CS 205a Fall 2008 Midterm 1 Solutions

1 Multiple Choice

1. (b) (c) (d)
2. (d)
3. (a) (b)
4. (a) (c)
5. (a) (c) (e)
6. (c) (d)

2 Error

1. Modeling errors: parts of the problem may be ignored; Truncation errors: errors from mathematical approximation of an equation; Rounding errors: errors from limited precision of machine numbers.
2. Other sources of errors (like modeling erros, rounding errors, etc.) may dominate truncation errors caused by approximating the equations.

3 QR

1. Given a $n \times n$ Householder matrix H, its eigenvalues are 1 (with $n-1$ multiplicity) and -1 (with 1 multiplicity). Because $H = I - 2vv^T/v^Tv$ is a reflection matrix, only components along one direction (eigenvector v) are reverted, which corresponds to eigenvalue -1, and all other compoenents remain, which correspond to eigenvalue 1.
2. 3
3. $(-2, 0, 0, 0)^T$
4. $(-2, 0, 0, 0)^T$
5. With QR factorization, the equation $Ax = b$ is transformed into $QRx = b$. We apply QT to both sides of the equation which leads to $Rx = Q^Tb$. Then we can use the backward substitution to solve this uppper triangular system. To show that QR solver minimizes the two-norm of the residual, first we know that the normal equations system $A^TAx = A^Tb$ gives the least squares solution as $x = (A^TA)^{-1}A^Tb$. Let $A = QR$, then $A^T = R^TQ^T$, therefore

$$A^TA = (R^TQ^T)(QR) = R^T(Q^TQ)R = R^TIR = R^TR$$

and

$$x = (A^TA)^{-1}A^Tb = (R^TR)^{-1}R^TQ^Tb = R^{-1}R^{-T}R^TQ^Tb = R^{-1}Q^Tb$$

which is $Rx = Q^Tb$.

4 Optimization

1. The secant method looks like

$$\begin{aligned} x_{k+1} =& x_k - \frac{(x_k - x_{k-1})f(x_k)}{f(x_k) - f(x_{k-1})} \\ =& \frac{f(x_k) - f(x_{k-1}))x_k - (x_k - x_{k-1})f(x_k)}{f(x_k) - f(x_{k-1})} \\ =& \frac{x_k f(x_{k-1}) - x_{k-1} f(x_k)}{f(x_k) - f(x_{k-1})} \end{aligned}$$

   which is the given formula.

2. From $f'(x) = Ax - b = 0$ and the Hessian matrix $H = A$ is symmetric positive definite, we know that the minimum of $f(x)$ is the solution of $Ax = b$. The first step of Newton's method looks like

$$\begin{aligned} x_1 &= x_0 - H^{-1} f'(x_0) \\ &= x_0 - A^{-1}(Ax_0 - b) \\ &= x_0 - x_0 + A^{-1}b \\ &= A^{-1}b \end{aligned}$$

   which is the exact solution of the minimization problem of $f(x)$.

3. The first step of steepest descent method looks like $x_1 = x_0 + r_0\alpha$, where $r_0 = b - Ax_0$, and α is solved to make the norm of $r_1 = b - Ax_1$ minimized. Since $Ax_0 = \lambda x_0$, $r_0 = b - \lambda x_0$, $x_1 = x_0 + r_0\alpha = x_0 + \alpha b - \alpha\lambda x_0$, and thus $r_1 = b - Ax_1 = b - Ax_0 + \alpha Ab - \alpha\lambda Ax_0 = b - \lambda x_0 + \alpha\lambda b - \alpha\lambda^2 x_0$. Clearly, if $\alpha = (\lambda x_0 - b)/\lambda(b - \lambda x_0)$, then $r_1 = 0$ and thus r_1 has the minimum norm. This is the exact solution of the problem. Therefore, if x_0 is an eigenvector of A, then steepest descent method needs only one step to converge.

5 Properties of Matrices

1. Given a positive definite matrix A, we have for any non-zero vector x, $x^T Ax > 0$. (Note: the inequility here enforce that even x is a complex vector, $x^T Ax$ is always real, therefore A must be a Hermitian matrix, whose eigenvalues are all real.) Assume that λ is an eigenvalue of A, and v is a corresponding eigenvector. Then $v^T Av = \lambda v^T v > 0$. Since $v^T v > 0$, and $v^T Av > 0$, we have $\lambda > 0$.

2. If A is idempotent, then $AA = A$. Thus for any non-zero vector x, we have $y = Ax$, such that $Ay = AAx = Ax = y$. If all such y are zero vectors, then $A = 0$, it has only zero eigenvalues. Otherwise, 1 is an eigenvalue of A and y is an eigenvector of A. Also, we have $z = x - Ax$, such that $Az = Ax - AAx = Ax - Ax = 0$. If all such z are zero vectors, then $A = I$, it has only 1 eigenvalues. Otherwise, 0 is an eigenvalue of A and z is an eigenvector of A.

3. First, it's trivial to prove that $A(A^T A)^{-1} A^T$ is symmetric. Then we multiply this matrix by itself

$$A(A^T A)^{-1} A^T A(A^T A)^{-1} A^T = A(A^T A)^{-1} \left(A^T A(A^T A)^{-1}\right) A^T = A(A^T A)^{-1} I A^T = A(A^T A)^{-1} A^T$$

   therefore it is a projection matrix. Given that this matrix is symmetric and a projection matrix, we can conclude that it is an orthogonal matrix. We know that the normal equation $A^T Ax = A^T b$ projects the residual $r = b - Ax$ onto the null space of A^T. Given any vector b, the solution of the normal equation $A^T Ax = A^T b$ is $x = (A^T A)^{-1} A^T b$. The residual $r = b - Ax = b - A(A^T A)^{-1} A^T b$ is the component of b in the null space of A^T. Since the column space and the null space of A^T are complementary, $A(A^T A)^{-1} A^T b = b - r$ is the compoenent of b projected onto the column space of A. Therefore $A(A^T A)^{-1} A^T$ is an orthogonal projecter onto the span of A.

4. As seen above, the projector $A(A^TA)^{-1}A^T$ simply projects the least-squares solution back to the column space of A.

5. This statement is wrong. It holds true iff the matrix is an orthogonal projection. Assume that A is a orthogonal projection matrix, then we have $AA = A$ and $A^TA = A$, therefore $AA = A^TA$, $A = A^T$. A counterexample of the original statement is

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$