164 Midterm 2 Solutions[1]

1. Question 1

Let $n \geq 1$. Let A be an $n \times n$ symmetric positive definite matrix. Let $b \in \mathbb{R}^n$. Suppose you want to solve for $x \in \mathbb{R}^n$ in the equation $Ax = b$.

Consider (i) inverting the matrix A and setting $x = A^{-1}b$, or (ii) using the Conjugate Gradient Method.

Describe the benefit of method (ii) over method (i), in terms of the number of arithmetic operations that can be required of each method (in the worst possible case). (It is possible to answer this question in around three sentences.) (In this question you can freely cite things we did in class.)

Solution. Inverting the matrix A using Gaussian elimination can require around n^3 arithmetic operations. On the other hand, the Conjugate Gradient method requires only around $n\max(m, n) = \max(nm, n^2)$ arithmetic operations, where m is the number of nonzero entries of A. So, the Conjugate Gradient method is much faster. There are even faster algorithms than the Conjugate Gradient method, but we will not discuss these in this class.

2. Question 2

Let A be an $m \times n$ real matrix with $m \geq n$. Assume that A has rank n. Show that $A^T A$ is positive definite.

Solution. Suppose A has rank n. Note that $A^T A$ is a real symmetric $n \times n$ matrix (since $(A^T A)^T = A^T A^{TT} = A^T A$), so $A^T A$ has n real eigenvalues. Also, for any $x \in \mathbb{R}^n$, $x^T A^T A x = (Ax)^T Ax \geq 0$. Therefore, all eigenvalues of $A^T A$ must be nonnegative. (If $A^T A$ had a negative eigenvalue λ with corresponding eigenvector $x \neq 0$, then $x^T A^T A x = x^T \lambda x = \lambda(x^T x) < 0$, a contradiction.) So, we know that $A^T A$ is positive semidefinite. It remains to show that $A^T A$ has no zero eigenvalues. We argue by contradiction. Suppose $A^T A x = 0$ for some $x \in \mathbb{R}^n$ with $x \neq 0$. Since A has rank n, the rank-nullity theorem says that A has nullity zero. That is, $Ax \neq 0$ if $x \neq 0$. But then $0 = x^T A^T A x = (Ax)^T Ax > 0$ since $Ax \neq 0$, a contradiction. We conclude that $A^T A$ is positive definite.

3. Question 3

Using the Simplex Algorithm, solve the following linear program:

$$\text{minimize} \quad -4x_1 - 2x_2 \quad \text{subject to the constraints}$$

$$x_1 + x_2 + x_3 = 6$$

$$2x_1 + x_2/2 + x_4 = 8, \qquad x \geq 0.$$

(Hint: start at the point $(x_1, x_2, x_3, x_4) = (0, 0, 6, 8)$.) (You may assume without proof that this linear program is non-degenerate.)

Solution. We will first check that $(0, 0, 6, 8)$ is a basic feasible solution. Let $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 1/2 & 0 & 1 \end{pmatrix}$ and let $b = \binom{6}{8}$. Let $S = \{3, 4\}$, and let $x = (0, 0, 6, 8)^T$. Note that $Ax = b$, $x \geq 0$, $x_i = 0$

[1]December 7, 2016,

when $i \notin S$, and the matrix $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is invertible. Therefore, $(0, 0, 6, 8)$ is a basic feasible solution, which allows us to begin the simplex method.

Let A_i denote the i^{th} column of A, $1 \leq i \leq 4$. We can now proceed in two different ways. Method 1. We pivot on the first column. Then $A_1 = A_3 + 2A_4$, so $A(1, 0, -1, -2)^T = 0$, and for any $t \in \mathbb{R}$, $A[(0, 0, 6, 8) + t(1, 0, -1, -2)] = b$. Since $x \geq 0$ in the feasible set, we choose $t = 4$, to get $A(4, 0, 2, 0) = b$. Then, if we define $S := \{1, 3\}$ and set $x := (4, 0, 2, 0)$, then the matrix $B = (A_1, A_3) = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$ is invertible, and $x_i = 0$ when $i \notin S$. Therefore, $(4, 0, 2, 0)$ is a basic feasible solution. At the previous basic feasible solution, we had $-4x_1 - 2x_2 = 0$, and we now have $-4x_1 - 2x_2 = -16$. So, the new basic feasible solution is smaller, so we maintain this solution and continue the algorithm.

We now pivot on the second column. Then $A_2 = (1, 1/2)^T = (1/4)(1, 2)^T + (3/4)(1, 0)^T$, and for any $t \in \mathbb{R}$, $A[(4, 0, 2, 0) + t(-1/4, 1, -3/4, 0)] = b$. Since $x \geq 0$ in the feasible set, we choose $t = 8/3$, to get $A(10/3, 8/3, 0, 0) = b$. Then, if we define $S := \{1, 2\}$ and set $x := (10/3, 8/3, 0, 0)$, then the matrix $B = (A_1, A_2) = \begin{pmatrix} 1 & 1 \\ 2 & 1/2 \end{pmatrix}$ is invertible, and $x_i = 0$ when $i \notin S$. Therefore, $(10/3, 8/3, 0, 0)$ is a basic feasible solution. At the previous basic feasible solution, we had $-4x_1 - 2x_2 = -16$, and we now have $-4x_1 - 2x_2 = -40/3 - 16/3 = -56/3 < -48/3$. So, the new basic feasible solution is smaller, so we maintain this solution and continue the algorithm.

