EE 159/CS 168 - Convex Optimization
Scott Fullenbaum

Homework 6

---

1. $\nabla f(x^*) = Px^* + q = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix} \begin{bmatrix} 1 \\ 0.5 \\ -1 \end{bmatrix} + \begin{bmatrix} -22 \\ -14.5 \\ 13 \end{bmatrix} = \begin{bmatrix} 21 \\ 29/2 \\ -11 \end{bmatrix} + \begin{bmatrix} -22 \\ -29/2 \\ -11 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$

   Now we need to verify that the gradient satisfies the first order condition for optimality. Let $y \in \mathbb{R}^3$, show $\nabla f(x^*)^T(y - x^*) \geq 0$ for all $y \in \mathbf{dom} f$. Our equation becomes: $\begin{bmatrix} -1 & 0 & 2 \end{bmatrix}(y - \begin{bmatrix} 1 \\ 1/2 \\ -1 \end{bmatrix}) = \begin{bmatrix} -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 - 1 \\ y_2 - 1/2 \\ y_3 + 1 \end{bmatrix} = -(y_1 - 1) + 2(y_3 + 1) \geq 0$. For our feasible set $\mathcal{F}$, $|y_i| \leq 1$ for all i. As $y_3 + 1 \geq 0$ for $y_3 \in \mathcal{F}$ and it is the same with $-(y_1 - 1)$, then f satisfies the optimality condition and $x^* = (1, 1/2, -1)$ is optimal.

2. First, show the equations in parts a and b are equivalent. For part b, define the function being minimized as $f(w)$, hold x constant and $u_i = (a_i^T x - b_i)$.Then, it follows that $\nabla f(w)$ $= \begin{bmatrix} -u_1^2/(w_1+1)^2 + M^2 \\ \cdots\cdots \\ -u_m^2(w_m+1)^2 + M^2 \end{bmatrix}$ Then set $\nabla f(w) = 0$ and for any i, $M^2 - u_i^2/(w_i+1)^2 = 0 \rightarrow$ $(|u_i|/M - 1) = w_i$, which is our global optimal solution. However, as $w \succeq 0$, then we can say for $w_i = \left\{ \begin{array}{cc} |u_i|/M - 1 & |u_i| > M \\ 0 & |u_i| \leq M \end{array} \right\}$

   To get the optimal solution, we can plug each w_i into an individual term in the summation, or $g(w_i) = (a_i^T x - b_i)^2/(w_i + 1) + M^2 w_i$. If $w_i = 0, g(0) = (a_i^T x - b_i)^2$. If $w_i = |a_i^T x - b_i|/M - 1$, this equals $(a_i^T x - b_i)^2/(|a_i^T x - b_i|/M) + M^2((a_i^T x - b_i/M) - 1)$ $= M|a_i^T x - b_i| + M|a_i^T x - b_i| - M^2 = M(2|a_i^T x - b_i| - M)$. So, we have:

   $$g(a_i^T x - b_i) = \left\{ \begin{array}{cc} M(2|a_i^T x - b_i| - M) & |a_i^T x - b_i| > M \\ (a_i^T x - b_i)^2 & |a_i^T x - b_i| \leq M \end{array} \right\}$$

   At this point, we can see $g(a_i^T x - b_i)$ is equivalent to the Huber loss function in part a, so minimizing the equation in part b is equivalent to solving the equation in part a.

   Next, let's show parts a and c are equivalent. First, at the optimal solution, $u + v = |Ax - b|$. We can make this assumption because if at the optimal solution, $u + v > |Ax - b|$ with $0 \preceq u \preceq M$ and $v \succeq 0$ if both are nonzero, we can decrease the values in u and v without violating the constraints. As our objective function is a summation, the objective function will decrease. Therefore, $u + v = |Ax - b|$ Next, we can rewrite $v = |Ax - b| - u$ or $v_i = |a_i^T x - b_i| - u_i$. We can transform the original equation and look at the constraints by looking at individual u_i, Taking into account that $v_i \geq 0$ leads to the reformulation:

$$\begin{aligned} &\text{minimize} \sum_{i=1}^{m} u_i^2 - 2Mu_i + 2M|a_i^T x - b_i| \\ &\text{subject to } 0 \leq u_i \leq \min\{M, |a_i^T x - b_i|\} \end{aligned}$$

If $|a_i^T x - b_i| \leq M$, then as we are taking a sum, the optimal value is $u_i = |a_i^T x - b_i|$ and then the objective function becomes $|a_i^T x - b_i|^2$. If $M \leq |a_i^T x - b_i|$ then the optimal value is $u_i = M$, and the value at the i-th term is $-M^2 + 2M|a_i^T x - b_i| = M(2|a_i^T x - b_i| - M)$. From here, we can see how the solution to this minimization problem is equivalent to minimizing the Huber loss function in part a, which shows these problems are equivalent.

