TW776: Assignment 1 (solutions)

Properties of the Eigenvalues & Vectors of Square Matrices

(a) Recall the formula on p. 92 in Ascher & Greif, namely $\det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$. Set $\lambda = 0$ to obtain $\det(-A) = (-1)^n \lambda_1\lambda_2\cdots\lambda_n$. By replacing A by $-A$ all eigenvalues change sign, so $\det(A) = (-1)^n(-1)^n\lambda_1\lambda_2\cdots\lambda_n = \lambda_1\lambda_2\cdots\lambda_n$. If A is singular, $\det(A) = 0$, implying that one of the eigenvalues must be zero in the product $\lambda_1\lambda_2\cdots\lambda_n$.

(b) We show only for the 3×3 case, the general case follows by induction. With $n = 3$ in the formula in part (a), $\det(\lambda I - A) = (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3) = \lambda^3 - (\lambda_1+\lambda_2+\lambda_3)\lambda^2 + \ldots$ On the other hand,

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{23} \end{vmatrix} = \lambda^3 - (a_{11}+a_{22}+a_{33})\lambda^2 + \ldots$$

Comparing the coefficients of λ^2 in the two formulas yields $\lambda_1+\lambda_2+\lambda_3 = a_{11}+a_{22}+a_{33}$.

(c) Assume A is nonsingular. Then we can multiply $A\boldsymbol{x} = \lambda\boldsymbol{x}$ either side by the inverse of A to obtain $\boldsymbol{x} = \lambda A^{-1}\boldsymbol{x}$. Dividing by λ yields $A^{-1}\boldsymbol{x} = (1/\lambda)\boldsymbol{x}$ $(\lambda \neq 0)$.

(d) Assume A is a real, symmetric matrix, with a complex eigenvalue, λ. Then

$$\lambda\,\boldsymbol{x} = A\boldsymbol{x}, \qquad \bar{\lambda}\,\bar{\boldsymbol{x}} = A\bar{\boldsymbol{x}},$$

where we took the complex conjugate of the first equation to get the second. Premultiply the first equation by $\bar{\boldsymbol{x}}^T$ and the second by $\boldsymbol{x}^T$ and subtract to get

$$(\lambda - \bar{\lambda})\,\boldsymbol{x}^T\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}}^T A\boldsymbol{x} - \boldsymbol{x}^T A\bar{\boldsymbol{x}},$$

where we have used the fact that $\bar{\boldsymbol{x}}^T\boldsymbol{x} = \boldsymbol{x}^T\bar{\boldsymbol{x}}$. Since A is symmetric, $A = A^T$, and the RHS can be expressed as $\bar{\boldsymbol{x}}^T A\boldsymbol{x} - \boldsymbol{x}^T A\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}}^T A\boldsymbol{x} - \left(\bar{\boldsymbol{x}}^T A\boldsymbol{x}\right)^T$. Since the transpose of a scalar is equal to the scalar, the term on the right vanishes. Therefore $(\lambda - \bar{\lambda})\,\boldsymbol{x}^T\bar{\boldsymbol{x}} = 0$, and since $\boldsymbol{x}$ is an eigenvector it cannot be the zero vector. Therefore $\lambda = \bar{\lambda}$, i.e., λ is real. [1]

(e) Using the same strategy as in the previous problem, let λ_1 and λ_2 be distinct real eigenvalues with corresponding real eigenvectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and consider

$$\lambda_1\,\boldsymbol{x}_1 = A\boldsymbol{x}_1, \qquad \lambda_2\,\boldsymbol{x}_2 = A\boldsymbol{x}_2$$

Premultiply the first equation by $\boldsymbol{x}_2^T$ and the second by $\boldsymbol{x}_1^T$ and subtract to get

$$(\lambda_1 - \lambda_2)\boldsymbol{x}_1^T\boldsymbol{x}_2 = \boldsymbol{x}_2^T A\boldsymbol{x}_1 - \boldsymbol{x}_1^T A\boldsymbol{x}_2$$

where we have used the fact that $\boldsymbol{x}_2^T\boldsymbol{x}_1 = \boldsymbol{x}_1^T\boldsymbol{x}_2$. Since A is symmetric, $A = A^T$, and the RHS can be expressed as $\boldsymbol{x}_2^T A\boldsymbol{x}_1 - \boldsymbol{x}_1^T A\boldsymbol{x}_2 = \boldsymbol{x}_2^T A\boldsymbol{x}_1 - \left(\boldsymbol{x}_2^T A\boldsymbol{x}_1\right)^T$, which is zero for the same reason as in part (d). Therefore $(\lambda_1 - \lambda_2)\boldsymbol{x}_1^T\boldsymbol{x}_2 = 0$, and since $\lambda_1 \neq \lambda_2$, it follows that $\boldsymbol{x}_1^T\boldsymbol{x}_2 = 0$.

