Mar 25, 2023

COL 726 Minor II

Give justifications for your answers. Without a valid reasoning, you may not get any marks. You may use any algorithm discussed in class.

1. **(4 marks)** Let A be a square matrix (which need not be Hermitian). We say that A is positive definite iff $x^*Ax > 0$ for all non-zero vectors x. Is it true that if A is positive definite, then all of its eigenvalues are positive real numbers? Is it true that if all the eigenvalues of A are positive real numbers, then it is positive definite?

   Solution: The first statement is true. Suppose λ is an eigenvalue of A and let v be an eigenvector corresponding to λ. Then

   $$v^*Av = \lambda \cdot v^*v.$$

   Since we know that RHS is a real positive number and v^*v is real positive, it follows that λ must be real positive.

   The converse is false. Consider for example the matrix

   $$A = \begin{pmatrix} 1 & 10 \\ 0 & 1 \end{pmatrix}.$$

   The eigenvalues of a triangular matrix are on the diagonal, and so the only eigenvalue of A, i.e., 1, is positive. But it is easy to check that if $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, then $x^*Ax < 0$.

2. **(4 marks)** While running Gaussian elimination with partial pivoting on a square matrix A, we say that the procedure *fails* if we are not able to find a non-zero pivot at some step. Is it possible that the procedure fails when A is an invertible square matrix? Either give a counter-example or a proof of this statement.

   Solution: The procedure will never fail. Indeed, suppose the procedure fails in iteration i, let A_i be the current matrix. Then the first i columns have 0 in rows i upto n. But then we essentially have i vectors, where each of them has $(i-1)$ non-zero coordinates. Any set of i vectors in $(i-1)$-dimensional subspace must be linearly independent. Then A_i is singular. Since A_i is obtained from A by left multiplying with invertible matrices, A is also singular, a contradiction.

3. **(3 marks)** A is a 6×6 square Hermitian matrix with eigenvalues $-10, -6, -2, 3, 7, 9$. Starting with a randomly chosen vector, you run the following two algorithms on A: power iteration and power iteration with shift $\mu = 2$. Which eigenvalues will the two procedures converge to? In which case will the convergence be faster?

Solution: Power iteration converges to the largest eigenvalue in absolute value, so it converges to -10. In the second case, the eigenvalues of $A - \mu I$ are $-12, -8, -4, 1, 5, 7$. The two smallest eigenvalues in absolute value are 1 and -4. So the inverse iteration would converge to 3 (or 1 of $A - \mu I$). The second convergence will be faster because the rate of convergence would be proportional to $1/4$, whereas in the first case it would be $9/10$.

4. **(5 marks)** You are given a 3×3 Hermitian matrix A whose eigenvalues are $5, -5, 2$ and the corresponding eigenvectors are $(1,1,1), (1,1,-2), (1,-1,0)$. You run power iteration on A starting from the initial vector $(0,1,1)$. Will the power iteration on A converge to a vector? If so, what will this (unit) vector be? Will the corresponding Rayleigh quotient converge to a value? If so, what is this value?

   Solution: Since there are two equal eigenvalues of same absolute value and the starting vector has non-zero components along both the corresponding eigenvectors, this will not converge to any vector. To see this, let v_1, v_2, v_3 be the three normalized (to unit length) eigenvectors for $5, -5, 2$ respectively, and suppose the starting vector is given by $v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$. Then $A^k v = 5^k \alpha_1 v_1 + (-5)^k \alpha_2 v_2 + 2^k \alpha_3 v_3$. For large k, when you normalize this vector it will be either $w = \alpha_1 v_1 + \alpha_2 v_2$ or $w' = \alpha_1 v_1 - \alpha_2 v_2$ depending on whether k is odd or even. So this will NOT converge.

   The Rayleigh quotient of the corresponding vector (this will be same for w' as well) will be
$$\frac{w^* A w}{w^* w} = \frac{5\alpha_1^2 - 5\alpha_2^2}{\alpha_1^2 + \alpha_2^2}.$$
   So it converges to the above value. A simple calculation shows that $\alpha_1 = \frac{2}{\sqrt{3}}, \alpha_2 = \frac{-1}{\sqrt{6}}$.

5. **(4 marks)** Let W be a $n \times n$ (real) symmetric positive definite matrix. For a vector $x \in \Re^n$, define $||x||_W$ as $\sqrt{x^T W x}$. Let A be a real $m \times n$ matrix, and b a real column vector of size n. Assume that $m \geq n$. Give an efficient algorithm to find a vector x such that $||Ax - b||_W$ is minimised.

   Solution: We reduce this to the usual least squares problem. We know that any positive definite matrix W can be written as $L^T L$ for a matrix L (Cholesky factorization). Now notice that $||Ax - b||_W^2 = (Ax - b)^T W (Ax - b) = (L(Ax - b))^T (L(Ax - b))$. So if A' denotes LA and b' denotes Lb, the problem reduces to finding x such that $||L'x - b'||$ is minimized. This is the usual least squares problem.