CS5321 Numerical Optimization Homework 4

Due April 25

1. (10%) What is the distance of a point $\vec{p}$ to a hyperplane $\vec{a}^T\vec{x} + b = 0$. Justify your answer.

   Assume $\vec{a} \neq 0$. First, let's consider the case that $b = 0$. In such case, we split the vector $\vec{p}$ into two vectors: one parallel to the normal vector of the hyperplane $\vec{a}$, and another perpendicular to $\vec{a}$,

   $$\vec{p} = \cos\angle(\vec{p}, \vec{a})\vec{p} + \sin\angle(\vec{p}, \vec{a})\vec{p}.$$

   The distance from $\vec{p}$ to the hyperplane is $\|\cos\angle(\vec{p}, \vec{a})\vec{p}\| = |\vec{a}^T\vec{p}|/\|\vec{a}\|$. For the case that $b \neq 0$ is just adding a shift to the distance. Therefore, it is

   $$|\vec{a}^T\vec{p} + b|/\|\vec{a}\|.$$

   For the case $\vec{a} = 0$, b must be 0 too. In this case, the distance is the distance to the origin $\|\vec{p}\|$.

2. (40%) Our frequently used matrix norms are called *subordinate matrix norm* because they are derived from corresponding vector norms. For an $n \times m$ matrix A, its 1-norm, 2-norm and infinite-norm are defined by

   $$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p,$$

   where $p = 1, 2, \infty$ respectively.

   (a) What is the matrix 1-norm? Justify your answer?

   Vector 1-norm of $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ is $\|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i|$.
   Let $A = (\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)$, where $\vec{a}_i$ is the ith column vector of A.
   Given an $\vec{x}$,

   $$\begin{aligned} A\vec{x} &= \textstyle\sum_{i=1}^n x_i\vec{a}_i \\ \|A\vec{x}\|_1 &= \|\textstyle\sum_{i=1}^n x_i\vec{a}_i\|_1 \\ &\leq \textstyle\sum_{i=1}^n |x_i|\|\vec{a}_i\|_1 && \text{(Triangle inequality)} \\ &\leq (\max_{i=1,..,n}\|\vec{a}_i\|_1)\textstyle\sum_{i=1}^n |x_i| && \text{(Find a largest } \|\vec{a}_i\|_1) \\ &= (\max_{i=1,..,n}\|\vec{a}_i\|_1)\|\vec{x}\|_1 \end{aligned}$$

   Therefore, $\|A\|_1 = \max_{i=1,..,n}\|\vec{a}_i\|_1$.

(b) What is the matrix ∞-norm? Justify your answer?

Vector ∞-norm of $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ is $\|\vec{x}\|_\infty = \max_{i=1,..,n} |x_i|$.

Let $A = \left(\vec{a}_1^T, \vec{a_2}^T, \ldots, \vec{a}_m^T\right)^T$, where $\vec{a}_i$ is the ith row vector of A.

$$A\vec{x} = \begin{pmatrix} \vec{a}_1^T\vec{x} \\ \vec{a}_2^T\vec{x} \\ \vdots \\ \vec{a}_m^T\vec{x} \end{pmatrix}$$

Given an $\vec{x}$,

$$\begin{aligned} \|A\vec{x}\|_\infty &= \max_{i=1,..,m} |\vec{a}_i^T\vec{x}| \\ &= \max_{i=1,..,m} \left|\textstyle\sum_{j=1}^n a_{i,j}x_j\right| \\ &\leq \max_{i=1,..,m} \textstyle\sum_{j=1}^n |a_{i,j}x_j| \\ &\leq \max_{i=1,..,m} \left(\textstyle\sum_{j=1}^n |a_{i,j}|\right) \max_{j=1,..,n} |x_j| \\ &= \max_{i=1,..,m} \|\vec{a}_i\|_1 \|\vec{x}\|_\infty \end{aligned}$$

Therefore, $\|A\|_1 = \max_{i=1,..,n} \|\vec{a}_i\|_1$.
(Note the meaning of $\vec{a}_i$ here is different from that is 2(a).)

(c) What is the matrix 2-norm? Justify your answer?

Vector 2-norm of $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ is $\|\vec{x}\|_\infty = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\vec{x}^T\vec{x}}$.
Let $A^TA = Q\Lambda Q^{-1}$ be the eigenvalue decomposition of A^TA, where Λ is diagonal with elements $\lambda_1, \lambda_2, \ldots, \lambda_n$.
Because A^TA is symmetric, one can make Q orthogonal, $Q^{-1} = Q^T$. In addition, A^TA is positive semidefinite, all λ_i are non-negative. Given an A and an $\vec{x}$, let $\vec{y} = Q^T\vec{x}$.

$$\begin{aligned} \|A\vec{x}\|_2^2 &= \vec{x}^TA^TA\vec{x} \\ &= \vec{x}^TQ\Lambda Q^T\vec{x} \\ &= \vec{y}^T\Lambda\vec{y} \\ &= \sum_{i=1}^n \lambda_i y_i^2 \\ &\leq \left(\max_{i=1,..n} \lambda_i\right) \sum_{i=1}^n y_i^2 \\ &\leq \left(\max_{i=1,..n} \lambda_i\right) \|\vec{y}\|_2^2 \end{aligned}$$

Because Q is orthogonal, $\|\vec{y}\|_2 = \|Q^T\vec{x}\|_2 = \|\vec{x}\|_2$. Therefore, $\|A\|_2 =$ the largest eigenvalue of A^TA, which is also the largest singular value of A.

(d) Show the condition number of an invertible matrix A, $\kappa(A)$, equations to σ_1/σ_n, where σ_1 is the largest singular value of A and σ_n is the smallest singular value of A.

Here we use 2-norm to define the condition number of a invertible matrix A: $\kappa(A) = \|A\|_2 \|A^{-1}\|_2$. Since A is invertible, all its singular values are nonzero. Also, if $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ are the singular values of A, $\sigma_n^{-1} \geq \sigma_{n-1}^{-1} \geq \ldots \geq \sigma_1^{-1} \geq 0$ are the singular values of A^{-1}. Therefore, $\kappa(A) = \sigma_1 \sigma_n^{-1}$.

3. (50%) Consider the following linear program:

$$\begin{array}{ll} \max_{x_1,x_2} & z = 8x_1 + 5x_2 \\ \text{s.t.} & 2x_1 + x_2 \leq 1000 \\ & 3x_1 + 4x_2 \leq 2400 \\ & x_1 + x_2 \leq 700 \\ & x_1 - x_2 \leq 350 \\ & x_1, x_2 \geq 0 \end{array}$$

(a) Transform it the standard form.

(b) Suppose the initial guess is $(0, 0)$. Use the simplex method to solve this problem. In each iterations, show

- Basic variables and non-basic variables, and their values.
- Pricing vector.
- Search direction.
- Ratio test result.