C&O367: Nonlinear Optimization
(Winter 2013)
Assignment 2
H. Wolkowicz

Posted ???

Due: Thursday, Jan. 31, 10:00AM (before class),

1 Convex Sets

1.1 Intersection of Convex Sets

Show that the intersection of convex sets is a convex set. Is the union of convex sets convex? (If not provide a counterexample.)

> **Solution:** Let $C_1, \cdots, C_k$ be convex sets, and define $C := \cap C_k$. Assume $x, y \in C$ and $\lambda \in [0,1]$. By definition of C we have $x, y \in C_i$, $i \in \{1, \cdots, k\}$, and by convexity we have $\lambda x + (1-\lambda) y \in C_i$, $i \in \{1, \cdots, k\}$. This means $\lambda x + (1-\lambda) y \in C$, so C is convex. (A similar argument holds for an infinite number of sets.)
>
> This is not true for the union of convex sets. Let $C_1 := \{0\}$ and $C_2 := \{1\}$ as subsets of $\mathbb{R}$. Both sets are clearly convex, but the union of them is not convex.

1.2 Midpoint Convex

A set is $C \subseteq \mathbb{R}^n$ is *midpoint convex* if whenever $x, y \in C$ we have $\frac{1}{2}(x+y) \in C$. It is clear that C convex implies that C is midpoint convex.

1. Suppose that C is a *closed* midpoint convex set. Show that C is a convex set.

   Solution: Let $x, y \in C$ and assume that z is a point on the line segment between x and y. We construct a sequence $\{z_i\} := \frac{x_i + y_i}{2}$, where $x_0 = x$, $y_0 = y$, and

   $$x_{i+1} = \begin{cases} z_i & z_i \leq z \\ x_i & z_i > z \end{cases} \qquad y_{i+1} = \begin{cases} z_i & z_i > z \\ y_i & z_i \geq z \end{cases}$$

   By this construction, z is always on the line segment between x_i and y_i, so we have $\|z_i - z\| \leq \|x - y\| 2^{-i}$ which means $\{z_i\}$ converges to z. By midpoint convexity of C we have $z_i \in C$ for all i, and because C is closed we have $z \in C$. This is true for all z on the line segment between x and y, so C is convex.

2. Find an example of a set C that is midpoint convex but is *not* a convex set.

Solution: Q, the set of all rational numbers is midpoint convex, but clearly is not convex.

1.3 Halfspaces

1. Give conditions that guarantee that one halfspace contains another, i.e.,

$$\{x \in \mathbb{R}^n : a^T x \leq b\} \subseteq \{x \in \mathbb{R}^n : \bar{a}^T x \leq \bar{b}\},$$

where both $a, \bar{a}$ are not zero.

Solution: Let's denote the first halfspace $\{x \in \mathbb{R}^n : a^T x \leq b\}$ by H_1 and the other one by H_2. First I prove that there exists $\alpha > 0$ such that $a = \alpha\bar{a}$, i.e., a and $\bar{a}$ are parallel. If not, there exists a vector w such that $a^T w = 0$ and $\bar{a}^T w > 0$. Now for any point $x \in H_1$, we have $a^T(x + \lambda w) = a^T x \leq b$. This means for all values $\lambda > 0$, $a^T(x + \lambda w) \in H_1$. However, we have $\bar{a}^T(x + \lambda w) \to \infty$ when λ goes to $+\infty$, and $\bar{a}^T(x + \lambda w)$ is not in H_2 for all values of λ which is a contradiction. Hence, $a = \alpha\bar{a}$. It's easy to show that for $H_1 \subseteq H_2$, we must have $b \leq \alpha\bar{b}$.

2. When are the halfspaces equal?

Solution: In this case we have $H_1 \subseteq H_2$ and $H_2 \subseteq H_1$. By the above explanation, there exists $\alpha > 0$ such that $a = \alpha\bar{a}$ and $b = \alpha\bar{b}$.

2 Convex Functions

2.1 Three slope rule

Suppose that $f : \mathbb{R} \to \mathbb{R}$ is convex and that $x < y < z$. Show that

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} \leq \frac{f(z) - f(y)}{z - y}.$$

Solution: By simple calculations we can write y as the convex combination of x and z as follows:

$$y = \frac{z - y}{z - x}x + \frac{y - x}{z - x}z,$$
$$\frac{z - y}{z - x} + \frac{y - x}{z - x} = 1$$

convexity of f implies that

$$f(y) \leq \frac{z - y}{z - x}f(x) + \frac{y - x}{z - x}f(z).$$

Both of the inequalities are derived by simple reformulations.

2.2 Convex and Strictly Convex Functions

Are the following functions convex or strictly convex on the specified sets:

1. $f(x, y) = 5x^2 + 2xy + y^2 - x + 2y + 3$ on $D = \mathbb{R}^2$.

