{R}$ is said to be a *Lagrangian for* $(\mathcal{P})$, provided

$$f(x) \;=\; \sup_{y\in\Lambda}\; \mathcal{L}(x,y) \qquad \text{for all } x \in C. \tag{20.2}$$

There are many examples of such Lagrangians. This notation is standard for optimization theory - but only rarely is related to the Lagrangians that arise in physics. The following are examples of Lagrangians for some of the constrained optimization problems we have already considered.

Example 20.1. Suppose, as in section 13, that $f : \mathbb{R}^n \to \mathbb{R}$ is a given differentiable function. Let A be a $m \times n$ matrix and $b \in \mathbb{R}^m$ be in the range of A. Assume $rank\, A = m < n$ and define C to be the set of all solutions of the equation

$$Ax \;=\; b \tag{20.3}$$

C is an affine subspace of $\mathbb{R}^n$. The primal problem is to minimize f on C. Define $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$\mathcal{L}(x, \lambda) \;=\; f(x) \;-\; \langle \lambda, Ax - b \rangle. \tag{20.4}$$

Then the function $F : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$F(x) \;:=\; \sup_{\lambda \in \mathbb{R}^m} \mathcal{L}(x, \lambda) \;=\; \begin{cases} f(x) & \text{if} \quad Ax = b \\ \infty & \text{otherwise.} \end{cases} \tag{20.5}$$

This shows that this $\mathcal{L}$ is a Lagrangian for our problem.

Example 20.2. Suppose again that $f : \mathbb{R}^n \to \mathbb{R}$ is a given differentiable function. Let A be a $m \times n$ matrix and $b \in \mathbb{R}^m$ be in the range of A. Assume $rank\, A = m \geq 1$ and, as in section 14, let C be the set of all solutions of the inequalities

$$Ax \;\leq\; b \tag{20.6}$$

C is a closed convex set, assume it is non-empty. The primal problem is to minimize f on C. Consider $\mathcal{L} : \mathbb{R}^n \times [0, \infty)^m \to \mathbb{R}$ defined by

$$\mathcal{L}(x, \mu) \;=\; f(x) \;+\; \langle \mu, Ax - b \rangle \;=\; f(x) \;+\; \sum_{j=1}^{m} \mu_j \left(\langle a^{(j)}, x \rangle - b_j \right). \tag{20.7}$$

Define the function $F : \mathbb{R}^n \to (-\infty, \infty]$ by

$$F(x) \;:=\; \sup_{\mu \in [0,\infty)^m} \mathcal{L}(x, \mu) \;=\; \begin{cases} f(x) & \text{if} \quad Ax \;\leq\; b \\ \infty & \text{otherwise.} \end{cases} \tag{20.8}$$

Here F may be regarded as an extension of f to $\mathbb{R}^n$ as it equals f on C. Thus $\mathcal{L}$ is a Lagrangian for this problem.

Example 20.3. Suppose now that $f, g : \mathbb{R}^n \to \mathbb{R}$ are given differentiable functions. Assume g is convex and let $C \;:=\; \{x \in \mathbb{R}^n : g(x) \;\leq\; c\}$ be non-empty. Then C is closed and convex. The primal problem is to minimize f on C and is the type of problem described in section 16.

Consider $\mathcal{L} : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}$ defined by

$$\mathcal{L}(x, \mu) \;=\; f(x) \;+\; \mu \left(g(x) - c \right). \tag{20.9}$$

Then the function $F : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$F(x) \;:=\; \sup_{\mu \in [0,\infty)} \mathcal{L}(x, \mu) \;=\; \begin{cases} f(x) & \text{if} \quad g(x) \;\leq\; c \\ \infty & \text{otherwise.} \end{cases} \tag{20.10}$$

so F is an extension of f, and this $\mathcal{L}$ is a Lagrangian for this problem.

In terms of the Lagrangian one sees that

$$\alpha(f, C) \;:=\; \inf_{x \in C} \sup_{y \in \Lambda} \; \mathcal{L}(x, y) \tag{20.11}$$

When $\mathcal{L}$ is a Lagrangian as in (20.2), then the dual problem $(\mathcal{P}^*)$ associated with this Lagrangian is to maximize $G : \Lambda \to [-\infty, \infty)$ defined by

$$G(y) \;:=\; \inf_{x \in C} \; \mathcal{L}(x, y). \tag{20.12}$$

The value of this problem will be

$$\beta(G, \Lambda) \;:=\; \sup_{y \in \Lambda} \; G(y) \;=\; \sup_{y \in \Lambda} \inf_{x \in C} \; \mathcal{L}(x, y). \tag{20.13}$$

The following result shows that whenever we have a Lagrangian function as above then the value of the dual problem is less than that of the primal problem.

Theorem 20.1 Suppose $\mathcal{L}$ is a Lagrangian function for the problem $(\mathcal{P})$ and G is the dual functional defined by (20.12). Then

$$G(y) \;\leq\; \beta(G, \Lambda) \;\leq\; \alpha(f, C) \;\leq\; f(x) \quad \text{for all } y \in \Lambda,\; x \in C. \tag{20.14}$$

Proof: From (19.2) and (20.12),

$$f(\tilde{x}) - G(\tilde{y}) \;=\; \sup_{y \in \Lambda} \; \mathcal{L}(\tilde{x}, y) \;-\; \inf_{x \in C} \; \mathcal{L}(x, \tilde{y}) \quad \text{for any } (\tilde{x}, \tilde{y}) \in C \times \Lambda.$$

This RHS equals

$$\sup_{y \in \Lambda} \sup_{x \in C} \quad [\mathcal{L}(\tilde{x}, y) - \mathcal{L}(x, \tilde{y})] \quad \text{and this is} \quad \geq \; 0.$$

Hence $f(\tilde{x}) \;\geq\; G(\tilde{y})$ for any $(\tilde{x}, \tilde{y}) \in C \times \Lambda$ and the result follows.

The central inequality here can be written as follows. It says that for any function $\mathcal{L}$ defined on a product set $C \times \Lambda$, the sup inf is less than, or equal to the sup inf of the function.

