Homework Assignment #3 Solutions

1. *More convex functions.*

   (a) *Proof by Hessian.* We prove convexity by showing that the Hessian is positive semidefinite:

$$\nabla^2 f(t,x) = \frac{2}{(t^2 - x^T x)^2} \begin{bmatrix} t^2 + x^T x & -2tx^T \\ -2tx & (t^2 - x^T x)I + 2xx^T \end{bmatrix}. \tag{1}$$

We make use of the Schur complement result (homwework 1, problem 2) A symmetric matrix of the form

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

is positive definite if and only if $C \succ 0$ and $A - BC^{-1}B^T \succ 0$.
Applying this to

$$A = t^2 + x^T x, \quad B = -2tx^T, \quad C = (t^2 - x^T x)I + 2xx^T,$$

we find that

$$C = (t^2 - x^T x)I + 2xx^T \succ 0$$

because $t^2 - x^T x > 0$. Also

$$C^{-1} = \frac{1}{t^2 - x^T x}\left(I - \frac{2xx^T}{t^2 + x^T x}\right)$$

which can be shown using the formula for the rank-one modification of identity,

$$(I + uu^T)^{-1} = I + \frac{1}{1 - u^T u} uu^T$$

valid for $\|u\| < 1$. Therefore we have

$$\begin{aligned} A - BC^{-1}B^T &= t^2 + x^T x - \frac{4t^2}{t^2 - x^T x} x^T \left(I - \frac{2xx^T}{t^2 + x^T x}\right) x \\ &= t^2 + x^T x - \frac{4t^2}{t^2 - x^T x}\left(x^T x - \frac{2x^T x x^T x}{t^2 + x^T x}\right) \\ &= t^2 + x^T x - \frac{4t^2 x^T x}{t^2 - x^T x}\left(1 - \frac{2x^T x}{t^2 + x^T x}\right) \\ &= t^2 + x^T x - \frac{4t^2 x^T x}{t^2 - x^T x}\left(\frac{t^2 - x^T x}{t^2 + x^T x}\right) \end{aligned}$$

$$
\begin{aligned}
&= t^2 + x^T x - \frac{4t^2 x^T x}{t^2 + x^T x} \\
&= \frac{(t^2 + x^T x)^2 - 4t^2 x^T x}{t^2 + x^T x} \\
&= \frac{(t^2 - x^T x)^2}{t^2 + x^T x} \\
&\geq 0
\end{aligned}
$$

so $\nabla^2 f(x,t) \succeq 0$ and we are done.

This way of proving convexity is tedious! Here is one more elegant proof.

Proof by composition. First, we note that the function

$$g(x,t) = \frac{x^T x}{t}$$

is convex on $\mathbf{dom}\, f$. Indeed, its epigraph $\mathbf{epi}\, g = \{(x,t,\alpha) \mid g(x,t) \leq \alpha\}$ is convex, since $g(x,t) \leq \alpha$ when $\|x\| < t$, is equivalent to

$$\begin{bmatrix} \alpha & x^T \\ x & tI \end{bmatrix} \succeq 0.$$

(We have used the Schur complement result proved in homework # 1.) By the composition rules, we conclude that $-\log g(x,t)$ is convex. Hence $f(x,t) = -\log t - \log(g(x,t)$ is also convex on its domain.

(b) The function has the epigraph

$$\mathbf{epi}\, f = \left\{(X,Y,t) \mid X \succ 0, \ \ \mathbf{Tr}\, Y^T X^{-1} Y \leq t\right\}.$$

For given $X \succ 0$, Y, t, the inequalities $X \succ 0$ and $\mathbf{Tr}\, Y^T X^{-1} Y \leq t$ hold if and only if there exist a Z such that $\mathbf{Tr}\, Z \leq t$, $Y^T X^{-1} Y \preceq Z$, which (by the Schur complement result of problem 2 in homework 1) can be rewritten

$$\begin{bmatrix} Z & Y^T \\ Y & X \end{bmatrix} \succeq 0.$$

Thus, $\mathbf{epi}\, f$ is the projection (on the space of (X,Y,t)-variables) of the set

$$\left\{(X,Y,Z,t) \mid X \succ 0, \ \begin{bmatrix} Z & Y^T \\ Y & X \end{bmatrix} \succeq 0, \ \ \mathbf{Tr}\, Z \leq t.\right\}.$$

The above set is convex, as it is the intersection of the cone of positive semidefinite matrices of size $(n+k) \times (n+k)$, with a halfspace. Thus, f is convex.

