MA2501, Spring 2020, Numerical Methods

February 6, 2020

Problem 1

We denote by $\sigma(A)$ the set of all eigenvalues of the square matrix A. Recall that the spectral radius of A is $\rho(A) := \max_{\lambda \in \sigma(A)} |\lambda|$. Recall that by definition the condition number of A with respect to the 2-norm is $\mathcal{K}_2(A) := \|A\|_2 \, \|A^{-1}\|_2$.

Prove that if A is a $n \times n$ real, symmetric matrix then

$$\mathcal{K}_2(A) = \frac{\max_{\lambda \in \sigma(A)} |\lambda|}{\min_{\lambda \in \sigma(A)} |\lambda|}.$$

Guidelines:

- Use in the proof the following definition of 2-norm, $\|A\|_2 = \rho(A^T A)^{\frac{1}{2}}$ (see the book of Süli and Mayers, theorem 2.9 for this result).
- Prove that, for any matrix A $n \times n$, if λ is an eigenvalue of A then λ^2 is an eigenvalue of A^2.
- Prove that for A symmetric $\|A\|_2 = \rho(A)$.
- Prove that, for any matrix A $n \times n$, if λ is an eigenvalue of A then λ^{-1} is an eigenvalue of A^{-1}.
- Prove that for A symmetric $\|A^{-1}\|_2 = \frac{1}{\min_{\lambda \in \sigma(A)} |\lambda|}$.

Problem 2

Householder transformations

An Householder transformation is a $n \times n$ matrix of the form

$$P = I - 2ww^T,$$

where I is the identity matrix $n \times n$, $w \in \mathbb{R}^n$ with w of 2-norm equal to 1. Given $x \in \mathbb{R}^n$, we can define a Householder transformation such that

$$Px = \gamma\,\mathbf{e}_1, \qquad \gamma \in \mathbb{R}$$

and $\mathbf{e}_1$ the first canonical vector. If $x = 0$, $P\,\mathbf{0} = 0\,\mathbf{e}_1$. If $x \neq 0$, this can be achieved by taking

$$w := \tilde{w}/\|\tilde{w}\|_2, \quad \text{where} \quad \tilde{w} := x \pm \|x\|_2\mathbf{e}_1. \tag{1}$$

a) Prove that a Hauseholder transformation is symmetric and orthogonal, that is $P = P^T$ and $P^TP = I$, where I is the identity matrix.

b) To compute the QR factorization of a matrix A $m \times n$, one can use n Householder transformations. Consider the matrix

$$A := \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix},$$

apply two Householder transformations to transformations P_1 and P_2 to obtain an upper triangular 2×3 matrix R such that

$$P_2P_1A = R$$

and

$$A = QR, \quad \text{where} \quad Q := P_1^TP_2^T.$$

The last row of R is the zero row vector. Note $P_2 = I - 2vv^T$ must be constructed such that the first row and column of P_1A will remain unchanged when applying P_2 to (P_1A), and only the last two elements of the second column of P_1A are affected by the multiplication by P_2. This can be achieved taking the vector $v \in \mathbb{R}^3$ with the first component equal to zero.

Show that Q is an orthogonal matrix, i.e. $Q^T Q = QQ^T = I$.

Solution:

$$P_1 = \begin{bmatrix} \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ 0 & 1 & 0 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \end{bmatrix}, \qquad P_1 A = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 1 \\ 0 & \sqrt{2} \end{bmatrix},$$

$$P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & -b & -a \end{bmatrix} \qquad P_2 P_1 A = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \\ 0 & 0 \end{bmatrix},$$

with $a = 1-2(1-\sqrt(3))^2/((1-\sqrt(3))^2+2)$, $b = -2(1-\sqrt(3))\sqrt(2)/((1-\sqrt(3))^2 + 2)$.

$$Q = P_1^T P_2^T, \qquad R = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \\ 0 & 0 \end{bmatrix}.$$

c) Consider the least squares problem

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

compute the normal equations and find the least squares solution. Use then the method based on the QR- factorization (Theorem 2.13) and verify that you obtain the same answer.

You will consider the QR-factorization from **b)** and to get $\hat{Q}$ and $\hat{R}$ as in theorem 2.13, you consider only the first two columns of Q and the first two rows and columns of R.

Solution:

Normal equations

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Least squares solution $x = 3/2 + 1/3 = 1.8333$, $y = -1/3 = -0.3333$.

To get $\hat{Q}$ and $\hat{R}$ from the theorem, we consider only the first two columns of Q and the first two rows and columns of R.

$$\hat{Q} = \begin{bmatrix} \sqrt{2}/2 & a \\ 0 & a \\ \sqrt{2}/2 & -a \end{bmatrix}, \quad \hat{R} = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix},$$

then computing $\hat{R}^{-1}\hat{Q}^T b$ one verifies that the obtained vector of $\mathbb{R}^2$ coincides with the solution of the normal equations.

d) Compute the SVD of A. The matrix A has rank equal to 2 (verify). Compute a rank one approximation of the matrix A using the SVD.

Solution:

$$V = \begin{bmatrix} 0.4472 & -0.8944 \\ 0.8944 & 0.4472 \end{bmatrix} \quad \Sigma = \mathrm{diag}(\sqrt{6}, 1), U = \begin{bmatrix} 0 & 0.9129 \\ 0.4472 & 0.3651 \\ -0.8944 & 0.1826 \end{bmatrix}$$

Appendix

We outline here a procedure to compute by hand the SVD of a matrix A $m \times n$ (as stated in theorem 2.14). Numerical algorithms for the SVD are not covered in this course.

- Compute $A^T A$, this is a symmetric and positive definite $n \times n$ matrix, compute the eigenvalues and a basis of orthonormal eigenvectors for $A^T A$.
- The square roots of the eigenvelues of $A^T A$ are the singular values σ_i, which are also the diagonal elements of the matrix Σ of the SVD.
- The orthonormal eigenvectors $v_1, \ldots, v_n$ are the clumns of the matrix V^T $n \times n$.
- Consider next $u_i := \frac{1}{\sigma_i} A v_i$, $i = 1, \ldots, n$, these are the first n columns of the matrix U $m \times n$. (Verify that these are orthonormal).