Corollary 20.2 Suppose $\mathcal{L}$ is any function on $C \times \Lambda$, then

$$\sup_{y \in \Lambda} \inf_{x \in C} \; \mathcal{L}(x, y) \;\leq\; \inf_{x \in C} \sup_{y \in \Lambda} \; \mathcal{L}(x, y). \quad \text{for all } y \in \Lambda,\; x \in C. \tag{20.15}$$

21. Conjugate Convex Functions.

To describe the dual problems it helps to use the theory of conjugate, (or dual or polar), convex functions. Their systematic use helps simplify many arguments and formulae. Unfortunately they are not treated in any of the recommended textbooks.

Let C be a convex subset of $\mathbb{R}^n$ and $f : C \to (-\infty, \infty]$ be a given function. The *effective domain* of f is defined to be the set

$$dom(f) \ := \ \{x \in C : \ f(x) < \infty\} \tag{21.1}$$

The function f is said to be nontrivial on C provided $dom(f) \neq \emptyset$.

The *conjugate convex* of f is the function $f^* : \mathbb{R}^n \to (-\infty, \infty]$ defined by

$$f^*(y) \ := \ \sup_{x \in C} [\ \langle y, x \rangle - f(x) \ ]. \tag{21.2}$$

If you prefer to use minimization this is equivalent to

$$-f^*(y) \ = \ \inf_{x \in C} [\ f(x) - \langle y, x \rangle \ ]. \tag{21.3}$$

Note that f^* is defined on all of $\mathbb{R}^n$ whenever C is a subset of $\mathbb{R}^n$. Moreover we don't require that f is continuous, convex or have any property except that it is not identically ∞. Observe that

$$f^*(0) \ = \ -\inf_{x \in C} \ f(x) \ = \ -\alpha(f, C).$$

This conjugate function is a simpler (and better) way of describing what used to be called the *Legendre transform* of a function. This has been that has been used in classical mechanics for 200 years and is the basis of the transition from Lagrangian mechanics to Hamiltonian mechanics. It also is a fundamental operation in thermodynamics. The following theorem says that f^* is always convex - whether or not f is.

Theorem 21.1 Suppose f is a function defined on a convex subset C of $\mathbb{R}^n$ with f bounded below and not identically ∞. Then f^* defined by (21.2) is nontrivial and convex on $\mathbb{R}^n$.

Proof: When f is bounded below on $\mathbb{R}^n$, then $0 \in dom(f^*)$ as above. It is convex as f^* is the supremum of a family of linear - hence convex - functions of y on $\mathbb{R}^n$.

Lecture 21; 11/8/2004

Example 21.1. Let A be a positive definite and symmetric $n \times n$ matrix and define

$$f(x) \ := \ (1/2)\langle Ax, x \rangle.$$

Then $f(x)$ is strictly convex and coercive on $\mathbb{R}^n$ and the infimum in equation (21.3) with this function occurs at a point $\hat{x}$ which satisfies $Ax \;=\; y$. Thus $\hat{x} = A^{-1}y$ since a positive definite matrix always has an inverse. Substitute back in (21.3) to find that

$$f^*(y) \;=\; (1/2)\langle A^{-1}y, y\rangle.$$

In particular when $A := I_n$, then $f(x) := (1/2)\|x\|_2^2$, and $f^*(y) \;=\; (1/2)\|y\|_2^2$.

Example 21.2. Let $f(x) \;:=\; (1/p)\;\|x\|_p^p \quad$ with $1 < p < \infty$.
Then $f(x)$ is strictly convex and coercive on $\mathbb{R}^n$ and the infimum in equation (21.3) occurs at a point $\hat{x}$ which you (should have) found in one of your homework problems (when n = 1) and is described in theorem 12.1. Then, from the fact that equality holds in Young's inequality,

$$f^*(y) \;=\; (1/p^*)\;\|x\|_{p^*}^{p^*}.$$

Example 21.3. An important convex function in probability theory, thermodynamics and statistical mechanics is the *entropy function* $h : [0,\infty)^n \to\; [0,\infty)$ defined by

$$h(x) \;:=\; \sum_{j=1}^{n}\; f(x_j) \quad \text{where} \quad f(s) := s\,\ln s \tag{21.4}$$

This function f was example 6.4 earlier. It is straightforward to show that $f^*(y) \;=\; e^{y-1}$ for $y \in \mathbb{R}$, so that

$$h^*(y) \;=\; e^{-1}\sum_{j=1}^{n}\; e^{y_j} \qquad \text{for } y \in \mathbb{R}^n \tag{21.5}$$

Example 21.4. Let C be a nonempty closed convex set in $\mathbb{R}^n$. The *indicator function of* C is $I_C : \mathbb{R}^n \to\; [0,\infty]$ defined by

$$I_C(x) \;:=\; \begin{cases} 0 & \text{if} \quad x \in C \\ \infty & \text{otherwise.} \end{cases} \tag{21.6}$$

The conjugate function of I_C is called the *support function* of C and is given by

$$s_C(y) \;:=\; \sup_{x\in C}\,\langle x, y\rangle. \tag{21.7}$$

When C is bounded, this function is finite for any $y \in \mathbb{R}^n$.

Example 21.5. Given $c \in \mathbb{R}^n$, define $f(x) \;:=\; \langle c, x\rangle \;+\; c_0$. This is an affine function and its conjugate is

$$f^*(y) \;=\; \begin{cases} -c_0 & \text{if} \quad y = c \\ \infty & \text{otherwise.} \end{cases} \tag{21.8}$$

When $c_0 = 0$, this conjugate function is the indicator function of the single point $c \in \mathbb{R}^n$.

Some simple properties of convex conjugation include the following. Here each f is a given function with values in $(-\infty, \infty]$.

1. If $f(0) = c_0$, then $f^*(y) \geq -c_0$ for every $y \in \mathbb{R}^n$.

2. If $g(x) := cf(x)$ with $c \in (0, \infty)$, then $g^*(y) = c\, f^*(y/c)$

3. If $g(x) := f(Lx)$ with L an invertible $n \times n$ matrix, then $g^*(y) = f^*((L^T)^{-1}y)$.

4. If $f(x) \leq g(x)$ for all $x \in C$, then $f^*(y) \geq g^*(y)$ for all $y \in \mathbb{R}^n$.

5. If $f(x^{(1)}, x^{(2)}) := f_1(x^{(1)}) + f_2(x^{(2)})$ is defined for $(x^{(1)}, x^{(2)}) \in C_1 \times C_2$ where C_1, C_2 are closed convex sets then $f^*(y^{(1)}, y^{(2)}) = {f_1}^*(y^{(1)}) + {f_2}^*(y^{(2)})$.