2. *Examples of conjuguate functions.*

(a) For the indicator function, the conjuguate is

$$\delta_C^*(y) = \sup_x (x^T y - \delta_C(x)) = \sup_{x \in C} x^T y.$$

The conjuguate of the indicator function is the support function of the set C. For the support function, we have

$$s_C^*(y) = \sup_x (x^T y - \sup_{z \in C} z^T x) = \sup_x \inf_{z \in C} x^T (y - z).$$

We obviously have

$$s_C^*(y) \leq \inf_{z \in C} \sup_x x^T (y - z).$$

For every $z \in C$, the supremum in the above expression is infinite, except when it is zero, for $y = z$. But y can only be equal to some $z \in C$ if $y \in C$. We obtain that, for every y,

$$s_C^*(y) \leq \delta_C(y).$$

Let us prove that equality holds when C is closed. We distinguish two cases: $y \notin C$ or $y \in C$.
(a) If $y \notin C$, then from the notes (page 23), there exist a hyperplane that separates the (bounded, convex) set $\{y\}$ from the closed convex set C. This means that there exists a vector x_0 such that $x_0^T y > 0$ and $x_0^T z < 0$ for every $z \in C$. Therefore,

$$\inf_{z \in C} x_0^T z \geq 0, \text{ thus } \inf_{z \in C} x_0^T (y - z) > 0.$$

With $x = \lambda x_0$, with $\lambda \to +\infty$, we obtain that

$$\sup_x \inf_{z \in C} x^T (y - z) = +\infty = \delta_C(y).$$

(b) If $y \in C$, then by choosing $z_0 = y \in C$ we obtain

$$\inf_{z \in C} x^T (y - z) \leq x^T (y - z_0) = 0$$

for *every* x, so that

$$\sup_x \inf_{z \in C} x^T (y - z) \leq 0 = \delta_C(y).$$

The upper bound is obviously attained (for $x = 0$), therefore equality holds.

3. *Subgradients and quasigradients.*

(a) Since f is differentiable, it has a unique subgradient at every point, given by the gradient

$$\nabla f(x) = 2Ax + b.$$

(b) For a given x, let i be such that $f_i(x) = \max(f_1(x), \ldots, f_n(x))$. Let $g_i(x)$ be a subgradient of f_i at x. We have for every y

$$f(y) \geq f_i(y) \geq f_i(x) + g_i(x)^T(y - x),$$

where the first inequality comes from the definition of pointwise maximum, and the second from the fact that $g_i(x)$ is a subgradient of f_i at x.

(c) Let (X, Y) be in $\mathbf{dom}\, f$. We start from the definition of

$$\lambda_{\max}(X, Y) = \lambda_{\max}(Y^{-1/2}XY^{-1/2}),$$

from which it is easily proven (using the results of problem 1, homework # 2), that

$$\lambda_{\max}(X, Y) = \sup_Z \left\{ \frac{\mathbf{Tr}\, XZ}{\mathbf{Tr}\, YZ} \mid Z \geq 0, \quad Z \neq 0 \right\}.$$

The supremum is achieved by the nonzero, positive semidefinite matrix $Z^* = Y^{-1/2}vv^TY^{-1/2}$, where v is a nonzero eigenvector of $Y^{-1/2}XY^{-1/2}$ corresponding to its largest eigenvalue.

