EE/M 520, Spring 2006

Exam 2: Due Session 26 (9:30am at the ECE front office)

Solutions (version: May 9, 2006, 12:31) Total 50 pts.

1. (10 pts.) Suppose that we have a speech signal, represented as a finite sequence of real numbers $x_1, x_2, \ldots, x_n$. Suppose that we record this signal onto magnetic tape. The recorded speech signal is represented by another sequence of real numbers $y_1, y_2, \ldots, y_n$.

Suppose we model the recording process as a simple scaling of the original signal (i.e., we believe that an good model of the relationship between the recorded signal and the original signal is $y_i = \alpha x_i$ for some constant α that does not depend on i). Suppose we know exactly the original signal $x_1, x_2, \ldots, x_n$ as well as the recorded signal $y_1, y_2, \ldots, y_n$. Use the least-squares method to find a formula for estimating the scale factor α given this data. (You may assume that at least one x_i is nonzero.)

Ans.: Let $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$ and $\boldsymbol{y} = [y_1, y_2, \ldots, y_n]^T$. This least-squares estimation problem can be expressed as

$$\text{minimize } \|\alpha \boldsymbol{x} - \boldsymbol{y}\|^2,$$

with α as the decision variable. Assuming that $\boldsymbol{x} \neq \boldsymbol{0}$, the solution is unique and is given by

$$\alpha^* = (\boldsymbol{x}^T\boldsymbol{x})^{-1}\boldsymbol{x}^T\boldsymbol{y} = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\boldsymbol{x}^T\boldsymbol{x}} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}.$$

2. (16 pts.) Consider the constrained optimization problem

$$\begin{array}{ll} \text{maximize} & f(\boldsymbol{x}) \\ \text{subject to} & \|\boldsymbol{x}\| = 1, \end{array}$$

where $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ and $\boldsymbol{Q} = \boldsymbol{Q}^T$. We wish to apply a projected gradient algorithm fixed step size to this problem:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{\Pi}[\boldsymbol{x}^{(k)} + \alpha \nabla f(\boldsymbol{x}^{(k)})],$$

where $\alpha > 0$ and $\boldsymbol{\Pi}$ is the usual projection operator defined by $\boldsymbol{\Pi}[\boldsymbol{x}] = \arg\min_{\boldsymbol{z}\in\Omega} \|\boldsymbol{z} - \boldsymbol{x}\|$ and Ω is the constraint set.

a. Find a simple formula for $\boldsymbol{\Pi}[\boldsymbol{x}]$ in this problem (an explicit expression in terms of $\boldsymbol{x}$), assuming $\boldsymbol{x} \neq \boldsymbol{0}$.

b. For the remainder of the question, suppose that

$$\boldsymbol{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Find the solution(s) to this optimization problem.

c. Let $y^{(k)} = x_1^{(k)}/x_2^{(k)}$. Derive an expression for $y^{(k+1)}$ in terms of $y^{(k)}$ and α.

d. Assuming $x_2^{(0)} \neq 0$, use parts b and c to show that for any $\alpha > 0$, $\boldsymbol{x}^{(k)}$ converges to a solution to the optimization problem (i.e., the algorithm works).

e. In part d, what if $x_2^{(0)} = 0$?

Ans.: a. By drawing a simple picture, it is easy to see that $\boldsymbol{\Pi}[\boldsymbol{x}] = \boldsymbol{x}/\|\boldsymbol{x}\|$, provided $\boldsymbol{x} \neq \boldsymbol{0}$.

b. By inspection, we see that the solutions are $[0, 1]^T$ and $[0, -1]^T$. (Or use Rayleigh's inequality.)