Homework set 5. (due 11/15/04)
Question 1. Prove the above 5 properties.
Question 2. Prove the formulae for h^* in example 21.3.
Question 3. Suppose that K is a closed convex cone in $\mathbb{R}^n$ and I_K is the indicator function of K as in example 21.4.
(a) Show that the support function of K is again an indicator function and describe the effective domain of s_K.
(b) If $K := [0, \infty)^n$, what is the effective domain of s_K?

The definition (21.2) leads to the following result.
Theorem 21.2 (Fenchel-Young inequality) Suppose f is a function defined on a convex subset C of $\mathbb{R}^n$ and f^* is defined by (21.2). Then

$$f(x) + f^*(y) \geq \langle x, y \rangle \qquad \text{for all } (x, y) \in C \times \mathbb{R}^n. \tag{21.9}$$

If f is convex and differentiable on C, equality holds here if and only if (x, y) satisfy $\nabla f(x) = y$.
Proof: (21.9) follows directly from the definition (21.2). When f is convex and differentiable on C, and the infimum in (21.3) is attained then the minimizer is a point which satisfies $\nabla f(x) = y$. Hence equality holds in (21.9) implies that $\nabla f(x) = y$. Conversely if $\nabla f(x) = y$, then such an x will minimize the RHS of (21.3) and this implies equality in (21.9).

Note that the Youngs' inequality of section 12 is the special case of this inequality with f as in example 21.2. Essentially this theorem generalizes Young's equality to the class of all functions that have a nontrivial conjugate function.

Conjugate convex functions often arise in describing dual optimization problems. For the Lagrangians associated with either equality or inequality constraints, we have the following dual problems.

Example 21.1. (continued) When $\mathcal{L} : C \times \Lambda \to \mathbb{R}$ is defined by equation (20.2), then the dual problem $(\mathcal{P}^*)$ from (20.12) is to maximize

$$G(\lambda) := \langle b, \lambda \rangle - f^*(A^T \lambda) \quad \text{for } \lambda \in \mathbb{R}^m. \tag{21.10}$$

This is an unconstrained m-dimensional optimization problem - which is often written as the problem of minimizing $-G(\lambda)$ on $\mathbb{R}^n$.

Example 21.2. (continued) When $\mathcal{L} : C \times [0,\infty)^m \to \mathbb{R}$ is defined by equation (20.7), then the dual problem $(\mathcal{P}^*)$ from (20.12) is to maximize

$$G(\mu) \;:=\; -\langle b, \mu\rangle - \; f^*(-A^T\,\mu) \quad \text{for } \mu \in [0,\infty)^m. \tag{21.11}$$

This is an maximization problem on the positive orthant in $\mathbb{R}^m$. Again it is often treated as a problem of minimizing $-G(\mu)$ on this positive orthant.

In both these examples, the dual problem involves maximizing a concave function on a relatively simple closed convex set - either all of $\mathbb{R}^m$ or the positive orthant $[0,\infty)^m$.

Lecture 22; 11/10/2004

22. Saddle Points and Dual Optimization Problems.

Suppose f, C are as in section 19 and $(\mathcal{P})$ is the primal problem of minimizing f on C. Let C, D be nonempty closed convex subsets of $\mathbb{R}^n, \mathbb{R}^m$ respectively and $\mathcal{L} : C \times D \to \mathbb{R}$ be a *Lagrangian for* $(\mathcal{P})$. A point $(\hat{x}, \hat{y})$ is said to be a *saddle point* for $\mathcal{L}$ on $C \times D$ provided

$$\mathcal{L}(\hat{x}, y) \;\le\; \mathcal{L}(\hat{x}, \hat{y}) \;\le\; \mathcal{L}(x, \hat{y}) \quad \text{for all } (x, y) \in C \times D. \tag{22.1}$$

This is a purely algebraic definition and is not necessarily the same as the definition you may have had in a calculus course. The dual problem $(\mathcal{P}^*)$ is to maximize $G : D \to \overline{\mathbb{R}}$ defined by

$$G(y) \;:=\; \inf_{x\in C}\; \mathcal{L}(x, y). \tag{22.2}$$

The following theorem says that if a Lagrangian has a saddle point then there are solutions of both the primal and dual optimization problem and their values are equal. Moreover, if the primal and dual problems both have solutions, and their values are equal, then the Lagrangian has a saddle point. It is the central theorem of (convex) duality theory.

Theorem 22.1 (Saddle-point duality) Suppose $\mathcal{L}$ is a Lagrangian for the problem $(\mathcal{P})$ and G is defined by (22.2). Then $(\hat{x}, \hat{y})$ is a saddle point for $\mathcal{L}$ on $C \times D$ if and only if

1. $\hat{x}$ is a minimizer of $(\mathcal{P})$,
2. $\hat{y}$ is a maximizer of $(\mathcal{P}^*)$, and
3. $\alpha(f, C) \;=\; \beta(G, D)$.

Proof: Choose $\tilde{x}$ in C and $\tilde{y}$ in D. Then the properties of $\mathcal{L}$ imply that

$$f(\tilde{x}) - G(\tilde{y}) \;=\; \sup_{y\in D}\; \mathcal{L}(\tilde{x}, y) - \inf_{x\in C}\; \mathcal{L}(x, \tilde{y})$$

$$= \sup_{y\in D} \sup_{x\in C} \mathcal{L}(\tilde{x}, y) - \mathcal{L}(x, \tilde{y}) \geq 0.$$

From the definition of a saddle point (22.1), we have

$$f(\hat{x}) \leq \mathcal{L}(\hat{x}, \hat{y}) \leq G(\hat{y})$$

These last two sets of inequalities then imply that

$$f(\hat{x}) = \mathcal{L}(\hat{x}, \hat{y}) = G(\hat{y}) \tag{22.3}$$

so properties 1-3 of the theorem hold.

