Homework Assignment 8

Matthew Tiger

May 2, 2016

Problem 1. Write at least two necessary conditions and at least two sufficient conditions for a function $f : \mathbb{R}^n \to \mathbb{R}$ to be concave.

Solution. By definition, a function $f : \mathbb{R}^n \to \mathbb{R}$ is concave over the convex set $\Omega \subset \mathbb{R}^n$ if $-f$ is convex over Ω. This definition will allow us to obtain results for concave functions by replacing f with $-f$ in previously obtained results concerning convex functions.

Using this definition and Theorem 22.2, we see that the condition that if for all $\alpha \in (0, 1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, we have that

$$f(\alpha\boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \geq \alpha f(\boldsymbol{x}) + (1 - \alpha)f(\boldsymbol{y})$$

is a necessary and sufficient condition for f to be concave on the convex set Ω.

Further, if the function f is $\mathcal{C}^1$-smooth, we see from the above definition and Theorem 22.4 that the condition that if for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega$, we have that

$$f(\boldsymbol{x}) \leq f(\boldsymbol{y}) + Df(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{y})$$

is a necessary and sufficient condition for f to be concave on the open convex set Ω.

Going one last step further, if the function f is $\mathcal{C}^2$-smooth, we see from the above definition and Theorem 22.5 that the condition that if for all $\boldsymbol{x} \in \Omega$, the Hessian matrix $\boldsymbol{F}(\boldsymbol{x})$ of f at $\boldsymbol{x}$ is a negative semi-definite matrix is a necessary and sufficient condition for f to be concave on the open convex set Ω. $\square$

Problem 2. Let $S \subset \mathbb{R}^n$ be a convex set and let $\boldsymbol{x}^* \in S$. Prove that a vector $\boldsymbol{d} \in \mathbb{R}^n$ is a feasible direction at $\boldsymbol{x}^*$ (relative to S) if and only there exists $t_0 > 0$ such that $\boldsymbol{x}^* + t_0\boldsymbol{d} \in S$ with $\boldsymbol{d} \neq \boldsymbol{0}$.

Solution. Suppose first that the vector $\boldsymbol{d} \in \mathbb{R}^n$ is a feasible direction at $\boldsymbol{x}^*$ (relative to S). By definition, the vector $\boldsymbol{d}$ is a feasible direction at $\boldsymbol{x}^* \in S$ if there exists $t_0 > 0$ such that $\boldsymbol{x}^* + t\boldsymbol{d} \in S$ for all $t \in [0, t_0]$ with $\boldsymbol{d} \neq \boldsymbol{0}$. Thus, choosing $t = t_0$, we have by the above definition that there exists $t_0 > 0$ such that $\boldsymbol{x}^* + t_0\boldsymbol{d} \in S$ with $\boldsymbol{d} \neq \boldsymbol{0}$, proving the first implication.

Now suppose that there exists $t_0 > 0$ such that $\boldsymbol{x}^* + t_0\boldsymbol{d} \in S$ with $\boldsymbol{d} \neq \boldsymbol{0}$. Since S is convex and $\boldsymbol{x}^* \in S$, any convex combination of $\boldsymbol{x}^*$ and $\boldsymbol{x}^* + t_0\boldsymbol{d}$ will also be in S, i.e. for all $\alpha \in [0, 1]$, we have that

$$\alpha\boldsymbol{x}^* + (1-\alpha)(\boldsymbol{x}^* + t_0\boldsymbol{d}) = \boldsymbol{x}^* + (1-\alpha)t_0\boldsymbol{d} \in S.$$

Since $t_0 > 0$, we have that the following two sets are equal:

$$\{(1-\alpha)t_0 \mid 0 \leq \alpha \leq 1\} = \{t \mid 0 \leq t \leq t_0\}.$$

Thus, if $\boldsymbol{x}^* + (1-\alpha)t_0\boldsymbol{d} \in S$ for all $\alpha \in [0, 1]$, then $\boldsymbol{x}^* + t\boldsymbol{d} \in S$ for all $t \in [0, t_0]$. Therefore, if $\boldsymbol{x}^* \in S$ with S a convex set and there exists $t_0 > 0$ such that $\boldsymbol{x}^* + t_0\boldsymbol{d} \in S$ with $\boldsymbol{d} \neq \boldsymbol{0}$, then $\boldsymbol{x}^* + t\boldsymbol{d} \in S$ for all $t \in [0, t_0]$, i.e. the vector $\boldsymbol{d}$ is a feasible direction at $\boldsymbol{x}^*$ (relative to S). $\square$

Problem 3. Recall that

$$\max\{\alpha, \beta\} := \begin{cases} \alpha & \text{if } \alpha \geq \beta \\ \beta & \text{if } \alpha < \beta \end{cases}.$$

Given two convex functions $f_1 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and $f_2 : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, prove that for $\boldsymbol{x} \in \mathbb{R}^n$, the function

$$f(\boldsymbol{x}) := \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$$

is convex.

