CS5321 Numerical Optimization Homework 6

Due May 24

1. (20%) (Farkas's Lemma) Let A be an $m \times n$ matrix and $\vec{b}$ an m vector. Prove that exact one of the following two statements is true:

   (a) There exists a $\vec{v} \in \mathbb{R}^n$ such that $A\vec{v} = \vec{b}$ and $\vec{v} \geq 0$.

   (b) There exists a $\vec{u} \in \mathbb{R}^m$ such that $A^T\vec{u} \geq 0$ and $\vec{b}^T\vec{u} < 0$.

   (Hint: prove if (a) is true, then (b) cannot be true, and vice versa.)

   Suppose (a) is true. If (b) is also true, then $\vec{u}^T A\vec{v} = \vec{u}^T\vec{b}$. Since $\vec{v} \geq 0$ and $A^T\vec{u} \geq 0$, $\vec{u}^T\vec{b} = \vec{b}^T\vec{u} \geq 0$, which contradicts with the statement in (b).

   Suppose (b) is true. If (a) is true,

   $$\vec{b}^T\vec{u} = \vec{v}^T A^T\vec{u} < 0.$$

   However, because $\vec{v} \geq 0$ and $A^T\vec{u} \geq 0$, their product cannot be negative.

   Those two cases have their geometric interpretations. For (a), $A\vec{v}$ for $\vec{v} \geq 0$ forms a *cone*. $A\vec{v} = \vec{b}$ just means $\vec{b}$ is a vector in the cone. For (b), $A^T\vec{u} \geq 0$ means the angles of $\vec{u}$ and the column vectors of A are less than 90 degree; and the angle of $\vec{u}$ and $\vec{b}$ is larger than 90 degree. If $\vec{u}$ is a normal vector of a hyperplane $\mathcal{P}$, then $\mathcal{P}$ separates $\vec{b}$ from the column vectors of A.

   Farkas lemma just says either $\vec{b}$ is inside the cone $A\vec{v}$ or outside the cone. One of them must be true, but cannot be both.

Farkas's Lemma (1902) plays an important role in the proof of the KKT condition. The most critical part in the proof of the KKT condition is to show that the Lagrange multiplier $\vec{\lambda}^* \geq 0$ for inequality constraints. We can say if the LICQ condition is satisfied at $\vec{x}^*$, then any feasible direction $\vec{y}$ at $\vec{x}^*$ must have the following properties:

- $\vec{y}^T \nabla f(\vec{x}^*) \geq 0$, since $\vec{x}^*$ is a local minimizer. (Otherwise, we find a feasible descent direction that decreases f.)
- $\vec{y}^T \nabla c_i(\vec{x}^*) = 0$ for equality constraints, $c_i = 0$.
- $\vec{y}^T \nabla c_i(\vec{x}^*) \geq 0$ for inequality constraints, $c_i \geq 0$.

Here is how Farkas Lemma enters the theme. Let $\vec{b}$ be $\nabla f(\vec{x}^*)$, $\vec{u}$ be $\vec{y}$ (any feasible direction at $\vec{x}^*$), the columns of A be $\nabla c_i(\vec{x}^*)$. Since no such $\vec{u}$ exists, according to the properties of $\vec{y}$, statement (a) must hold. The vector $\vec{v}$ in (a) corresponds to $\vec{\lambda}^*$, which just gives us the desired result of the KKT condition.

2. (50%) Consider the following constrained minimization problem

$$\min_{x_1, x_2} -x_1 + x_2^2 \text{ subject to } \begin{cases} (1-x_1)^3 - x_2 & \geq 0 \\ x_1 + x_2 - 1 & \geq 0 \end{cases}$$

(a) Plot the feasible region of the problem, and use it to find the optimal solution $\vec{x}^*$.

$\vec{x}^* = (0, 1)^T$.

(b) Write its Lagrangian function $\mathcal{L}(\vec{x}, \vec{\lambda})$, and use KKT condition to compute $\vec{\lambda}^*$.

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = -x_1 + x_2^2 - \lambda_1((1-x_1)^3 - x_2) - \lambda_2(x_1 + x_2 - 1).$$

At the optimal solution, $\nabla_x \mathcal{L}(\vec{x}^*, \vec{\lambda}) = 0$.

$$\nabla_x \mathcal{L}(\vec{x}, \vec{\lambda}) = \begin{pmatrix} -1 + 3\lambda_1(1-x_1)^2 - \lambda_2 \\ 2x_2 + \lambda_1 - \lambda_2 \end{pmatrix}.$$

To make $\nabla_x \mathcal{L}((0,1)^T, \vec{\lambda}) = 0$,

$$\nabla_x \mathcal{L}(\vec{x}, \vec{\lambda}) = \begin{pmatrix} -1 + 3\lambda_1 - \lambda_2 \\ 2 + \lambda_1 - \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The solution is $\lambda_1^* = 3/2, \lambda_2^* = 7/2$.