We have so far checked three basic feasible solutions. We will now show that $(10/3, 8/3, 0, 0)$ achieves the minimum value of the linear program. From Proposition 4.24 in the notes, we know the Simplex Algorithm will terminate at the minimum value. So, it suffices to show that the simplex algorithm terminates. We consider the two available pivots. If we pivot on the third column, we will return to $(4, 0, 2, 0)$, which must increase the linear program. And if we pivot on the last column, we have $A_4 = (0, 1)^T = (2/3)[(1, 2)^T - (1, 1/2)^T]$, and for any $t \in \mathbb{R}$, $A[(10/3, 8/3, 0, 0) + t(-2/3, 2/3, 0, 1)] = b$. We choose $t = 5$ and consider the point $x = (0, 6, 0, 5)$. This is a basic feasible solution, but $-4x_1 - 2x_2 = -12 > -56/3$. That is, moving from $(10/3, 8/3, 0, 0)$ to any other basic feasible solution (as in the Simplex Algorithm) produces a point that increases the linear program. We conclude that $(10/3, 8/3, 0, 0)$ is the minimal point.

Note. The set of all basic feasible solutions is: $(0, 0, 6, 8)$, $(4, 0, 2, 0)$, $(10/3, 8/3, 0, 0)$, $(0, 6, 0, 5)$. (We cannot use $S = \{1, 4\}$, since then $x_1 = 6$, and then no solution $x \geq 0$ exists with $x_2 = x_3 = 0$. Also, we cannot use $S = \{2, 3\}$, since then $x_2 = 16$, and then no solution $x \geq 0$ exists with $x_1 = x_4 = 0$.)

4. Question 4

Describe in detail how the ellipsoid method works. In your description you should answer the following questions.

- What is a polytope? What is an ellipsoid?
- What are the input and output of the algorithm?
- What problem does the ellipsoid method solve?
- Under what assumptions does the algorithm terminate?
- What happens in "one step" of the algorithm?

(You do not have to provide every single detail as we did in class, but try to include as much detail as you can. Your goal to try to convince the grader, as best you can, that you understand how the ellipsoid method works.) (Only answering the above five questions does not constitute a complete answer.) (Also drawing a picture might be helpful, though a correct answer can be given without drawing any pictures.)

Solution. A polytope is the intersection of a finite number of half spaces. An ellipsoid is a set of the form $\{x \in \mathbb{R}^n \colon x^T A x \leq 1\}$ where A is a real symmetric positive definite $n \times n$ matrix. The input of the algorithm is a nonempty polytope P, given in the form

$$P = \{x \in \mathbb{R}^n \colon Ax \leq b\}$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$. We are also given a Euclidean ball B such that $P \subseteq B$. The goal of the algorithm is to find a point $p \in P$. It is assumed that $\mathrm{vol}_n(B)/\mathrm{vol}_n(P) \leq c'$ for some $c' \geq 1$. Under this assumption, the algorithm terminates after at most $2n\log(c')$ steps.

The algorithm begins with the Euclidean ball B. We then ask if the center c of this ellipsoid B is in P or not. If the center is in P, we are done, since we have found a point in P. If the center is not in P, then one of the inequalities of $Ax \leq b$ has been violated. That is, there exists some row $a^{(j)}$ of A, $1 \leq j \leq m$, such that $\langle a^{(j)}, x\rangle > b_i$. By definition of P, $P \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x\rangle > b_j\} = \emptyset$. Since $P \subseteq B$, we conclude that

$$\begin{aligned} P &\subseteq B \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x\rangle \leq b_j\} \subseteq B \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x\rangle \leq \langle a^{(j)}, y_k\rangle\} \\ &= B \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x - y_k\rangle \leq 0\}. \end{aligned}$$

That is, P is contained in the half ellipsoid $B \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x - y_k\rangle \leq 0\}$. (This set is called a half ellipsoid, since it is the intersection of the ellipsoid B with the half space.) We then let E_1 be the Löwner-John ellipsoid of $B \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x - y_k\rangle \leq 0\}$. That is, E_1 is an ellipsoid $E_1 \supseteq B \cap \{x \in \mathbb{R}^n \colon \langle a^{(j)}, x - y_k\rangle \leq 0\}$, so that $P \subseteq E_1$. And the volume of E_1 decreases relative to B, i.e. $\mathrm{vol}_n(E_1)/\mathrm{vol}_n(B) < e^{-1/n}$, as shown on the homework. We now repeat the above procedure, replacing E_1 with B. We repeat this procedure $N > 2n\log(c')$ steps. Then

$$\frac{\mathrm{vol}_n(E_N)}{\mathrm{vol}_n(B)} = \frac{\mathrm{vol}_n(E_N)}{\mathrm{vol}_n(E_0)} = \prod_{k=0}^{N-1} \frac{\mathrm{vol}_n(E_{k+1})}{\mathrm{vol}_n(E_k)} \overset{(*)}{<} (e^{-1/2n})^{2n\log(c')} = (c')^{-1}.$$

That is, $\mathrm{vol}_n(E_N) < \mathrm{vol}_n(B)(c')^{-1}$. On the other hand, $\mathrm{vol}_n(B)/\mathrm{vol}_n(P) \leq c'$, so $\mathrm{vol}_n(P) \geq (c')^{-1}\mathrm{vol}_n(B)$. That is, $\mathrm{vol}_n(E_N) < \mathrm{vol}_n(P)$. So, it is not possible that $P \subseteq E_N$. That is, the algorithm terminates after at most N steps. At termination, we must have found a point in P, as desired.