EE 227BT 10/7/14
L. El Ghaoui

Midterm: Solutions

1. (10 points) *Conjugates of some functions.* Express the conjugate of the following functions in closed form. Make sure to define precisely the domain of the conjugate functions, and their values.

   (a) $f(x) = a^T x + b$, with $a \in \mathbf{R}^n$, $b \in \mathbf{R}$.

   (b) The function with domain $\{x : \|x\|_2 \le 1\}$, and values on that domain given by $f(x) = a^T x + b$, with $a \in \mathbf{R}^n$, $b \in \mathbf{R}$.

   (c) $f(x) = -\|x\|_2$. (Note that f is not convex!)

   (d) $f(x) = \frac{1}{p}(x_1^p + \ldots + x_p^p)$, with $p > 1$, $x \in \mathbf{R}_+^p$. (It will be convenient to use $q = p/(p-1)$.)

   (e) $f(x) = \log(1 + e^{c^T x})$, with $c \in \mathbf{R}^n$, $\|c\|_2 = 1$, given.

   Solution:

   (a) $f(x) = a^T x + b$, with $a \in \mathbf{R}^n$, $b \in \mathbf{R}$: we have

   $$f^*(y) = \max_x \; y^T x - (a^T x + b) = \begin{cases} -b & \text{if } y = a, \\ +\infty & \text{otherwise.} \end{cases}$$

   (b) The function with domain $\{x : \|x\|_2 \le 1\}$, and values on that domain given by $f(x) = a^T x + b$, with $a \in \mathbf{R}^n$, $b \in \mathbf{R}$: we have

   $$\begin{aligned} f^*(y) &= -b + \max_x \; y^T x - a^T x \; : \; \|x\|_2 \le 1 \\ &= \|y - a\|_2 - b. \end{aligned}$$

   (c) $f(x) = -\|x\|_2$: in this case,

   $$f^*(y) \;=\; \max_x \; y^T x + \|x\|_2 = +\infty,$$

   irrespective of the vector y. The domain of f^* is empty.

   (d) $f(x) = \frac{1}{p}(x_1^p + \ldots + x_p^p)$, with $p > 1$, $x \in \mathbf{R}_+^p$: the function is convex, and differentiable in its domain. Taking derivatives of

   $$y_i x_i - \frac{1}{p} x_i^p$$

leads to $y_i = x_i^{p-1}$. This requires $y_i \geq 0$, otherwise $f^*(y) = +\infty$. In such a case, we proceed and obtain the optimal point $x_i^* = y_i^{1/(p-1)}$, $i = 1, \ldots, p$, and

$$\begin{aligned} f^*(y) &= \max_x \sum_{i=1}^n \left(y_i x_i - \frac{1}{p} x_i^p \right) \\ &= \sum_{i=1}^n \left(y_i x_i^* - \frac{1}{p} (x_i^*)^p \right) \\ &= \frac{1}{q} \sum_{i=1}^p y_i^q, \end{aligned}$$

with $1/p + 1/q = 1$.

(e) $f(x) = \log(1 + e^{c^T x})$, with $c \in \mathbf{R}^n$, $\|c\|_2 = 1$, given: any $x \in \mathbf{R}^n$ can be written as $x = \alpha c + r$, with $\alpha \in \mathbf{R}$, and $r \in \mathbf{R}^n$, with $c^T r = 0$. Using α, r as new variables, we have

$$\begin{aligned} f^*(y) &= \max_x \; y^T x - \log(1 + e^{c^T x}) \\ &= \max_\alpha \; \alpha y^T c - \log(1 + e^\alpha) + \max_{r \,:\, c^T r = 0} y^T r. \end{aligned}$$

The second term is finite (and in fact, zero) if and only y, c are parallel, which, due to the fact that $c \neq 0$, is the same as $y = \beta c$ for some $\beta \in \mathbf{R}$. We obtain in that case

$$f^*(y) = \max_\alpha \; \alpha\beta - \log(1 + e^\alpha).$$

Taking derivates leads to

$$e^\alpha = \frac{\beta}{1 - \beta},$$

which requires $\beta \in [0, 1]$. If that is not the case, the dual function is $+\infty$. Otherwise, we have

$$f^*(y) = \beta \log \beta + (1 - \beta) \log(1 - \beta).$$

To conclude we have

$$f^*(y) = \begin{cases} \beta \log \beta + (1 - \beta) \log(1 - \beta) & \text{if } y = \beta c \text{ for some } \beta \in [0, 1] \\ +\infty & \text{otherwise.} \end{cases}$$

2. (10 points) *Convexity of a function via Hessian.*

Consider the function $f : \mathbf{R}^n_{++} \to \mathbf{R}$, with values

$$f(x) = \sum_{i=1}^n \frac{1}{x_i} - \frac{1}{\sum_{i=1}^n x_i}.$$

Show that the function is convex, using the following steps.