   Solution: The Hessian of the function is

   $$\nabla^2 f(x, y) = \begin{bmatrix} 10 & 2 \\ 2 & 2 \end{bmatrix},$$

   which is positive definite, so the function is strictly convex.

2. $f(x, y) = c_1 x_1 + c_2/x_1 + c_3 x_2 + c_4/x_2$ on $D = \mathbb{R}^2_{++}$, where the c_i are all positive numbers.

   Solution: The Hessian of the function is

   $$\nabla^2 f(x, y) = \begin{bmatrix} 2c_2 x^{-3} & 0 \\ 0 & 2c_4 y^{-3} \end{bmatrix},$$

   which is positive definite for $(x, y) \in \mathbb{R}^2_{++}$ and $c_i > 0$, so the function is strictly convex on the domain.

2.3 Quadratic Convex Functions

A quadratic function is one that can be expressed as

$$q(x) = x \cdot Ax + b \cdot x + \alpha,$$

where $A \in \mathcal{S}^n$, the space of real symmetric $n \times n$ matrices, $b \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$.

1. Show that the function

   $$q(x) = (x_1 - x_2)^2 + (x_1 + 2x_2 + 1)^2 - 8x_1 x_2$$

   is a quadratic function by finding the appropriate A, b, α.

   Solution: If you expand $q(x)$, you can see that it's a quadratic function with the following parameters:

   $$A = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}, \quad b = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \alpha = 1.$$

2. Compute the gradient and Hessian of $q(x)$ and express them using the A, b, α computed above in Item 1.

   Solution: It is easy to check that $\nabla q(x) = 2Ax + b$ and $\nabla^2 q(x) = 2A$ which is true for any quadratic function.

3. Show that a quadratic function $q(x)$ is convex if, and only if, A is positive semidefinite and it is strictly convex if A is positive definite.

   Solution: By the fact that $\nabla^2 q(x) = 2A$, the result is clear from the theorem proved for a general convex function. But we use the following proof which also solves the next question. For any $x, y \in \mathbb{R}^n$ and any $\lambda \in (0, 1)$, let's define $Q(x, y, \lambda)$ as

   $$Q(x, y, \lambda) := \lambda q(x) + (1-\lambda)q(y) - q(\lambda x + (1-\lambda)y) = \lambda(1-\lambda)(x-y)^T A(x-y)$$

   $q(x)$ is convex if and only if $Q(x, y, \lambda) \geq 0$ for all $x, y \in \mathbb{R}^n$ and any $\lambda \in (0, 1)$. As $\lambda(1-\lambda) > 0$, this is true if and only if A is positive semidefinite. $q(x)$ is strictly convex if and only if $Q(x, y, \lambda) > 0$ for all $x, y \in \mathbb{R}^n$ and any $\lambda \in (0, 1)$, which is true if and only if A is positive definite.

4. Suppose that the quadratic function $q(x)$ is strictly convex. Is the matrix A positive definite?

   Solution: Yes, Proved in the previous question.

5. Suppose that $q(x)$ is a quadratic function of n variables such that the corresponding matrix A is positive definite. Show that $0 = 2Ax + b$ has a unique solution and that this solution is the strict global minimizer of $q(x)$.

   Solution: A is positive definite, so is invertible. Then $2Ax + b = 0$ has the unique solution $x^* = -\frac{1}{2}A^{-1}b$. x^* is a critical point and $q(x)$ is strictly convex, so x^* is the unique global minimizer.

6. Suppose that $q(x)$ is a quadratic function of n variables. Find necessary and sufficient conditions for $q(x)$ to be bounded below.

   Solution: I claim that $q(x)$ is bounded below if and only if A is positive semidefinite and $b \in \mathcal{R}(A)$ (range of A). First note that for a convex function, point x^* is a global minimizer if and only if it's gradient is equal to zero (this is true only for convex functions). Assume that A is positive semidefinite and $b \in \mathcal{R}(A)$, then by the above results, $q(x)$ is convex. $b \in \mathcal{R}(A)$ means that $2Ax + b = 0$ has a solution which is a global minimizer; $q(x)$ is bounded below.

   Now, assume that $q(x)$ is bounded below. If A is not positive semidefinite, then there exists y such that $y^T Ay < 0$. We may assume that $b^T y \leq 0$, otherwise take $-y$ instead. Then $q(\lambda y)$ goes to $-\infty$ when λ goes to $+\infty$ which is a contradiction. Hence, A is positive semidefinite. If $b \notin \mathcal{R}(A)$, then b is not orthogonal to the null space of $A^T = A$. Hence, there exists y such that $y^T Ay = 0$ and $b^T y < 0$. Again, $q(\lambda y)$ goes to $-\infty$ when λ goes to $+\infty$ which is a contradiction; we must have $b \in \mathcal{R}(A)$.