Conversely if $(\hat{x}, \hat{y})$ satisfy properties 1-3, then

$$\sup_{y\in D} \mathcal{L}(\hat{x}, y) = \inf_{x\in C} \mathcal{L}(x, \hat{y}) \leq \mathcal{L}(\hat{x}, \hat{y}).$$

Thus

$$\mathcal{L}(\hat{x}, y) \leq \mathcal{L}(\hat{x}, \hat{y}) \quad \text{for all } y \in D.$$

Similarly property 3 implies that

$$G(\hat{y}) = \inf_{x\in C} \mathcal{L}(x, \hat{y}) = f(\hat{x}) = \sup_{y\in D} \mathcal{L}(\hat{x}, y) \geq \mathcal{L}(\hat{x}, \hat{y})$$

Thus $\mathcal{L}(x, \hat{y}) \geq \mathcal{L}(\hat{x}, \hat{y})$ for all $x \in C$, so the other part of (22.1) holds. That is, $(\hat{x}, \hat{y})$ is a saddle point for $\mathcal{L}$ on $C \times D$.

Thus a pair of dual optimization problems will both have solutions whenever the associated Lagrangian function has a saddle point. Usually to prove that a function $\mathcal{L}$ has a saddle point in the sense of (22.1), we need some convexity type conditions on $\mathcal{L}$.

The function $\mathcal{L}$ is said to be *convex-concave* on $C \times D$ provided C and D are convex sets and
(i) $\mathcal{L}(., y)$ is convex on C for each $y \in D$, and
(ii) $\mathcal{L}(x, .)$ is concave on D for each $x \in C$.

Remember that a function G is concave on a convex set D whenever $-G$ is convex on D. Note that when $\mathcal{L}$ is convex-concave on $C \times D$, then from the property 2 of operations on convex functions in lecture 7, we have
(i) $F(x) := \sup_{y\in D} \mathcal{L}(x, y)$ is convex on C, and
(ii) $G(y) := \inf_{x\in C} \mathcal{L}(x, y)$ is concave on D.

Example 22.1 Suppose $f : \mathbb{R}^n \to R$ and $g : \mathbb{R}^m \to \mathbb{R}$ are continuous convex functions and B is an $m \times n$ matrix. Define $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$\mathcal{L}(x, y) := f(x) + \langle Bx, y\rangle - g(y).$$

Then $\mathcal{L}$ will be convex-concave. Moreover the functions $F : \mathbb{R}^n \to \mathbb{R}$ and $G : \mathbb{R}^m \to \mathbb{R}$ are given by

$$F(x) \;=\; f(x) \;+\; g^*(Bx) \qquad \text{and} \quad G(y) \;=\; -g(y) - f^*(-B^T y).$$

Thus if a Lagrangian is convex-concave, both the primal and the dual problems are equivalent to minimizing a convex function on a convex set. This holds as maximizing the concave function G on D is equivalent to minimizing $-G$ - which is convex - on D.

The basic theorem on saddle points is the following result which was originally proved as a result in the theory of 2-person games.

Theorem 22.2 (von Neumann's minimax) Suppose $\mathcal{L}$ is a convex-concave function, C, D are closed nonempty compact convex sets in $\mathbb{R}^n, \mathbb{R}^m$ respectively and $\mathcal{L}$ is continuous. Then there is at least one saddle point of $\mathcal{L}$ on $C \times D$.

This is theorem 5.2, chapter 3 of Berkovitz, page 117ff and he takes 4 pages for the proof. In fact we have the following result when C, D are not compact - which is the case in the examples we've had

Corollary 22.3 Suppose $\mathcal{L}$ is a continuous convex-concave function and C,D are nonempty, closed unbounded convex sets in $\mathbb{R}^n, \mathbb{R}^m$ respectively. Suppose also
(i) there is a $y_0 \in D$, such that

$$\lim_{||x|| \to \infty, x \in C} \mathcal{L}(x, y_0) \;=\; \infty, \quad \text{and}$$

(ii) there is an $x_0 \in C$, such that

$$\lim_{||y|| \to \infty, y \in D} \mathcal{L}(x_0, y) \;=\; -\infty,$$

then there is at least one saddle point of $\mathcal{L}$ on $C \times D$.

The proof of this is very similar to that of the theorem. There are some other general theorems on the existence of saddle points for various types of functions; the above are just the most often used results.

In either case we have the following which depends on the fact that the set of minimizers of a convex function on a convex set will itself be a convex set.

Theorem 22.4 Suppose $\mathcal{L}$ is a continuous convex-concave function and C,D are nonempty, closed convex sets in $\mathbb{R}^n, \mathbb{R}^m$ respectively. Then the set of saddle points of $\mathcal{L}$ on $C \times D$ is a closed convex set in $C \times D$.

Proof: From theorem 21.1, a point $(\tilde{x}, \tilde{y})$ in $C \times D$ will be a saddle point of $\mathcal{L}$ if and only if $\tilde{x}$ minimizes F on C and $\tilde{y}$ maximizes G on D. The function F is convex so from Corollary 9.4, the set of all minimizers of F on C is a closed convex set C_m in $\mathbb{R}^n$. Similarly the set of all maximizers of G on D is a closed convex set D_m in $\mathbb{R}^m$. Thus the set of all

saddle points of $\mathcal{L}$ will be $C_m \times D_m$ and is a closed convex set since the product of two closed convex sets is again a closed convex set.

A natural question is what differential conditions hold at the saddle points of a differentiable function $\mathcal{L}$? The following is the necessary condition.

Theorem 22.5 Suppose C, D are nonempty, open, convex sets in $\mathbb{R}^n, \mathbb{R}^m$ respectively and $\mathcal{L}$ is a continuous function. If $(\hat{x}, \hat{y})$ is a saddle point of $\mathcal{L}$ and $\mathcal{L}$ is G-differentable at $(\hat{x}, \hat{y})$, then

$$\nabla_x \mathcal{L}(\hat{x}, \hat{y}) \;=\; 0 \quad \text{and} \quad \nabla_y \mathcal{L}(\hat{x}, \hat{y}) \;=\; 0 \tag{22.4}$$

Proof: Just like the proof of theorem 8.1.

We also have the following analog of theorem 9.3.

Theorem 22.6 Suppose C, D are nonempty, open, convex sets in $\mathbb{R}^n, \mathbb{R}^m$ respectively and $\mathcal{L}$ is a differentiable convex-concave function. The point $(\hat{x}, \hat{y})$ is a saddle point of $\mathcal{L}$ on $C \times D$ if and only if (22.4) holds.

