DSA3102: Practice Final Solutions

1. Consider the following Boolean linear program:

$$\min_x \; c^T x \quad \text{s.t.} \quad Ax \preceq b, \quad x_i(x_i - 1) = 0, \quad i = 1, \dots, n$$

where $c \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$ are given.

(a) (15 points) Show that the dual function of the above problem is

$$g(\lambda, \nu) = \begin{cases} -b^T\mu - \frac{1}{4}\sum_{i=1}^n \frac{(c_i + a_i^T\mu - \nu_i)^2}{\nu_i}, & \nu_i \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

where $\mu \in \mathbf{R}^m$ and $\nu \in \mathbf{R}^n$ are the Lagrange multipliers associated to the first and second sets of constraints respectively and a_i is the i-th column of A.

Solution: the Lagrangian is

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x + \lambda^T(Ax - b) + \nu^T x + x^T \operatorname{diag}(\nu) x \\ &= x^T \operatorname{diag}(\nu) x + (c + A^T\lambda - \nu)^T x - b^T\lambda \end{aligned}$$

This is because $\sum_i \nu_i x_i(x_i - 1)$ is exactly $\nu^T x + x^T \operatorname{diag}(\nu) x$. Minimizing over x gives the dual function $g(\lambda, \nu)$ since

$$x^* = -\frac{1}{2}\operatorname{diag}(\nu)^{-1}(c + A^T\lambda - \nu)$$

(b) (10 points) Show that a lower bound on the optimal value of the above problem can be obtained as the optimal value of the following problem

$$\max_\mu \; -b^T\mu + \sum_{i=1}^n \min\{0, c_i + a_i^T\mu\}, \quad \text{s.t.} \quad \mu \succeq 0$$

Solution: The resulting dual problem is

$$\max \; -b^T\lambda - \frac{1}{4}\sum_{i=1}^n (c_i + a_i^T\lambda - \nu_i)^2/\nu_i \quad \text{s.t.} \quad \nu \succeq 0$$

To simplify the dual, we optimize analytically over ν, by noting that

$$\sup_{\nu_i \geq 0} \left(-\frac{(c_i + a_i^T\lambda - \nu_i)^2}{\nu_i} \right) = \min\{0, c_i + a_i^T\lambda\}$$

This allows us to eliminate ν from the problem and simplify it as

$$\max \; -b^T + \sum_{i=1}^n \min\{0, c_i + a_i^T\lambda\} \quad \text{s.t.} \quad \lambda \succeq 0.$$

2. (a) (15 points) Solve for the optimal x using the KKT conditions:

$$\max_{x_1,x_2} \; 14x_1 - x_1^2 + 6x_2 - x_2^2 + 7 \quad \text{s.t.} \quad x_1 + x_2 \le 2 \quad x_1 + 2x_2 \le 3$$

Solution: The Lagrangian is

$$L(x,\mu) = -14x_1 + x_1^2 - 6x_2 + x_2^2 - 7 + \mu_1(x_1 + x_2 - 2) + \mu_2(x_1 + 2x_2 - 3)$$

The first order necessary conditions are

$$\begin{aligned} -14 + 2x_1 + \mu_1 + \mu_2 &= 0 \\ -6 + 2x_2 + \mu_1 + 2\mu_2 &= 0 \\ \mu_1 &\ge 0 \\ \mu_2 &\ge 0 \\ \mu_1(x_1 + x_2 - 2) &= 0 \\ \mu_2(x_1 + 2x_2 - 3) &= 0 \end{aligned}$$

We again try various combinations of active constraints.
Case 1: (inactive, inactive)

$$\begin{aligned} -14 + 2x_1 &= 0 \\ -6 + 2x_2 &= 0 \\ x_1 &= 7 \\ x_2 &= 3 \end{aligned}$$

However, $x_1 + x_2 = 10$ which is not less than 2, so this solution does not satisfy the constraints.
Case 2: (inactive, active)

$$\begin{aligned} -14 + 2x_1 + \mu_2 &= 0 \\ -6 + 2x_2 + \mu_2 &= 0 \\ x_1 + 2x_2 - 3 &= 0 \end{aligned}$$

Solving this system of equations yields $x_1 = 5$, $x_2 = -1$, $\mu_2 = 4$. However, $x_1 + x_2 = 4$ which is not less than 2, so this solution does not satisfy the constraints.
Case 3: (active, inactive)

$$\begin{aligned} -14 + 2x_1 + \mu_1 &= 0 \\ -6 + 2x_2 + \mu_1 &= 0 \\ x_1 + x_2 - 2 &= 0 \end{aligned}$$

Solving this system of equations yields $x_1 = 3$, $x_2 = -1$, $\mu_1 = 8$. We can easily check that this solution satisfies the first-order necessary conditions.
Case 3: (active, active)

$$\begin{aligned} -14 + 2x_1 + \mu_1 + \mu_2 &= 0 \\ -6 + 2x_2 + \mu_1 + \mu_2 &= 0 \\ x_1 + x_2 - 2 &= 0 \\ x_1 + 2x_2 - 3 &= 0 \end{aligned}$$

Solving this system of equations yields $x_1 = 1$, $x_2 = 1$, $\mu_1 = 20$, $\mu_2 = -8$. However, one of the Lagrange multipliers is negative, so this solution is not valid.
Thus only $x^* = (3, -1)$ and $\mu^* = (8, 0)$ satisfies the KKT conditions, Furthermore, since $\nabla^2 f(x) = [-1, 0; 0 - 1] < 0$, we have that f is strictly concave and since the constraint set is convex, x^* is the global maximum.