(f) Take the 2-norm both sides of the eigenvalue expression $A\boldsymbol{x} = \lambda\boldsymbol{x}$, that is $\|A\boldsymbol{x}\|_2 = \|\lambda\boldsymbol{x}\|_2$. On the left, use the formula at the bottom of the page 85 in Ascher & Greif, that is $\|A\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$ if A is orthogonal. On the right, we have $\|\lambda\boldsymbol{x}\|_2 = |\lambda|\,\|\boldsymbol{x}\|_2$, which follows from the second property of a vector norm, page 79. Thus $\|\boldsymbol{x}\|_2 = |\lambda|\,\|\boldsymbol{x}\|_2$, and since $\boldsymbol{x}$ is an eigenvector it is not the zero vector so the factor $\|\boldsymbol{x}\|_2$ can be cancelled both sides. Hence $|\lambda| = 1$. (Note: we cannot conclude from this that $\lambda = \pm 1$, as the eigenvalues may be complex.)

---

[1] In this problem we used the conjugate transpose, $\bar{\boldsymbol{x}}^T$, which is more commonly denoted by, $\boldsymbol{x}^H$, known as the Hermitian transpose. See exercise 14, p. 147, in Ascher & Greif.

Vector and Matrix Norms

(a) We prove the three properties on p. 79 of Ascher & Greif:

(i) $\|\boldsymbol{x}\|_A = \sqrt{\boldsymbol{x}^T A \boldsymbol{x}} > 0$ for $\boldsymbol{x} \neq \boldsymbol{0}$ since A is positive definite (recall definition, top of page 84).

(ii) $\|\alpha\boldsymbol{x}\|_A = \sqrt{(\alpha\boldsymbol{x})^T A(\alpha\boldsymbol{x})} = |\alpha|\sqrt{\boldsymbol{x}^T A \boldsymbol{x}}$

(iii) Since A is s.p.d. it has a Cholesky factorisation[2], say $A = LL^T$. On the left, we can therefore write

$$\|\boldsymbol{x}+\boldsymbol{y}\|_A = \sqrt{(\boldsymbol{x}+\boldsymbol{y})^T LL^T(\boldsymbol{x}+\boldsymbol{y})} = \sqrt{(L^T\boldsymbol{x}+L^T\boldsymbol{y})^T(L^T\boldsymbol{x}+L^T\boldsymbol{y})} = \left\|L^T\boldsymbol{x}+L^T\boldsymbol{y}\right\|_2$$

On the right, we have

$$\|\boldsymbol{x}\|_A + \|\boldsymbol{y}\|_A = \sqrt{\boldsymbol{x}LL^T\boldsymbol{x}} + \sqrt{\boldsymbol{y}LL^T\boldsymbol{y}} = \left\|L^T\boldsymbol{x}\right\|_2 + \left\|L^T\boldsymbol{y}\right\|_2$$

For the 2-norm the triangle inequality holds, i.e.,

$$\left\|L^T\boldsymbol{x}+L^T\boldsymbol{y}\right\|_2 \le \left\|L^T\boldsymbol{x}\right\|_2 + \left\|L^T\boldsymbol{y}\right\|_2,$$

and therefore

$$\|\boldsymbol{x}+\boldsymbol{y}\|_A \le \|\boldsymbol{x}\|_A + \|\boldsymbol{y}\|_A$$

(b) This is not a well-defined matrix norm; the property

$$\|AB\| \le \|A\|\,\|B\|$$

does not necessarily hold. A counter-example is

$$A = B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

with

$$\|AB\| = 5 \qquad \text{and} \qquad \|A\|\,\|B\| = 4.$$

[2] The Cholesky factorisation only appears in Section 5.5, which was not covered at the time of this assignment. Please accept your instructor's apologies.