(a) Show that the Hessian at $x \in \mathbf{R}^n_{++}$ satisfies

$$\frac{s(x)^3}{2} \nabla^2 f(x) = D(x) - \mathbf{1}\mathbf{1}^T,$$

where $s(x) = \mathbf{1}^T x$, with $\mathbf{1}$ the vector of ones, and $D(x)$ a diagonal matrix, which you will determine.

(b) Use Schur complements (twice) to show that the function f is convex.

Solution:

(a) Using the chain rule, we easily obtain

$$\frac{1}{2} \nabla^2 f(x) = \mathbf{diag}(\frac{1}{x_1^3}, \ldots, \frac{1}{x_n^3}) - \frac{1}{(x_1 + \ldots + x_n)^3} \mathbf{1}\mathbf{1}^T,$$

from which derives the desired expression, with

$$s(x) = x_1 + \ldots + x_n = \mathbf{1}^T x, \quad D(x) = \mathbf{diag}(d(x)),$$

where

$$d_i(x) = \frac{s(x)^3}{x_i^3}, \quad i = 1, \ldots, n.$$

(b) We have to show that $D(x) - \mathbf{1}\mathbf{1}^T$ is positive semidefinite. Using Shcur complements, this condition is the same as

$$\begin{pmatrix} D(x) & \mathbf{1} \\ \mathbf{1}^T & 1 \end{pmatrix} \succeq 0.$$

Using Schur complements again, and exploiting the fact that $D(x) \succ 0$ on $\mathbf{dom}\, f$, we obtain the equivalent condition

$$1 \geq \mathbf{1}^T D(x)^{-1} \mathbf{1} = \sum_{i=1}^n \frac{1}{d_i(x)} = \frac{\sum_{i=1}^n x_i^3}{(\sum_{i=1}^n x_i)^3},$$

which is trivially true whenever $x \in \mathbf{R}^n_{++}$.

3. (10 points) *Some duals.* In the following questions, make sure to prove strong duality when needed.

   (a) For a given $a \in \mathbf{R}^n$, find the dual of the problem

   $$p^* = \min_x \; a^T x + \lambda \|x\|_2,$$

   and find a closed-form expression for p^*. *Hint:* express the problem in min-max form, and apply Sion's theorem.

   (b) For a given $a \in \mathbf{R}^n$, and $\lambda \geq 0$, consider the problem:

   $$p^* = \min_x \; \|x - a\|_2 + \lambda \|x\|_1.$$

   Find the dual.

   (c) Show that, for the problem above, when $\lambda > 1$, $x = 0$ is optimal for the primal problem. *Hint:* observe that the dual constraints on the dual variable u are of the form $\|u\|_2 \leq 1$, $\|u\|_\infty \leq \lambda$, and argue that for $\lambda > 1$ the second constraint is inactive, then show that $p^* = \|a\|_2$.

Solution:

(a) We have

$$p^* = \min_x \; a^T x + \lambda \|x\|_2 = \min_x \; \max_{u \,:\, \|u\|_2 \leq 1} \; (a - u)^T x$$

Sion's theorem applies since the variable u is constrained to a compact set. We obtain

$$p^* = \max_{u \,:\, \|u\|_2 \leq 1} \; \min_x \; (a - u)^T x$$

The inner problem's value is 0 when $a = u$, and $-\infty$ otherwise. We obtain the dual form

$$p^* = \max_u \; 0 \;:\; u = a, \;\; \|u\|_2 \leq \lambda.$$

This means that

$$p^* = \begin{cases} 0 & \text{if } \|a\|_2 \leq \lambda, \\ -\infty & \text{otherwise.} \end{cases}$$

(b) We have

$$\begin{aligned} p^* &= \min_x \; \|x - a\|_2 + \lambda \|x\|_1 \\ &= \min_x \; \max_{u,v} \; u^T(a - x) + v^T x \;:\; \|u\|_2 \leq 1, \;\; \|v\|_\infty \leq \lambda. \end{aligned}$$

Applying Sion's theorem, we obtain

$$\begin{aligned} p^* &= \max_{u,v} \; \min_x \; u^T(a - x) + v^T x \;:\; \|u\|_2 \leq 1, \;\; \|v\|_\infty \leq \lambda \\ &= \max_{u,v} \; u^T a \;:\; \|u\|_2 \leq 1, \;\; \|v\|_\infty \leq \lambda, \;\; u = v \\ &= \max_u \; u^T a \;:\; \|u\|_2 \leq 1, \;\; \|u\|_\infty \leq \lambda. \end{aligned}$$

(c) We observe that if $\lambda > 1$, then no matter what the optimal u is, we have

$$\|u\|_\infty \leq \|u\|_2 \leq 1 < \lambda.$$

Hence the second constraint is not active at optimum in that case, and we can remove it:

$$p^* \quad = \quad \max_u \ u^T a \ : \ \|u\|_2 \leq 1 = \|a\|_2.$$

This means that $x = 0$ is optimal, as it achieves the optimal value p^*.