Example 22.1 Let A be a real symmetric, positive definite $n \times n$ matrix and B be an $n \times n$ skew-symmetric matrix. That is $B^T = -B$. Consider the function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by

$$\mathcal{L}(x, y) \;:=\; \frac{1}{2}\langle Ax, x\rangle \;+\; \langle f, y - x\rangle \;-\; \langle Bx, y\rangle \;-\; \frac{1}{2}\langle Ay, y\rangle \tag{22.5}$$

This function is convex-concave and it has the derivatives

$$\nabla_x \mathcal{L}(x, y) \;=\; Ax \;-\; f \;-\; B^T y \quad \text{and} \quad \nabla_y \mathcal{L}(x, y) \;=\; f \;-\; Bx \;-\; Ay. \tag{22.6}$$

Theorem 22.7 Suppose $A, B, \mathcal{L}$ as above, then there is a unique saddle point $(\hat{x}, \hat{x})$ of $\mathcal{L}$ on $\mathbb{R}^n \times \mathbb{R}^n$. Moreover $\hat{x}$ is the unique solution of the linear equation

$$(A \;+\; B)\, x \;=\; f \tag{22.7}$$

Proof: This Lagrangian satisfies the conditions of Corollary 21.3 so it has at least one saddle point. Use the formulae for the derivatives (22.6) and theorem 21.6, then any such saddle point $(\hat{x}, \hat{y})$ satisfies

$$Ax \;-\; B^T y \;=\; f \quad \text{and} \quad Bx \;+\; Ay = f.$$

Subtract these two equations, then

$$(A \;-\; B)\,(\hat{x} \;-\; \hat{y}) \;=\; 0$$

Take inner products with $\hat{x} \;-\; \hat{y}$. Then

$$\langle A(\hat{x} \;-\; \hat{y}), \hat{x} \;-\; \hat{y}\rangle \;=\; 0$$

This implies $\hat{x} = \hat{y}$ as A is positive definite. Thus the theorem follows.

The primal and dual problems associated with this Lagrangian provided optimization problems for the solution of (22.7).

Note that any matrix equation of the form $Cx = f$ with C being an $n \times n$ matrix and $\langle Cx, x\rangle > 0$ for $x \in \mathbb{R}^n \setminus \{0\}$ can be written in the form (22.7). Thus there are both optimization, and saddle point characterizations of solutions of linear equations of this form.

Lecture 23; 11/15/2004

23. Duality for Quadratic Programming problems.

Here we shall look at the dual problems associated with a programming problem of the form $(\mathcal{P})$; minimize

$$f(x) := (1/2)\langle Ax, x\rangle - \langle c, x\rangle \quad \text{subject to} \quad Bx \leq d \tag{23.1}$$

Here A is a positive definite symmetric $n \times n$ matrix, B is an $m \times n$ matrix, c, x are in $\mathbb{R}^n$ and $d \in \mathbb{R}^m$. This is the problem treated in Chapter 5, section 6 of Berkovitz. This is called a *quadratic programming* problem as it involves minimizing a convex quadratic function subject to linear constraints. We shall further assume that the set

$$C := \{x \in \mathbb{R}^n : Bx \leq d\}. \tag{23.2}$$

contains more than 1 point. It is a closed convex subset of $\mathbb{R}^n$. Note that A is positive definite implies that f is strictly convex and coercive on $\mathbb{R}^n$, so this optimization problem has a unique minimizer whenever C is non-empty.

All of the portfolio optimization problems of section 18 can be put into this form as linear equality constraints correspond to 2 linear inequality constraints. That is

$$\langle b, x\rangle = d_0 \quad \text{is the same as} \quad \langle b, x\rangle \leq d_0 \text{ and } \langle b, x\rangle \geq d_0.$$

The Lagrangian for this problem is described in example 19.2 and is $\mathcal{L} : \mathbb{R}^n \times [0, \infty)^m \to \mathbb{R}$ defined by

$$\mathcal{L}(x, \mu) = (1/2)\langle Ax, x\rangle - \langle c, x\rangle + \langle \mu, Bx - d\rangle \tag{23.3}$$

The primal problem associated with this Lagrangian is the original problem $(\mathcal{P})$. For fixed μ, this function $\mathcal{L}(x, \mu)$ is strictly convex and coercive in x, so it has a unique minimizer $\hat{x}(\mu)$ which will be the unique solution of

$$Ax = c - B^T\mu \tag{23.4}$$

That is

$$\hat{x}(\mu) \;=\; A^{-1}(c \;-\; B^T\mu)$$

This dual function defined by (19.11) is $G : [0,\infty)^m \to \mathbb{R}$ where

$$G(\mu) \;:=\; (-1/2)\,\langle A^{-1}(B^T\mu - c), (B^T\mu - c)\rangle \;-\; \langle \mu, d\rangle \tag{23.5}$$

The dual problem $(\mathcal{P}^*)$ is to maximize G on $[0,\infty)^m$. This is a problem of maximizing a strictly concave function on the positive orthant of $\mathbb{R}^m$. It has a unique maximizer $\hat{\mu}$ - which will be the vector of KKT multipliers for this problem.

Then the unique solution of $(\mathcal{P})$ will be given by the unique solution $\hat{x}$ of (23.4) with $\hat{\mu}$ in place of μ.

Recapitulation: Suppose the primal problem $(\mathcal{P})$ is to minimize the coercive, strictly convex quadratic function f on C. From the existence theorems, there is a unique solution and one could try to find this minimizer directly using a standard algorithm. If this is done then Theorem 16.1 - the second KKT theorem, says that the minimizer will be a solution of (23.4) for some KKT multipliers $\hat{\mu} \in [0,\infty)^m$ with the property that

$$\langle \hat{\mu}, B\hat{x} - d\rangle \;=\; 0$$

(Actually we may have to improve theorem 16.1 a little when there are some equality constraints in the definition of C but the result still holds for this problem). That is, KKT type theorem says that there will be some $\hat{\mu}$ such that the last two equations hold.