(b) (10 points) Consider the following problem

$$\max_x \ x_1^{a_1} x_2^{a_2} \dots x_n^{a_n} \quad \text{s.t.} \quad \sum_{i=1}^{n} x_i = 1 \quad x_i \geq 0, \quad i = 1, \dots, n$$

where a_i are given positive scalars. Find a global maximum and show that it is unique.

Solution: We can equivalently maximize $\sum_{i=1}^n a_i \log x_i$. Clearly none of the x_i will be 0. By first-order conditions,

$$\nabla f(x)^T d = \sum_i \frac{a_i}{x_i} d_i \leq 0$$

for all d such that

$$\sum_i d_i = 0, \sum_i x_i = 1, x_i \geq 0, i = 1, \dots, n$$

One way to satisfy the first inequality is to make it an equality. By comparison to the second equation, one way to satisfy the equality is if

$$\frac{a_1}{x_1} = \frac{a_2}{x_2} = \dots = \frac{a_n}{x_n}$$

The final two constraints are satisfied if we choose

$$x_i^* = \frac{a_i}{\sum_j a_j}$$

Since $f(x)$ is strictly concave, this x^* must be the unique global maximum

3. (a) (15 points) Formulate the following problem as an equivalent linear program

$$\min_x \ \|Ax - b\|_1 + \|x\|_\infty$$

where $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$ are given.

Solution: Equivalent LP

$$\min \ \mathbf{1}^T y + t, \quad \text{s.t.} \quad -y \leq Ax - b \leq y, -t\mathbf{1} \leq x \leq t\mathbf{1}$$

with variables $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$ and $t \in \mathbf{R}$.

(b) (10 points) Is the following set convex?

$$\mathcal{S} := \{x \in \mathbf{R}^2 : x_1 > 0, x_2 > 0, x_1 \log x_1 + x_2 \log x_2 \leq 2\}.$$

Solution: Yes. It is an intersection of convex sets. Indeed $f(x_1, x_2) = x_1 \log x_1 + x_2 \log x_2$ is a convex function on $\mathbf{R}^2_{++}$.

4. (a) (20 points) Consider the gradient descent method with bounded error,

$$x^{(k+1)} = x^{(k)} - s(\nabla f(x^{(k)}) + \epsilon^{(k)})$$

where s is a constant stepsize, $\epsilon^{(k)}$ are error terms satisfying

$$\|\epsilon^{(k)}\| \leq \delta$$

for all k, and f is the positive definite quadratic function

$$f(x) = \frac{1}{2}(x - x^*)^T Q (x - x^*)$$

Let

$$q := \max\{|1 - s\lambda_{\min}(Q)|, |1 - s\lambda_{\max}(Q)|\}$$

and assume that $q < 1$. Show that for all k, we have

$$\|x^{(k)} - x^*\| \le \frac{s\delta}{1-q} + q^k \|x^{(0)} - x^*\|$$

Solution: We have

$$\begin{aligned}
x^{k+1} &= x^k - s(Q(x^k - x^*) + \epsilon^k) \\
\|x^{k+1} - x^*\| &= \|x^k - x^* - s(Q(x^k - x^*) + \epsilon^k)\| \\
&= \|(I - sQ)(x^k - x^*) - s\epsilon^k\| \\
&\le \|(I - sQ)(x^k - x^*)\| + s\delta \\
&= \sqrt{(x^k - x^*)^T (I - sQ)^2 (x^k - x^*)} + s\delta \\
&\le \sqrt{\lambda_{\max}((I - sQ)^2)} \|x^k - x^*\| + s\delta \\
&= q\|x^k - x^*\| + s\delta
\end{aligned}$$

where $q = \max\{|1 - s\lambda_{\min}(Q)|, |1 - s\lambda_{\max}(Q)|\}$. Iterating this recursion,

$$\begin{aligned}
\|x^k - x^*\| &\le q^k \|x^0 - x^*\| + \sum_{i=0}^{k} q_j \cdot s\delta \\
&\le q^k \|x^0 - x^*\| + \sum_{i=0}^{\infty} q_j \cdot s\delta \\
&= \frac{s\delta}{1-q} + q^k \|x^0 - x^*\|
\end{aligned}$$

(b) (5 points) Consider the function f in part (a) of this problem. Suppose we now use Newton's method to find the minimum. Does Newton's method converge? If so how many iterations (Newton steps) are required for convergence to within $\varepsilon = 10^{-10}$ of x^*, i.e., find the minimum k such that $\|x^{(k)} - x^*\| \le \varepsilon$?

Solution: Yes and in 1 iteration because the function is quadratic.