From the results in previous problem, and also lecture 6 (page 132), we can guess that the quasigradient of f at (X, Y) is obtained by computing the quasigradients of the function

$$g(X, Y) = \frac{\mathbf{Tr}\, XZ}{\mathbf{Tr}\, YZ}$$

where $Z = Z^*$ is fixed. Remember: Z^* is actually a *function* of X, Y, but in the above it is considered fixed. The above function is differentiable on $\mathbf{dom}\, f$, and its derivatives with respect to X, Y are the two matrices

$$G_X = \frac{1}{\mathbf{Tr}\, YZ} Z, \quad G_Y = -\frac{\mathbf{Tr}\, XZ}{(\mathbf{Tr}\, YZ)^2} Z.$$

Let us prove that the pair (G_X, G_Y) is indeed the quasigradient of f at (X, Y). We must prove that, for every pair $(X', Y') \in \mathbf{dom}\, f$, we have

$$\mathbf{Tr}\, ((X' - X)G_X + (Y' - Y)G_Y) \geq 0 \text{ implies } \lambda_{\max}(X', Y') \geq \lambda_{\max}(X, Y).$$

A simple computation shows that

$$\begin{aligned} \mathbf{Tr}\, ((X' - X)G_X + (Y' - Y)G_Y) &= \frac{\mathbf{Tr}\, Y'Z}{\mathbf{Tr}\, YZ} \left(\frac{\mathbf{Tr}\, X'Z}{\mathbf{Tr}\, Y'Z} - \frac{\mathbf{Tr}\, XZ}{\mathbf{Tr}\, YZ} \right) \\ &= \frac{\mathbf{Tr}\, Y'Z}{\mathbf{Tr}\, YZ} \left(\frac{\mathbf{Tr}\, X'Z}{\mathbf{Tr}\, Y'Z} - \lambda_{\max}(X, Y) \right) \\ &\leq \frac{\mathbf{Tr}\, Y'Z}{\mathbf{Tr}\, YZ} (\lambda_{\max}(X', Y') - \lambda_{\max}(X, Y)). \end{aligned}$$

We have $\mathbf{Tr}\, Y'Z > 0$ (since $Z \succeq 0$, $Z \neq 0$, and $Y' \succ 0$). Similarly, $\mathbf{Tr}\, YZ > 0$. Therefore,

$$\mathbf{Tr}\, ((X' - X)G_X + (Y' - Y)G_Y) \geq 0$$

implies $\lambda_{\max}(X', Y') \geq \lambda_{\max}(X, Y)$.

4. Define $A(x) := A_0 + x_1 A_1 + \cdots + x_k A_k$ so that $A(x)_{ij}$, the (i,j) entry of $A(x)$, is equal to

$$A(x)_{ij} = a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}$$

where $a_{p,ij}$ is the (i,j) entry of A_p for $p = 1, \ldots . k$. Therefore, the problem of minimizing the max-row-sum

$$\text{minimize } \|A(x)\|_{\infty,\mathrm{i}} \tag{2}$$

over x can be expressed as

$$\text{minimize } \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}|. \tag{3}$$

This is a minimax problem and we use the standard trick of introducing a new (slack) variable $t \in \mathbf{R}$ to get the equivalent formulation

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \sum_{j=1}^{n} |a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}| \leq t, \quad i = 1, \ldots, m. \end{array} \tag{4}$$

It is easy to verify that the optimal solution of (3) and (4) are the same. In (4) since we are minimizing t, at least one of the constraints $\sum_{j=1}^{n} |a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}| \leq t$ become active at the optimal solution, because otherwise, we could simply reduce t to make the constraint with the largest $\sum_{j=1}^{n} |a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}|$ active, which contradicts the fact that t has been minimized. Therefore, at the optimal solution, $t = \max_{i=1,\ldots,m} \sum_{j=1}^{n} |a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}|$ so (3) and (4) are equivalent.

The formulation (4) is not a linear program yet. Somehow we need to get rid of the absolute values. This can be done by introducing additional variables s_{ij} for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ such that

$$|a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}| \leq s_{ij} \tag{5}$$

or

$$-s_{ij} \leq a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij} \leq s_{ij}. \tag{6}$$

Now it is easy to see that the optimization problem

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \sum_{j=1}^{n} s_{ij} \leq t, \quad i = 1, \ldots, m \\ & -s_{ij} \leq a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij} \leq s_{ij} \end{array} \tag{7}$$

is an equivalent formulation of (4) which is in fact a linear program in the variables t, $x_1, \ldots, x_k$ and s_{ij} for $i = 1, \ldots, m$ and $j = 1, \ldots, n$ (clearly, all constraints are in the form of linear inequalities in t, $x_1, \ldots, x_k$ and s_{ij}).

