CS5321 Numerical Optimization Homework 5

Due May 9

1. (30%) Use the optimality conditions of constrained optimization problems to verify the following properties.

   (a) The optimal solution of the total least square problem is $A^T A\vec{x} = \lambda\vec{x}$ for some λ.

   The total least square problem is

   $$\begin{array}{ll} \min_{\vec{x}} & \vec{x}^T A^T A\vec{x} \\ \text{s.t.} & \vec{x}^T\vec{x} = 1 \end{array}$$

   The Lagrangian function is

   $$\mathcal{L} = \vec{x}^T A^T A\vec{x} - \lambda(\vec{x}^T\vec{x} - 1).$$

   The KKT condition shows

   $$\nabla_x \mathcal{L} = 2A^T A\vec{x} - 2\lambda\vec{x} = 0$$

   (b) For the trust region method, the optimal solution $\vec{p}^*$ of the local model

   $$\min_{\vec{p}\in\mathbb{R}^n} m(\vec{p}) = \vec{g}^T\vec{p} + \frac{1}{2}\vec{p}^T A\vec{p} \ \text{ s.t. } \ \vec{p}^T\vec{p} \le \Delta^2,$$

   satisfies

   $(A+\lambda I)\vec{p}^* = -\vec{g}$, $\lambda(\Delta - \|\vec{p}^*\|) = 0$, and $(A+\lambda I)$ is positive semi-definite.

The Lagrangian function for the trust region method is

$$\mathcal{L} = \vec{g}^T\vec{p} + \frac{1}{2}\vec{p}^T A\vec{p} - \mu(\Delta^2 - \vec{p}^T\vec{p}).$$

The KKT condition shows

$$\nabla_x \mathcal{L} = \vec{g} + A\vec{p} + 2\mu\vec{p} = 0.$$

Let $\lambda = 2\mu$.

$$(A + \lambda I)\vec{p}^* = -\vec{g}.$$

The complementarity condition shows

$$\lambda(\Delta^2 - \|\vec{p}^*\|^2) = 0.$$

Either $\lambda = 0$ or $\Delta^2 = \|\vec{p}^*\|^2$, one can show that $\lambda(\Delta - \|\vec{p}^*\|) = 0$.

If $\lambda = 0$, the constrain is inactive, $\|\vec{p}^*\| < \Delta$. It is an unconstrained optimization problem. So by the second order condition of an unconstrained optimization problem, $B + \lambda I = B$ is positive semi-definite.

If $\lambda > 0$, by the complementarity condition, $\|\vec{p}^*\| = \Delta$. Thus, we only need to consider the position $\vec{p}$ such that $\|\vec{p}\| = \Delta$ (critical cone).

Since $\vec{p}^*$ is the minimizer, $m(\vec{p}^*) \leq m(\vec{p})$, which implies

$$\vec{g}^T\vec{p}^* + \frac{1}{2}(\vec{p}^*)^T A\vec{p}^* \quad \leq \quad \vec{g}^T\vec{p} + \frac{1}{2}\vec{p}^T A\vec{p} \tag{1}$$

We use the first condition: $\vec{g} = -(A+\lambda I)\vec{p}^*$ to obtain $\vec{g}^T\vec{p}^* = -(\vec{p}^*)^T(A+\lambda I)\vec{p}^*$ and $\vec{g}^T\vec{p} = -\vec{p}^T(A+\lambda I)\vec{p}^*$. Substituting them to (1), one has

$$-(\vec{p}^*)^T(A+\lambda I)\vec{p}^* + \frac{1}{2}(\vec{p}^*)^T A\vec{p}^* \leq -\vec{p}^T(A+\lambda I)\vec{p}^* + \frac{1}{2}\vec{p}^T A\vec{p}$$

Since $\|\vec{p}^*\|^2 = \|\vec{p}\|^2 = \Delta^2$, we add $\frac{1}{2}\lambda\Delta^2$ to both sides, and get

$$-(\vec{p}^*)^T(A+\lambda I)\vec{p}^* + \frac{1}{2}(\vec{p}^*)^T(A+\lambda I)\vec{p}^* \leq -\vec{p}^T(A+\lambda I)\vec{p}^* + \frac{1}{2}\vec{p}^T(A+\lambda I)\vec{p}$$

which is equivalent to

$$0 \leq (\vec{p}^* - \vec{p})^T(A+\lambda I)(\vec{p}^* - \vec{p}).$$

Since the only constraint of $\vec{p}$ is $\|\vec{p}\| = \Delta$, $(\vec{p}^* - \vec{p})$ can be any vector. Thus, $(A+\lambda I)$ is positive semi-definite.

2. (30%) For a quadratic programming,

$$\min_{\vec{x}} g(\vec{x}) = \frac{1}{2}\vec{x}^T G\vec{x} + \vec{x}^T\vec{c}$$

$$\text{s.t. } A\vec{x} = \vec{b},$$

Prove that if A has full row-rank and the reduced Hessian Z^TGZ is positive definite, where span$\{Z\}$ is the null space of span$\{A^T\}$, then the KKT matrix $K = \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}$ is nonsingular. (Hint: Prove that every vector $\begin{bmatrix} \vec{w} \\ \vec{v} \end{bmatrix}$ making $\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} \vec{w} \\ \vec{v} \end{bmatrix} = 0$ is a zero vector. Using the property that $\vec{w}^T G\vec{w} > 0$.)

Suppose $\begin{bmatrix} \vec{w} \\ \vec{v} \end{bmatrix}$ making $\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \vec{w} \\ \vec{v} \end{bmatrix} = 0$. which means

$$G\vec{w} + A^T\vec{v} = 0 \quad (2)$$
$$A\vec{w} = 0. \quad (3)$$

Pre-multiply $\vec{w}^T$ to (2), one has $\vec{w}^T G\vec{w} + \vec{w}^T A^T\vec{v} = \vec{w}^T G\vec{w} + 0^T\vec{v} = \vec{w}^T G\vec{w} = 0$. Since $A\vec{w} = 0$, $\vec{w}$ is in A's null space. $\vec{w} = Z\vec{u}$.

$$\vec{w}^T G\vec{w} = \vec{u}^T Z^T GZ\vec{u} = 0$$

and $Z^T GZ$ is positive definite implies $\vec{u} = 0$, so is $\vec{w}$.

Since $\vec{w} = 0$, from (2), $A^T\vec{v} = 0$. But A is of full row rank. Therefore, $\vec{v}$ must be a zero vector.

Thus, matrix K is nonsingular.

3. (40%) Consider the quadratic programming problem with bounded constraints

$$\min_{x_1,x_2,x_3} (x_1 - 4)^2 + (x_2 - 3)^2 + (x_3 - 2)^2$$

$$\text{s.t.} \quad 0 \le x_1, x_2, x_3 \le 2$$

Use gradient projection method to find its optimal solution with $\vec{x}_0 = 0$. Write down the trace, like

$$\vec{x}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \to \vec{x}_1 = \begin{pmatrix} 2 \\ 3/2 \\ 1 \end{pmatrix} \to \vec{x}_2 = \cdots$$