This Lagrangian formulation provides an optimization problem for the KKT multipliers. This dual problem is a quadratic optimization problem on the positive orthant in $\mathbb{R}^m$ which has a unique solution for $\hat{\mu}$ and the corresponding maximum value is equal to the value of the primal problem. Then the minimizer $\hat{x}$ is found by solving a positive definite symmetric linear equation. This is a standard linear algebra problem - which is usually resolved using the Cholesky factorization. Thus the Lagrangian formulation provides two associated optimation problems; the primal problem and one for the KKT multipliers.

In fact one can show that $(\hat{x}, \hat{\mu})$ is a saddle point of this Lagrangian. The Lagrangian approach provides extra information about the problem - as it provides a direct description of the KKT multipliers. When m is small compared to n, it is usually better to use the dual problem obtained from this Lagrangian approach. When m is large, and n is small then a direct method for the primal problem may be preferred.

A different description of this duality is given in Chapter 5, section 6 of Berkovitz (page 210).

24. Linear Programming problems.

A linear programming problem is a problem $(\mathcal{P})$ of the form; maximize, (or minimize) the function

$$f(x) \ := \ \langle c, x \rangle \quad \text{subject to} \quad x \ \geq \ 0 \quad \& \quad Bx \ \leq \ d. \tag{24.1}$$

Here B is an $m \times n$ matrix, $c \in \mathbb{R}^n$ and we shall assume that the set

$$C \ := \ \{x \in \mathbb{R}^n \ : x \ \geq \ 0 \ \& \ Bx \ \leq \ d\} \tag{24.2}$$

contains at least two points. This C is a closed convex set in $\mathbb{R}^n$.

When C is bounded, it will be the convex hull of a finite number of points in $[0, \infty)^n$. These points are called the vertices of C and a careful argument leads to the following theorem.

Theorem 24.1 When C as above is bounded, the linear function $f(x)$ attains its minimum, and maximum, at vertices of C.

This result is geometrically obvious in 2 dimensions. Thus in linear programming, the effort was generally devoted to finding the vertices of the domain (or constraint or feasible set) C. The simplex method is an algorithm for generating nearby vertices from a given vertex. It is described in chapter VI of Berkovitz and will not be detailed here. Here we shall describe a duality theory for linear programming problems.

A Lagrangian for this LP problem is similar to those described before and is $\mathcal{L} : [0, \infty)^n \times [0, \infty)^m \to \mathbb{R}$ defined by

$$\mathcal{L}(x, \mu) \ = \ -\langle c, x \rangle \ + \ \langle \mu, Bx - d \rangle. \tag{24.3}$$

This can be rewritten as

$$\mathcal{L}(x, \mu) \ = \ \langle B^T \mu - c, x \rangle \ - \ \langle \mu, d \rangle.$$

One can verify that maximizing $\mathcal{L}$ over μ gives a function whose minimization leads to the primal problem $(\mathcal{P})$.

The dual function associated with this Lagrangian is $G(\mu)$ on $[0, \infty)^m$ where

$$G(\mu) \ := \ \begin{cases} -\langle \mu, d \rangle & \text{if} \quad B^T \mu \ \geq \ c \\ -\infty & \text{otherwise.} \end{cases} \tag{24.4}$$

The dual problem is to maximize G on $[0, \infty)^m$. Since we want this maximum to be finite, we consider the domain of the dual problem to be the set of points where $G(\mu) > -\infty$. Then the dual problem $(\mathcal{P}^*)$ is to maximize $g(\mu) := -\langle d, \mu \rangle$ on the set

$$D \ := \ \{\mu \in \mathbb{R}^m : \ \mu \geq 0 \ \& \ B^T \mu \ \geq \ c\} \tag{24.5}$$

This dual problem is again a LP problem involving the maximization of a linear function on a closed convex set of the same form as C. The vectors c and d have been interchanged, B^T replaces B and the inequality sign has been reversed. The main result about this is the following.

Theorem 24.2 Assume f, g, C, D as above. The problem $(\mathcal{P})$ has a solution $\hat{x}$ if and only if the dual problem $(\mathcal{P}^*)$ has a solution and then the values of the primal and dual problems are equal and the Lagrangian has a saddle point. If $(\mathcal{P})$ has no solution, then D is empty and $\mathcal{L}$ does not have a saddle point.

The proof of this is not hard - but will be omitted here. This is a version of theorem 7.1, chapter 5 of Berkovitz - where he also considers the case where C is empty. Rather than prove these results we'll treat the the following two examples which illustrate how some LP problems work.

Example 24.1 Consider the problem of maximizing $f(x) := \langle c, x\rangle$ over the unit simplex $\Delta_n \subset \mathbb{R}^n$ defined by equation 12.7. That is,

$$\sum_{j=1}^{n} x_j \;\leq\; 1 \quad \text{and} \quad x_j \;\geq\; 0 \quad \text{for } 1 \leq j \leq n. \tag{24.6}$$

This set Δ_n is the closed convex hull of the set $\{0, e^{(1)}, \ldots, e^{(n)}\}$. It has the form of (24.2) with $B := e$ being a $1 \times n$ matrix and $d = 1$.

This description of the vertices of Δ_n and theorem 23.1, implies, that f satisfies

$$\min\{0, c_1, \ldots, c_n\} \;\leq\; f(x) \;\leq\; \max\{0, c_1, \ldots, c_n\} \quad \text{for all} \quad x \in \Delta_n. \tag{24.7}$$

Please convince yourself that this is true. Asssume that $c_J = \min\{c_1, \ldots, c_n\}$ and $c_K = \max\{c_1, \ldots, c_n\}$. If $c_J \leq 0$ then $e^{(J)}$ is a minimizer of f on Δ_n and when $c_J \geq 0$ then 0 is a minimizer . Similarly $e^{(K)}$ is a maximizer of f on Δ_n when $c_K \geq 0$, and 0 is a maximizer when $c_K \leq 0$.