c. Now,

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{\Pi}[\boldsymbol{x}^{(k)} + \alpha \nabla f(\boldsymbol{x}^{(k)})] = \beta_k(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{Q}\boldsymbol{x}^{(k)}) = \beta_k(\boldsymbol{I} + \alpha\boldsymbol{Q})\boldsymbol{x}^{(k)},$$

where $\beta_k = 1/\|(\boldsymbol{I} + \alpha\boldsymbol{Q})\boldsymbol{x}^{(k)}\|$. For the particular given form of $\boldsymbol{Q}$, we have

$$\begin{aligned} x_1^{(k+1)} &= \beta_k(1+\alpha)x_1^{(k)} \\ x_2^{(k+1)} &= \beta_k(1+2\alpha)x_2^{(k)}. \end{aligned}$$

Hence,

$$y^{(k+1)} = \left(\frac{1+\alpha}{1+2\alpha}\right) y^{(k)}.$$

d. Assuming $x_2^{(0)} \neq 0$, $y^{(0)}$ is well defined. Hence, by part c, we can write

$$y^{(k)} = \left(\frac{1+\alpha}{1+2\alpha}\right)^k y^{(0)}.$$

Because $\alpha > 0$,

$$\frac{1+\alpha}{1+2\alpha} < 1,$$

which implies that $y^{(k)} \to 0$. But

$$\begin{aligned} 1 &= \|\boldsymbol{x}^{(k)}\| \\ &= \sqrt{(x_1^{(k)})^2 + (x_2^{(k)})^2} \\ &= \sqrt{(x_2^{(k)})^2 \left(\frac{(x_1^{(k)})^2}{(x_2^{(k)})^2} + 1\right)} \\ &= |x_2^{(k)}|\sqrt{(y^{(k)})^2 + 1}, \end{aligned}$$

which implies that

$$|x_2^{(k)}| = \frac{1}{\sqrt{(y^{(k)})^2 + 1}}.$$

Because $y^{(k)} \to 0$, we have $|x_2^{(k)}| \to 1$. By the expression for $x_2^{(k+1)}$ in part c, we see that the sign of $x_2^{(k)}$ does not change with k. Hence, we deduce that either $x_2^{(k)} \to 1$ or $x_2^{(k)} \to -1$. This also implies that $x_1^{(k)} \to 0$. Hence, $\boldsymbol{x}^{(k)}$ converges to a solution to the problem.

e. If $x_2^{(0)} = 0$, then $x_2^{(k)} = 0$ for all k, which means that $x_1^{(k)} = 1$ or -1 for all k. In this case, the algorithm is stuck at the initial condition $[1, 0]^T$ or $[-1, 0]^T$ (which are in fact the minimizers).

3. (10 pts.) Suppose a computer supplier has two warehouses, one located in city A and another in city B. The supplier receives orders from two customers, one in city C and another in city D. The customer in city C orders 50 units, and the customer in city D orders 60 units. The number of units at the warehouse in city A is 70, and the number of units at the warehouse in city B is 80. The cost of shipping each unit from A to C is 1, from A to D is 2, from B to C is 3, and from B to D is 4.

Formulate the problem of deciding how many units from each warehouse should be shipped to each customer to minimize the total shipping cost (assuming the values of units to be shipped are real numbers). Express the problem as an equivalent standard form linear programming problem, and write down the *tableau* for the problem.

Ans.: Let x_1 and x_2 represent the number of units to be shipped from A to C and to D, respectively, and x_3 and x_4 represent the number of units to be shipped from B to C and to D, respectively. Then, the given problem can be formulated as the following linear program:

$$\begin{array}{ll} \text{minimize} & x_1 + 2x_2 + 3x_3 + 4x_4 \\ \text{subject to} & x_1 + x_3 = 50 \\ & x_2 + x_4 = 60 \\ & x_1 + x_2 \le 70 \\ & x_3 + x_4 \le 80 \\ & x_1, x_2, x_3, x_4 \ge 0. \end{array}$$

Introducing slack variables x_5 and x_6, we have the standard form problem

$$\begin{array}{ll} \text{minimize} & x_1 + 2x_2 + 3x_3 + 4x_4 \\ \text{subject to} & x_1 + x_3 = 50 \\ & x_2 + x_4 = 60 \\ & x_1 + x_2 + x_5 = 70 \\ & x_3 + x_4 + x_6 = 80 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \ge 0. \end{array}$$