(c) Verify the LICQ condition at $\vec{x}^*$.

Let $c_1(\vec{x}) = (1 - x_1)^3 - x_2$ and $c_2(\vec{x}) = x_1 + x_2 - 1$.

$$\nabla c_1(\vec{x}) = \begin{pmatrix} -3(1-x_1)^2 \\ -1 \end{pmatrix}, \nabla c_2(\vec{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Let $A = [\nabla c_1(\vec{x}^*)\ \nabla c_2(\vec{x}^*)] = \begin{pmatrix} -3 & 1 \\ -1 & 1 \end{pmatrix}$.

Since $\det(A) = -2 \neq 0$, $\nabla c_1(\vec{x}^*)$ and $\nabla c_2(\vec{x}^*)$ are linearly independent. The LICQ holds.

(d) Verify the KKT condition at $\vec{x}^*$.

- The first condition $\mathcal{L}(\vec{x}^*, \vec{\lambda}^*)$ holds if we choose $\vec{\lambda}^* = (3/2, 7/2)$.
- There is no equality constraints.
- The inequality constraints, $c_1(\vec{x}^*) = (1-0)^3 - 1 = 0$ and $c_2(\vec{x}^*) = 0 + 1 - 1 = 0$ are also satisfied.
- $\lambda_1^* = 3/2 > 0$ and $\lambda_2^* = 7/2 > 0$.
- The complementarity condition, $\lambda_1^* c_1(\vec{x}^*) = 3/2 * 0 = 0$, and $\lambda_2^* c_2(\vec{x}^*) = 7/2 * 0 = 0$ are also true.

(e) Compute the Lagrangian Hessian at $\vec{x}^*$ and the critical cone, and verify the second order optimality condition.

The Lagrangian Hessian is

$$\nabla_{xx}\mathcal{L}(\vec{x}, \vec{\lambda}) = \begin{pmatrix} -6\lambda_1(1-x_1) & 0 \\ 0 & 2 \end{pmatrix}.$$

which equals to $\begin{pmatrix} -9 & 0 \\ 0 & 2 \end{pmatrix}$ at $\vec{x}^* = (0, 1)$ and $\vec{\lambda}^* = (3/2, 7/2)^T$.

Since both constraints are active, the critical cone is

$$\{\vec{w} | \vec{w}^T \nabla c_1(\vec{x}^*) = 0 \text{ and } \vec{w}^T \nabla c_2(\vec{x}^*) = 0\}.$$

The only vector in the cone is $\vec{w} = (0, 0)^T$.

$\vec{w}^T \nabla_{xx}\mathcal{L}(\vec{x}, \vec{\lambda})\vec{w} = 0$.

3. (30%) Consider the following problem

$$\min_{x_1, x_2} \frac{1}{2}\alpha x_1^2 + \frac{1}{2}x_2^2 + x_1 \text{ subject to } x_1 \geq 1.$$

Determine the solution to this problem for $\alpha = 1$ and $\alpha = 0$. For each case, formulate the dual, and determine whether the primal and the dual have the same optimal solution.

Case I

$$\min_{x_1,x_2} z_1 = \frac{1}{2}x_2^2 + x_1 \text{ subject to } x_1 \geq 1.$$

The optimal solution is 1 at $(1, 0)$. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}x_2^2 + x_1 - \lambda(x_1 - 1) = \frac{1}{2}x_2^2 + (1-\lambda)x_1 + \lambda.$$

The objective function of the dual problem is obtained by

$$\min_{x_1,x_2} \mathcal{L}(\lambda) = \begin{cases} -\infty, & \text{if } \lambda > 1; \\ \lambda, & \text{if } \lambda \leq 1. \end{cases}$$

The dual problem becomes

$$\max_{\lambda} \lambda \text{ subject to } 0 \leq \lambda \leq 1,$$

whose solution is 1.

Case II

$$\min_{x_1,x_2} z_2 = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + x_1 \text{ subject to } x_1 \geq 1.$$

The optimal solution is $3/2$ at $(1, 0)$. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + x_1 - \lambda(x_1 - 1).$$

The objective function of the dual problem is obtained by

$$\begin{aligned} \min_{x_1,x_2} \mathcal{L}(\lambda) &= \frac{1}{2}x_1^2 + (1-\lambda)x_1 + \lambda \\ &= \frac{1}{2}(x_1 + (1-\lambda))^2 - \frac{1}{2}(1-\lambda)^2 + \lambda \\ &= -\frac{1}{2}\lambda^2 + 2\lambda - \frac{1}{2} \end{aligned}$$

The dual problem is

$$\max_{\lambda} -\frac{1}{2}\lambda^2 + 2\lambda - \frac{1}{2}. \text{ subject to } \lambda \geq 0,$$

whose optimal solution is at $\lambda = 2$ and optimal objective function value is $3/2$, equal to the primal's.