A Lagrangian for this problem is the function $\mathcal{L} : [0, \infty)^n \times [0, \infty) \to \mathbb{R}$ defined by

$$\mathcal{L}(x, \mu) \;:=\; -\langle c, x\rangle \;+\; \mu(\langle e, x\rangle \;-\; 1). \tag{24.8}$$

This can be rewritten as

$$\mathcal{L}(x, \mu) \;=\; \langle \mu e - c, x\rangle \;-\; \mu.$$

You should verify that this is, in fact a Lagrangian for this problem. From equation 19.11, the dual function will be $G : [0, \infty) \to \overline{\mathbb{R}}$ defined by

$$G(\mu) \;:=\; \inf_{x \,\geq\, 0} \mathcal{L}(x, \mu) \;=\; \begin{cases} -\,\mu & \text{if} \quad \mu e \;\geq\; c \\ -\,\infty & \text{if} \quad \mu e < c. \end{cases} \tag{24.9}$$

Now the inequality $\mu e \ \geq \ c$ holds if and only if $\mu \ \geq \ c_K := \max\{c_1, \ldots, c_n\}$. When $c_K \leq 0$ then $G(\mu) = -\mu$ for all $\mu \geq 0$. In this case, $G(\mu) \leq G(0) = 0$ for all $\mu \geq 0$ so the value of the dual problem is 0 and it is attained at $\hat{\mu} = 0$.

When $c_K > 0$, then

$$G(\mu) \ = \ \begin{cases} -\mu & \text{if} \quad \mu \ \geq \ c_K \\ -\infty & \text{if} \quad 0 \leq \mu < c_K. \end{cases}$$

This function is maximized when $\mu = c_K$ and the value of this dual problem will then be $-c_K$.

Now use theorem 24.2, to infer that the value of the primal problem is either 0 when $c_K \leq 0$, or $-c_K$ when $c_K \geq 0$. This implies that the maximum of $\langle c, x \rangle$ on Δ_n will be either 0 or c_K as was earlier found in (24.7). Moreover this maximum is attained at $e^{(K)}$ whenever $c_K \geq 0$ or at 0 otherwise.

If you work carefully with the Lagrangian (24.8), you can show that (0,0) will be a saddle point when $c_K \leq 0$, and that $(e^{(K)}, c_K)$ is a saddle point when $c_K \geq 0$.

A similar analysis can be used to investigate the minimization of $\langle c, x \rangle$ on Δ_n. In place of (24.8), one should use the Lagrangian with c in place of $-c$, so

$$\mathcal{L}(x, \mu) \ := \ \langle c, x \rangle \ + \ \mu(\langle e, x \rangle \ - \ 1) \ = \ \langle \mu e + c, x \rangle \ - \ \mu. \tag{24.10}$$

In this Lagrangian version, the dual problem involves maximizing a linear function on an interval - and the only issue is to explicitly determine the interval. This is a much simpler problem than the primal problem. Once we know the solution of this dual problem, then theorem 24.2 says that the primal problem has a solution and its value is either $-c_K$ or 0 respectively. This leads to the detemination of which vertex (or vertices) are minimizers of f on Δ_n. So the solution of the primal problem is helped by knowing the solution of the dual problem.

The following example illustrates what happens when there is no solution of the primal problem.

Example 24.2 Consider the problem of maximizing $f(x) := \langle c, x \rangle$ over the closed convex set $C \subset \mathbb{R}^n$ defined by

$$\sum_{j=1}^{n-1} x_j \ \leq \ 1 \quad \text{and} \quad x_j \ \geq \ 0 \quad \text{for } 1 \leq j \leq n. \tag{24.11}$$

This is an unbounded closed convex set as if $a := (a_1, \ldots, a_{n-1}, 0)$ lies in C then $a \ + \ te^{(n)} \in C$ for all $t \geq 0$. Moreover

$$f(a \ + \ te^{(n)}) \ = \ f(a) \ + \ tc_n$$

When $c_n > 0$, this function will have $\sup_{x \in C} \; f(x) \; = \; \infty$, and there is no maximizer of f on C.

This problem has the form of (24.2) with $B := b \; := \; (1, \ldots, 1, 0)$ being a $1 \times n$ matrix and $d = 1$. The Lagrangian for this maximization problem is similar to (24.8) but with b in place of e. Thus

$$\mathcal{L}(x, \mu) \; := \; \langle \mu b \; - \; c, x \rangle \; - \; \mu. \tag{24.12}$$

When $c_n \; \geq \; 0$, then one finds that the dual functional $G(\mu) \; = \; -\infty$ for all $\mu \; \geq \; 0$. Hence the value of the dual problem is $-\infty$ and the essential domain of the dual problem is empty as was stated in theorem 23.2.

Lecture 25; 11/22/2004

25. Sensitivity of Critical Points and Lagrange Multipliers.

In this section, yet another property of the Lagrange multipliers for an equality constrained optimization problem will be described. Consider the problem $(\mathcal{P})$ of minimizing a continuously differentiable function f on a set S defined by L equality constraints as in section 18. We will be interested in seeing how the solutions and the value of the problem change when the constraints vary.

Let $\{h_1, \ldots, h_L\}$ be continuously differentiable real valued functions on $\mathbb{R}^n$ and define

$$S(t) \; := \; \{x \in \mathbb{R}^n : h_l(x) \; = \; t \, d_l \quad \text{for } 1 \leq l \leq L\} \tag{25.1}$$

Here $d := (d_1, \ldots, d_L)$ is a fixed vector and assume w.l.o.g. $-1 < t < 1$.

A point $\hat{x}(t) \in S(t)$ is a *constrained critical point* for f on $S(t)$ provided it satisfies the Lagrange multiplier system (18.2).

$$\nabla f(x) \; + \; \sum_{l=1}^{L} \; \lambda_l \; \nabla \, h_l(x) \quad = \quad 0 \tag{25.2}$$

$$h_l(x) \quad = \quad t \, d_l \tag{25.3}$$

For each value of $t \; \in \; (-1, 1)$, this is a system of $n + L$ equations for $n + L$ unknowns; namely the n entries in $\hat{x}(t)$ and the L Lagrange multipliers $\lambda_l(t)$. Suppose that the curve $\{\hat{x}(t) : -1 < t < 1\}$ is a differentiable curve in $\mathbb{R}^n$.

