CS532100 Numerical Optimization Homework 2

Student ID name

Due Dec 10

1. Consider the linear least square problem:

$$min_{\vec{x}\in\mathbb{R}^2}||A\vec{x}-\vec{b}||^2,$$

where

$$A=\begin{bmatrix}4 & 8\\ 2 & 4\\ 1 & 2\end{bmatrix}, \vec{b}=\begin{pmatrix}21/4\\ 0\\ 0\end{pmatrix}$$

   (a) (10%) Write its normal equation.
   Answers are put here.

   (b) (10%) Express $\vec{b}=\vec{b}_1+\vec{b}_2$ such that $\vec{b}_1$ is in the subspace spanned by A's column vectors, and $\vec{b}_2$ is orthogonal to A's column vectors.
   Answers are put here.

   (c) (10%) Show that $\vec{z}\in\mathbb{R}^2$ is a least square solution for $A\vec{x}=\vec{b}$ if and only if $\vec{z}$ is part of a solution to the larger linear system:

$$\begin{bmatrix}0 & A^T\\ A & I\end{bmatrix}\begin{bmatrix}\vec{z}\\ \vec{y}\end{bmatrix}=\begin{bmatrix}0\\ \vec{b}\end{bmatrix}$$

   Answers are put here.

2. In Note05 (Page 16), memoryless BFGS iteration matrix H_{k+1} can be derived from considering the Hestenes–Stiefel form of the nonlinear conjugate gradient method. Recalling that $\vec{s}_k=\alpha_k\vec{p}_k$, we have that the search direction for this method is given by

$$\vec{p}_{k+1}=-\nabla f_{k+1}+\frac{\nabla f_{k+1}^T\vec{y}_k}{\vec{y}^T\vec{p}_k}\vec{p}_k$$

$$=-\nabla f_{k+1}+\frac{\nabla f_{k+1}^T\vec{y}_k}{\vec{y}^T\vec{s}_k}\vec{s}_k$$

$$=-(I-\frac{\vec{s}_k\vec{y}_k^T}{\vec{y}^T\vec{s}_k})\nabla f_{k+1}$$

$$=-\hat{H}_{k+1}\nabla f_{k+1}$$

However, the matrix $\hat{H}_{k+1}$ is neither symmetric nor positive definite.

(a) (10%) Please show that the matrix $\hat{H}_{k+1}$ is singular.
(You can only prove it for the case $\nabla f_k, \vec{p}_k, \vec{y}_k, \vec{s}_k \in \mathbb{R}^2$ for all $k \in \mathbb{N}$.)

Answers are put here.

(b) (0%) Please read the reference book (Page 180) to understand the derivation of the inverse hessian formula in Note05 (Page 16).
(you don't need to write anything in this subproblem.)

$$H_{k+1} = (I - \frac{\vec{s}_k \vec{y}_k^T}{\vec{y}_k^T \vec{s}_k})(I - \frac{\vec{y}_k \vec{s}_k^T}{\vec{y}_k^T \vec{s}_k}) + \frac{\vec{s}_k \vec{s}_k^T}{\vec{y}_k^T \vec{s}_k}$$

3. (10%) The total least square problem is to solve the following problem

$$\min_{\vec{x}, \|\vec{x}\|=1} \vec{x}^T A^T A \vec{x}$$

where A is an $m \times n$ matrix. Here we assume $m > n$. Let $A = U\Sigma V^T$ be the SVD of A, where U is the matrix of left singular vectors, V is the matrix of right singular vectors, and Σ is a diagonal matrix with diagonal elements $\sigma_1, \sigma_2, \ldots, \sigma_n$. Moreover, U and V are orthogonal matrices, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. Show the solution of the above problem is the σ_n^2.

Answers are put here.

4. Consider the following linear programming problem:

$$\begin{array}{ll} \max_{x_1,x_2} & z = x_1 + x_2 \\ \text{s.t.} & x_1 + 2x_2 \leq 4 \\ & 4x_1 + 2x_2 \leq 12 \\ & -x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{array}$$

(a) (10%) Please refer Note08 (Page 2) to draw the figure of the constraints by any means, and use that to solve the problem.
Answers are put here.

(b) (10%) Derive its dual problem and solve the dual problem by any means. Compare the solutions of the primal and the dual problems.
Answers are put here.

(c) (10%) Verify the complementarity slackness condition.
Answers are put here.

(d) (10%) Transform the problem to the standard form.
Answers are put here.

(e) (10%) Solve it by the simplex method, as provided in Figure 1, using $\vec{x}_0 = (0, 0)$. Indicate $B_k, N_k, \vec{s}_k, \vec{d}_k, p_k, q_k, \gamma_k$ in each step.
Answers are put here.

(1)	Given a basic feasible point $\vec{x}_0$ and the corresponding index set $\mathcal{B}_0$ and $\mathcal{N}_0$.	
(2)	For $k = 0, 1, \ldots$	
(3)		Let $B_k = A(:, \mathcal{B}_k), N_k = A(:, \mathcal{N}_k)$, $\vec{x}_B = \vec{x}_k(\mathcal{B}_k), \vec{x}_N = \vec{x}_k(\mathcal{N}_k)$, and $\vec{c}_B = \vec{c}_k(\mathcal{B}_k), \vec{c}_N = \vec{c}_k(\mathcal{N}_k)$.
(4)		Compute $\vec{s}_k = \vec{c}_N - N_k^T (B_k^{-1})^T \vec{c}_B$ (pricing)
(5)		If $\vec{s}_k \geq 0$, return the solution $\vec{x}_k$. (found optimal solution)
(6)		Select $q_k \in \mathcal{N}_k$ such that $\vec{s}_k(i_q) < 0$, where i_q is the index of q_k in $\mathcal{N}_k$
(7)		Compute $\vec{d}_k = B_k^{-1} A_k(:, q_k)$. (search direction)
(8)		If $\vec{d}_k \leq 0$, return `unbounded`. (unbounded case)
(9)		Compute $[\gamma_k, i_p] = \min\limits_{i, \vec{d}_k(i) > 0} \dfrac{\vec{x}_B(i)}{\vec{d}_k(i)}$ (ratio test) (The first return value is the minimum ratio; the second return value is the index of the minimum ratio.)
(10)		$x_{k+1} \begin{pmatrix} \mathcal{B} \\ \mathcal{N} \end{pmatrix} = \begin{pmatrix} \vec{x}_B \\ \vec{x}_N \end{pmatrix} + \gamma_k \begin{pmatrix} -\vec{d}_k \\ \vec{e}_{i_q} \end{pmatrix}$ ($\vec{e}_{i_q} = (0, \ldots, 1, \ldots, 0)^T$ is a unit vector with i_qth element 1.)
(11)		Let the i_pth element in $\mathcal{B}$ be p_k. (pivoting) $\mathcal{B}_{k+1} = (\mathcal{B}_k - \{p_k\}) \cup \{q_k\}$, $\mathcal{N}_{k+1} = (\mathcal{N}_k - \{q_k\}) \cup \{p_k\}$

Figure 1: The simplex method for solving (minimization) linear programming