3. Since we want to minimize total cost, the objective function is $C = \sum_{i,j=1}^{n} c_{ij}x_{ij}$. For each x_{ij} there is the constraint $l_{ij} \leq x_{ij} \leq u_{ij}$. Additionally, as the flow in and out of any node must be 0, then $b_i + \sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = 0$. So our problem is:

$$\min C = \sum_{i,j=1}^{n} c_{ij}x_{ij}$$

$$\text{subject to: } b_i + \sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = 0 \quad i = 1, ..., n$$

$$l_{ij} \leq x_{ij} \leq u_{ij} \quad i, j = 1........, n$$

4. We can express this general problem as:

$$\begin{aligned} &\text{maximize } \alpha_4 T r^2 \\ &\text{subject to: } \alpha_1 T w^{-1} + \alpha_2 r + \alpha_3 r w \leq C_{max} \\ &T_{min} \leq T \leq T_{max} \\ &r_{min} \leq r \leq r_{max} \\ &w_{min} \leq w \leq w_{max} \end{aligned}$$

To reformulate as a geometric program, maximizing the objective function is equivalent to minimizing it's inverse. To reformulate the constraints, all constraints must be rewritten as posynomials and that they are ≤ 1. This leads to the following problem reformulation:

$$\begin{aligned} &\text{minimize } (1/\alpha_4) T^{-1} r^{-2} \\ &\text{subject to: } (1/C_{max})(\alpha_4 T w^{-1} + \alpha_2 r + \alpha_3 r w) \leq 1 \\ &T/T_{max} \leq 1, \quad T_{min}/T \leq 1 \\ &r/r_{max} \leq 1, \quad r_{min}/r \leq 1 \\ &w/w_{max} \leq 1, \quad w_{min}/w \leq 1 \\ &10w/r \leq 1 \end{aligned}$$

This is a geometric program, and all constraints are posynomials where the variables are T, r, and w.

6. (a) As $\lambda_1(x)$ is the max eigenvalue, $A(x) \preceq \lambda_1(x)I$ and therefore $A(x) - \lambda_1(x) \preceq 0$. There exists some $u \in \mathbb{R}$ such that $A(x) \preceq \lambda_1(x)I \preceq uI$. So, our problem can be thought of as minimizing u on these bounds. This leads to the following problem:

$$\begin{gathered}\text{minimize } u \\ \text{subject to } A(x) - uI \preceq 0\end{gathered}$$

This is in the form of a SDP, concluding this problem.

(b) Using variables u_1 and u_2 and the equation. Let u_1 and $u_2 \in \mathbb{R}$. If $\lambda_1(x) \leq u_1$ then $A(x) \preceq u_1 I$. We can do the reverse for u_2 and state that if $u_2 \leq \lambda_m(x)$, then $A(x) \succeq u_2 I$. This gives the following problem:

$$\begin{gathered}\text{minimize } u_1 - u_2 \\ \text{subject to } u_2 I \preceq A(x) \preceq u_1 I\end{gathered}$$

This can be reformulated as:

$$\begin{gathered}\text{minimize } u_1 - u_2 \\ \text{subject to } u_2 I - A(x) \preceq 0 \\ A(x) - u_1 I \preceq 0\end{gathered}$$

Which can be combined and written as one SDP.

(c) Since A is positive definite, it follows that A^{-1} is also positive definite. Because of this, we can say there is some $u \in \mathbb{R}$ such that $c^T A^{-1} c \preceq uI$ or $c^T A^{-1} c - uI \preceq 0$. So our problem is minimizing u, which can be written out as:

$$\begin{gathered}\text{minimize u} \\ \text{subject to } c^T A(x)^{-1} c - uI \preceq 0\end{gathered}$$

This can be combined into a block matrix which is then a SDP.

(d) Part d is similar to the previous part. However, as there are multiple c_i, we now have a system of constraints. So, using the same logic in part c, all c_i must satisfy the constraint for our minimizing variable u.

$$\begin{gathered}\text{minimize u} \\ \text{subject to } c_i^T A(x)^{-1} c_i - uI \preceq 0 \; i = 1, ..., K\end{gathered}$$