To prove that (7) is an equivalent formulation for minimizing the max-row-sum problem (2), first define $S \in \mathbf{R}^{m \times n}$ to have (i,j) entry equal to s_{ij}. The problem (7) can be re-written using compact matrix notation as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & S\mathbf{1}_n \preceq t\mathbf{1}_m \\ & -S \preceq A(x) \preceq S. \end{array}$$

The inequalities in $-S \preceq A(x) \preceq S$ are componentwise, *i.e.*, they are equivalent to $-s_{ij} \leq A(x)_{ij} \leq s_{ij}$ for $i = 1, \dots, m$, $j = 1, \dots, n$. The vectors $\mathbf{1}_n \in \mathbf{R}^n$ and $\mathbf{1}_m \in \mathbf{R}^m$ have all components one.

We first show that if x, S, t satisfy the constraints, then $t \geq \|A(x)\|_{\infty,\mathrm{i}}$.

The inequalities $-S \preceq A(x) \preceq S$ imply that $s_{ij} \geq |A(x)_{ij}|$, $i = 1, \dots, m$, $j = 1, \dots, n$. Therefore

$$(S\mathbf{1}_n)_i = \sum_{j=1}^{n} s_{ij} \geq \sum_{j=1}^{n} |A(x)_{ij}|,$$

$i = 1, \dots, m$. Combined with the second constraint, $S\mathbf{1}_n \leq t\mathbf{1}_m$, this means that $t \geq \|A(x)\|_{\infty,\mathrm{i}}$. Therefore we also have

$$t \geq \|A(x^\star)\|_{\infty,\mathrm{i}},$$

where $x^\star$ denotes any solution of the max-row-sum problem (2). In other words, the optimal value of the LP can never be less than $\|A(x^\star)\|_{\infty,\mathrm{i}}$.

Now it is also easy to verify that $x = x^\star$, $t = \|A(x^\star)\|_{\infty,\mathrm{i}}$, and $s_{ij} = |A(x^\star)_{ij}|$, $i = 1, \dots, m$, $j = 1, \dots, n$, are feasible in the LP, with objective value $\|A(x^\star)\|_{\infty,\mathrm{i}}$. Hence the optimal value of the LP cannot be greater that $\|A(x^\star)\|_{\infty,\mathrm{i}}$ and we are done.

Remarks. One might ask why not introduce the variables s_{ij} as $|a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}| = s_{ij}$ (with equality) instead of (5) (with inequality). The reason is simply that the constarint $|a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}| = s_{ij}$ is not convex in x and s_{ij} and therefore the problem cannot be cast as a convex or linear program this way. Also, you might think of rewriting the constraints $\sum_{j=1}^{n} |a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}| \leq t$ as linear inequality constraints by introducing 2^n vectors $c_l \in \mathbf{R}^n$ $(l = 1, \dots, 2^n)$ with elements that are either $+1$ or -1. In this case, these constraints can be written as

$$\sum_{j=1}^{n} c_{l,j}(a_{0,ij} + x_1 a_{1,ij} + \cdots + x_k a_{k,ij}) \leq t, \quad l = 1, \dots, 2^n, \quad i = 1, \dots, m$$

or

$$(\sum_{j=1}^{n} c_{l,j} a_{0,ij}) + (\sum_{j=1}^{n} c_{l,j} a_{1,ij}) x_1 + \cdots + (\sum_{j=1}^{n} c_{l,j} a_{k,ij}) x_k \leq t, \quad l = 1, \dots 2^n, \quad i = 1, \dots, m$$

which are linear constraints in $x_1, \dots, x_k$ and t. Although completely correct, this formulation has the drawback of having an "exponential" number of constraints. Even for relatively small n, 2^n can be very large and such a formulation might not be very useful in practice.