The problem tableau is

$$\begin{array}{cccccc c} 1 & 0 & 1 & 0 & 0 & 0 & 50 \\ 0 & 1 & 0 & 1 & 0 & 0 & 60 \\ 1 & 1 & 0 & 0 & 1 & 0 & 70 \\ 0 & 0 & 1 & 1 & 0 & 1 & 80 \\ 1 & 2 & 3 & 4 & 0 & 0 & 0 \end{array}$$

4. (14 pts.) Suppose we are given a linear programming problem in standard form (written in the usual notation), and we are told that the vector $\boldsymbol{x} = [1, 0, 2, 3, 0]^T$ is a basic feasible solution with corresponding relative cost coefficient vector $\boldsymbol{r} = [0, 1, 0, 0, -1]^T$ and objective function value 6. We are also told that the vector $[-2, 0, 0, 0, 4]^T$ lies in the nullspace of $\boldsymbol{A}$.

a. Write down the canonical tableau corresponding to the above given basic feasible solution, filling in as many values of entries as possible (use the symbol $*$ for entries that cannot be determined from the given information). Clearly indicate the dimensions of the tableau.

b. Find a feasible solution with an objective function value that is strictly less than 6.

Ans.: a. From the given information, we have the 4×6 canonical tableau

$$\begin{bmatrix} 1 & * & 0 & 0 & 1/2 & 1 \\ 0 & * & 1 & 0 & 0 & 2 \\ 0 & * & 0 & 1 & 0 & 3 \\ 0 & 1 & 0 & 0 & -1 & -6 \end{bmatrix}$$

Explanations:

- The given vector $\boldsymbol{x}$ indicates that $\boldsymbol{A}$ is 3×5.
- In the above tableau, we assume that the basis is $[\boldsymbol{a}_1, \boldsymbol{a}_3, \boldsymbol{a}_4]$, in this order. Other permutations of orders will result in interchanging rows among the first three rows of the tableau.
- The fifth column represents the coordinates of $\boldsymbol{a}_5$ with respect to the basis $[\boldsymbol{a}_1, \boldsymbol{a}_3, \boldsymbol{a}_4]$. Because $[-2, 0, 0, 0, 4]^T$ lies in the nullspace of $\boldsymbol{A}$, we deduce that $-2\boldsymbol{a}_1 + 4\boldsymbol{a}_5 = \boldsymbol{0}$, which can be rewritten as $\boldsymbol{a}_5 = (1/2)\boldsymbol{a}_1 + 0\boldsymbol{a}_3 + 0\boldsymbol{a}_4$, and hence the coordinate vector is $[1/2, 0, 0]^T$.

b. Let $\boldsymbol{d}_0 = [-2, 0, 0, 0, 4]^T$. Then, $\boldsymbol{A}\boldsymbol{d}_0 = \boldsymbol{0}$. Therefore, the vector $\boldsymbol{x}' = \boldsymbol{x} + \varepsilon\boldsymbol{d}_0$ also satisfies $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Now, $\boldsymbol{x}' = \boldsymbol{x} + \varepsilon\boldsymbol{d}_0 = [1 - 2\varepsilon, 0, 2, 3, 4\varepsilon]^T$. For $\boldsymbol{x}'$ to be feasible, we must have $\varepsilon \leq 1/2$. Moreover, the objective function value of $\boldsymbol{x}'$ is $\boldsymbol{c}^T\boldsymbol{x}' = z_0 + r_5 x_5' = 6 - 4\varepsilon$, where z_0 is the objective function value of $\boldsymbol{x}$. So, if we pick any $\varepsilon \in (0, 1/2]$, then $\boldsymbol{x}'$ will be a feasible solution with objective function value strictly less than 6. For example, with $\varepsilon = 1/2$, $\boldsymbol{x}' = [0, 0, 2, 3, 2]^T$ is such a point. (We could also have obtained this solution by pivoting about the element $(1, 5)$ in the tableau of part a.)