Theorem 25.1 Suppose f, h_l as above are continuously differentiable and $\{\hat{x}(t) : -1 < t < 1\}$ is a differentiable curve of constrained critical points of f on $S(t)$. Let

$$\alpha(t) := f(\hat{x}(t)), \qquad \text{then} \quad \frac{d \, \alpha(t)}{dt} \; = \; -\langle \lambda(t), d \rangle \quad \text{for} \quad -1 < t < 1. \tag{25.4}$$

Proof: From the chain rule

$$\frac{d\,\alpha(t)}{dt} = \langle \nabla\, f(\hat{x}(t)), \frac{d\hat{x}(t)}{dt} \rangle \tag{25.5}$$

$$= -\sum_{l=1}^{L} \lambda_l\, \langle \nabla\, h_l(\hat{x}(t)), \frac{d\hat{x}(t)}{dt} \rangle \tag{25.6}$$

Differentiate both sides of (25.3) wrt t, then from the chain rule,

$$\langle \nabla\, h_l(\hat{x}(t)), \frac{d\hat{x}(t)}{dt} \rangle \;=\; d_l$$

Substitute this in the preceding equation to obtain (25.4) as claimed.

In particular this shows that λ_l will be the rate of change of the value function with respect to the $l-th$ constraint when all the other constraints are kept constant. Thus λ_l measures the *sensitivity* of the value of the function f with respect to a change in the $l-th$ constraint. In particular applications, especially in economics and finance, the Lagrange multipliers often are observable quantities.

26. Augmented Lagrangians for Equality Constrained Optimization.

In section 18, the Lagrange multiplier rule was developed for equality constrained optimization using a penalty functions formulation. In example 20.1, a Lagrangian for a linear equality constrained problem was described. These may be combined to provide probably the most effective way of finding solutions of equality constrained problems which is the augmented Lagrangian method due to Hestenes and Powell in 1969.

Assume we have a problem $(\mathcal{P})$ as in the preceding section and take $t = 0$. Define the function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^L \times (0,\infty) \to \mathbb{R}$ by

$$\mathcal{L}(x,\lambda,\sigma) \;:=\; f(x) \;-\; \sum_{l=1}^{L} \lambda_l\, h_l(x) \;+\; \frac{\sigma}{2} \sum_{l=1}^{L} h_l(x)^2. \tag{26.1}$$

Let $h : \mathbb{R}^n \to \mathbb{R}^L$ be defined by $h(x) := (h_1(x), \ldots, h_L(x))$, then this can be written

$$\mathcal{L}(x,\lambda,\sigma) \;:=\; f(x) \;-\; \langle \lambda, h(x) \rangle \;+\; \frac{\sigma}{2} \|h(x)\|_2^2. \tag{26.2}$$

Note that this Lagrangian is linear with respect to (λ,σ). Consider the problem of maximizing this with respect to (λ,σ). Then

$$\sup_{(\lambda,\sigma)} \mathcal{L}(x,\lambda,\sigma) = \begin{cases} f(x) & \text{if} \quad x \in S \\ \infty & \text{otherwise.} \end{cases} \tag{26.3}$$

so this function is a Lagrangian in the sense of section 20 for this problem.

Moreover when f, h are continuously differentiable on $\mathbb{R}^n$, then this Lagrangian is differentiable and its critical points are given by the solutions of

$$\nabla_x \mathcal{L}(x,\lambda,\sigma) \;=\; \nabla f(x) \;-\; \langle \lambda, \nabla h(x)\rangle \;+\; \sigma \sum_{l=1}^{L} h_l(x)\,\nabla h_l(x) \;=\; 0. \quad (26.4)$$

$$\frac{d}{d\lambda_l}\mathcal{L}(x,\lambda,\sigma) \;=\; -\,h_l(x) \;=\; 0 \qquad \text{for } 1 \le l \le L \quad (26.5)$$

If $\tilde{x}$ is a local minimizer of f on S, and it is a regular point for the equality constraints, then from the Lagrange multiplier theorem 18.1, $(\tilde{x}, \lambda)$ will be a critical point of this augmented Lagrangian for any choice of $\sigma > 0$.

So the *relaxed* version of the primal problem $(\mathcal{P})$ is the problem $(\mathcal{P}_\sigma)$ of minimizing $\mathcal{L}(x,\lambda,\sigma)$ with respect to $x \in \mathbb{R}^n$ for fixed (λ,σ). If this minimum exists, then it would be $G(\lambda,\sigma)$, the dual function for this Lagrangian defined as in (20.12). If we could find an (unconstrained) minimizer of this problem which also is in S, then we would have a possible minimizer of f on S.

Many algorithms have been developed to find solutions of this problem. Essentially we choose values of (λ_k, σ_k) and find an approximate minimizer. Then "update" the choice of the Lagrange multiplier, and perhaps increase σ. Note that the larger that σ is the greater the "penalty" for violating a constraint. Then minimize in x again. Good strategies for doing this depend on the problem itself, but for most problems mathematicians have found efficient ways to find minima in this manner.

Homework Number 6. (due 12/1/2004)

Question 1. Given a finite set of distinct points $S := \{a^{(1)}, \ldots, a^{(m)}\} \subset \mathbb{R}^n$ find the point $\bar{x}$ which minimizes $F : \mathbb{R}^n \times \mathbb{R} \to [0,\infty)$ defined by

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

where each $c_j > 0$.

(i) show that this function is convex and coercive and that this problem has a minimizer.

(ii) what conditions must hold at a minimizer of this problem?

(iii) find the solution when $m = 2, n > 1$.

(iv) find the solution when $n = 2, m = 3$, each $c_j \;=\; 1$ and $S := \{0, e^{(1)}, e^{(2)}\}$.

Question 2. Suppose h is the entropy function defined in example 21.3 and C is the closed convex subset of points in $[0,\infty)^n$ that satisfy

$$\langle c, x\rangle = 1 \qquad \text{with each } c_j > 0$$

(i) Find the conditions that must hold at a minimizer of h on C.
(ii) Hence find the unique minimizer of this problem.
(iii) Describe a Lagrangian for this problem.
(iv) Give an explicit formula for the dual function associated with the Lagrangian you have in (iii).

Question 3. Define $g : (0, \infty)^n \to (0, \infty)$ by

$$g(x) := \sum_{j=1}^{n} \frac{1}{x_j}$$

Determine the gradient of this function. Is it convex on its domain?

Let C be the subset of $(0, \infty)^n$ of points which also satisfy $\langle e, x \rangle = 1$.
Find the possible solutions of the extremality equations that are satisfied by a local minimizer of g on C. Hence find the minimizer of g on C and the value of g at this minimizer. (Hint; Try the case n = 2 first.)