Solution. Note that it is clear that the set $\mathbb{R}^n$ is convex. Therefore, the function $f(\boldsymbol{x}) := \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$ is convex if for all $\alpha \in (0, 1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we have that

$$f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha)f(\boldsymbol{y}). \tag{1}$$

If either $f(\boldsymbol{x}) = +\infty$ or $f(\boldsymbol{y}) = +\infty$, then for all $\alpha \in (0, 1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we see that $\alpha f(\boldsymbol{x}) + (1-\alpha)f(\boldsymbol{y}) = +\infty$ and inequality (1) holds showing the convexity of f in this case.

Now suppose that both $f(\boldsymbol{x})$ and $f(\boldsymbol{y})$ are finite. Without loss of generality, we may assume that at the point $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}$, we have that the max of f_1 and f_2 at that point occurs at f_1, i.e.

$$\begin{aligned} f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &= \max\{f_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}), f_2(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y})\} \\ &= f_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}). \end{aligned}$$

Then, due to the convexity of f_1, we have that for all $\alpha \in (0, 1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$f_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f_1(\boldsymbol{x}) + (1-\alpha)f_1(\boldsymbol{y}).$$

From the above definition of max, we readily see that

$$\begin{aligned} \alpha f_1(\boldsymbol{x}) + (1-\alpha)f_1(\boldsymbol{y}) &\leq \alpha \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\} + (1-\alpha)\max\{f_1(\boldsymbol{y}), f_2(\boldsymbol{y})\} \\ &= \alpha f(\boldsymbol{x}) + (1-\alpha)f(\boldsymbol{y}). \end{aligned}$$

Therefore, combining, we have that for all $\alpha \in (0, 1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) = f_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha)f(\boldsymbol{y})$$

showing that inequality (1) holds and that the function f is convex. $\square$

Problem 4. Consider the pair of linear programming problems in asymmetric duality:

$$
\begin{array}{llll}
(P_a) & \text{minimize} & f(\boldsymbol{x}) = \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\
 & \text{subject to} & A\boldsymbol{x} = \boldsymbol{b} \\
 & & \boldsymbol{x} \geq \boldsymbol{0}
\end{array}
\qquad
\begin{array}{lll}
(D_a) & \text{maximize} & F(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^\mathsf{T}\boldsymbol{b} \\
 & \text{subject to} & \boldsymbol{\lambda}^\mathsf{T}A \leq \boldsymbol{c}^\mathsf{T}
\end{array}
$$

a. Prove that (D_a) is a convex programming problem.

b. Write the KKT conditions for (D_a).

c. Suppose that $\boldsymbol{x}^*$ is feasible for (P_a) and $\boldsymbol{\lambda}^*$ is feasible for (D_a). Use the KKT conditions to prove that if $(\boldsymbol{c}^\mathsf{T} - \boldsymbol{\lambda}^{*\mathsf{T}}A)\boldsymbol{x}^* = 0$, then $\boldsymbol{\lambda}^*$ is optimal for (D_a).

Solution. For the problem above, we assume that $\boldsymbol{x} \in \mathbb{R}^n$ and that A is an $m \times n$ matrix with $m < n$. This implies that $\boldsymbol{\lambda} \in \mathbb{R}^m$.

a. Note that (D_a) is a convex programming problem if the constraint set

$$\Omega = \{\boldsymbol{\lambda} \in \mathbb{R}^m \mid \boldsymbol{\lambda}^\mathsf{T}A \leq \boldsymbol{c}^\mathsf{T}\}$$

is a convex set and if $F : \Omega \to \mathbb{R}^m \cup \{+\infty\}$ where $F(\boldsymbol{\lambda}) := \boldsymbol{\lambda}^\mathsf{T}\boldsymbol{b}$ is a convex function.

