Homework 4 Solutions
CS322: Summer 2004
July 21, 2004

3. From Exercise 2, we see that $Q^T a = H a = (\|a\|_2, 0, \ldots, 0)^T = R$, where $H = I - 2w w^T / \|w\|^2$ and $v = a - \|a\|_2 e_1$. So $Q = H^T = H$. The LS solution to $ax \approx b$ is $x^* = R^{-1} H(1,:) b = H(1,:) b / \|a\|_2$.

   To show that columns of $Q$ form an orthonormal basis, first recall that by definition of orthogonal matrix, $Q(:, i)^T Q(:, i) = 1$ for any $i$ and $Q(:, i)^T Q(:, j) = 0$ for any $j \neq i$. Hence its columns are orthonormal.

   To show that it is a basis for $\mathbb{R}^n$, it suffices to show that these $n$ vectors are linearly independent. Suppose $y$ is a vector such that $Q(:, 1) y_1 + Q(:, 2) y_2 + \ldots + Q(:, n) y_n = 0$. Then $0 = Q(:, i)^T 0 = Q(:, i)^T Q(:, 1) y_1 + Q(:, i)^T Q(:, 2) y_2 + \ldots + Q(:, i)^T Q(:, n) y_n = y_i$. Therefore $y_i = 0$. Since $i$ is chosen arbitrarily, the vector $y$ must be zero. Therefore, the columns of $Q$ are linearly independent.

   To see that $a / \|a\|_2$ is one of the basis vectors, note that $Q e_1$ is the first column of $Q$. Let $\alpha$ denote $\|a\|_2$. Simplify using the same technique as in question 4.3 of PS4 results in

$$Q e_1 = e_1 - \frac{2(a - \alpha e_1)(a^T e_1 - \alpha e_1^T e_1)}{(a^T - \alpha e_1^T)(a - \alpha e_1)}$$

$$= e_1 - \frac{(a - \alpha e_1)(e_1^T a - \alpha)}{-\alpha(e_1^T a - \alpha)}$$

$$= e_1 + \frac{1}{\alpha}(a - \alpha e_1)$$

$$= \frac{a}{\|a\|_2}.$$  \hspace{1cm} (1)

4. Let $w$ be any arbitrary nonzero vector of length $k$, then $w^T X^T A X w = (X w)^T A (X w) = y^T A y$, where $y = W x \in \mathbb{R}^n$. Since $X$ is of full
rank, columns of $X$ are linearly independent. Therefore, $y = Xw = \sum_{i=1}^{k} w_i X(:, i) \neq 0$ as $w$ is nonzero. By positive definiteness of $A$, $w^{T}X^{T}AXw = y^{T}Ay \neq 0$ for any nonzero $y$. So, $X^{T}AX$ is positive definite.

Note that it is important to argue that $Xw$ cannot be zero. If it were possible, then there would be a nonzero vector $w$ such that $w^{T}X^{T}AXw = 0^{T}A0 = 0$, which would imply that $X^{T}AX$ is not positive definite.

5. For the first part of the hint, let $x \in \mathbb{R}^n$ be defined as

$$x_k = \begin{cases} 
1, & k = i, \\
-1, & k = j, \\
0, & \text{otherwise.}
\end{cases}$$

By positive definiteness of $A$, we see that $x^{T}Ax = A(i, i) - A(i, j) - A(j, i) + A(j, j) = A(i, i) - 2A(i, j) + A(j, j) \geq 0$ ($A(i, j) = A(j, i)$ by symmetry). This inequality implies that

$$A(i, j) \leq (A(i, i) + A(j, j))/2. \tag{2}$$

Similarly, define $y$ as

$$y_k = \begin{cases} 
1, & k = i \text{ or } k = j, \\
0, & \text{otherwise.}
\end{cases}$$

We have $y^{T}Ay = A(i, i) + A(i, j) + A(j, i) + A(j, j) = A(i, i) + 2A(i, j) + A(j, j) \geq 0$. This inequality implies that

$$-A(i, j) \leq (A(i, i) + A(j, j))/2. \tag{3}$$

Finally, (2) and (3) imply $|A(i, j)| \leq (A(i, i) + A(j, j))/2$. 