It is straight-forward to show that Ω is a convex set. Let $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ be given. Then $\boldsymbol{x}^\mathsf{T}A \leq \boldsymbol{c}^\mathsf{T}, \boldsymbol{y}^\mathsf{T}A \leq \boldsymbol{c}^\mathsf{T}$ and for all $\alpha \in [0, 1]$

$$\alpha\boldsymbol{x}^\mathsf{T}A \leq \alpha\boldsymbol{c}^\mathsf{T} \quad \text{and} \quad (1-\alpha)\boldsymbol{y}^\mathsf{T}A \leq (1-\alpha)\boldsymbol{c}^\mathsf{T}.$$

Thus, for all $\alpha \in [0, 1]$

$$(\alpha\boldsymbol{x}^\mathsf{T} + (1-\alpha)\boldsymbol{y}^\mathsf{T})A = \alpha\boldsymbol{x}^\mathsf{T}A + (1-\alpha)\boldsymbol{y}^\mathsf{T}A \leq \alpha\boldsymbol{c}^\mathsf{T} + (1-\alpha)\boldsymbol{c}^\mathsf{T} = \boldsymbol{c}^\mathsf{T},$$

i.e. the convex combination $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \Omega$ and that Ω is a convex set.

Since F is a linear function, it is of course convex showing that the problem (D_a) is a convex programming problem.

b. The KKT conditions for the convex programming problem (D_a) are stated below:

If the function $F \in \mathcal{C}^1$ is a convex function on the convex set of feasible points

$$\Omega = \{\boldsymbol{\lambda} \in \mathbb{R}^m \mid \boldsymbol{\lambda}^\mathsf{T}A \leq \boldsymbol{c}^\mathsf{T}\} = \{\boldsymbol{\lambda} \in \mathbb{R}^m \mid \boldsymbol{g}(\boldsymbol{\lambda}) = A^\mathsf{T}\boldsymbol{\lambda} - \boldsymbol{c} \leq \boldsymbol{0}\}$$

where $\boldsymbol{g} \in \mathcal{C}^1$ and if there exists $\boldsymbol{\lambda}^* \in \Omega$, $\boldsymbol{\mu}^* \in \mathbb{R}^n$ such that

i $\boldsymbol{\mu}^* \geq \boldsymbol{0}$.

ii $-DF(\boldsymbol{\lambda}^*) + \boldsymbol{\mu}^{*\mathsf{T}}D\boldsymbol{g}(\boldsymbol{\lambda}^*) = -\boldsymbol{b}^\mathsf{T} + \boldsymbol{\mu}^{*\mathsf{T}}A^\mathsf{T} = \boldsymbol{0}^\mathsf{T}$.

iii $\boldsymbol{\mu}^{*\mathsf{T}}\boldsymbol{g}(\boldsymbol{\lambda}) = \boldsymbol{\mu}^{*\mathsf{T}}(A^\mathsf{T}\boldsymbol{\lambda}^* - \boldsymbol{c}) = (\boldsymbol{\lambda}^{*\mathsf{T}}A - \boldsymbol{c}^\mathsf{T})\boldsymbol{\mu}^* = 0$.

then $\boldsymbol{\lambda}^*$ is a global maximizer of F over Ω.

c. Suppose that $\boldsymbol{x}^*$ is feasible for (P_a) and $\boldsymbol{\lambda}^*$ is feasible for (D_a) and that $(\boldsymbol{c}^\top - \boldsymbol{\lambda}^{*\top} A)\boldsymbol{x}^* = 0$. Since $\boldsymbol{\lambda}^*$ is feasible for (D_a), we know that $\boldsymbol{\lambda}^* \in \Omega$. Choose $\boldsymbol{\mu}^* = \boldsymbol{x}^*$. Then $\boldsymbol{\mu}^*$ satisfies the KKT conditions above, which we will now demonstrate.

Since $\boldsymbol{x}^*$ is feasible for (P_a), condition i. is readily seen to be true and since $A\boldsymbol{x}^* = \boldsymbol{b}$ or $\boldsymbol{b} - A\boldsymbol{x}^* = \mathbf{0}$, we see that

$$\mathbf{0}^\top = \boldsymbol{b}^{*\top} - (A\boldsymbol{x}^*)^\top = \boldsymbol{b}^{*\top} - \boldsymbol{x}^{*\top} A^\top = \boldsymbol{b}^{*\top} - \boldsymbol{\mu}^{*\top} A^\top$$

or that $-\boldsymbol{b}^{*\top} + \boldsymbol{\mu}^{*\top} A^\top = \mathbf{0}^\top$ and condition ii. is satisfied. Since by assumption we have that $(\boldsymbol{c}^\top - \boldsymbol{\lambda}^{*\top} A)\boldsymbol{x}^* = (\boldsymbol{c}^\top - \boldsymbol{\lambda}^{*\top} A)\boldsymbol{\mu}^* = 0$, we see also that $(\boldsymbol{\lambda}^{*\top} A - \boldsymbol{c}^\top)\boldsymbol{\mu}^* = 0$ and condition iii. is satisfied. Therefore, since $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ both satisfy the KKT conditions stated above, $\boldsymbol{\lambda}^*$ is optimal for (D_a).

$\square$