Alex Gittens
Matrix Theory

June 26, 2006

HOMEWORK 3

Problem 1. For matrices $A \in M_{m,n}$ and $B \in M_{n,m}$, show by direct calculation that $\operatorname{tr} AB = \operatorname{tr} BA$. Use this fact to show that for $A \in M_n$ and nonsingular $S \in M_n$, $\operatorname{tr} SAS^{-1} = \operatorname{tr} A$. The matrix SAS^{-1} is called a similarity of A, and this result says that the trace is a similarity invariant.

Proof. Let A and B be as given, then

$$\operatorname{tr} AB = \sum_{i=1}^{m} (AB)_{ii} = \sum_{i=1}^{m} \sum_{k=1}^{n} A_{ik} B_{ki} = \sum_{i=1}^{n} \sum_{k=1}^{m} A_{ki} B_{ik} = \sum_{i=1}^{n} (BA)_{ii} = \operatorname{tr} BA.$$

Let $S \in M_n^{-1}$, then by this result

$$\operatorname{tr} SAS^{-1} = \operatorname{tr} S(AS^{-1}) = \operatorname{tr}(AS^{-1})S = \operatorname{tr} A.$$

$\square$

Problem 2. Give an example of a diagonalizable matrix $A \in M_n$ that does not have distinct eigenvalues.

Let $D = \operatorname{diag}(1, \underbrace{2, \ldots, 2}_{n-1})$, and let $S \in M_n$ be the elementary matrix which, when multiplied by a matrix A on the right, returns the matrix formed by replacing the first row of A with the sum of its first and second rows and leaves the other rows unchanged:

$$S = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix},$$

then S^{-1} is the elementary matrix which, when multiplied by a matrix A on the right, returns the matrix formed by replacing the first row of A with the difference between the first row and the second row and leaves the other rows unchanged:

$$S^{-1} = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \Rightarrow (S^{-1})^t = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

and $AS^-1 = ((S^{-1})^t A^t)^t$, so mutliplication of a matrix A by S^{-1} on the right returns the transpose of the matrix formed by replacing the first row of A^t with the difference between the second and first rows of A^t. That is, AS^{-1} is the matrix formed by replacing the first column of A with the difference between the second and first columns of A. Therefore

$$SD = \begin{pmatrix} 1 & 2 & 0 & \dots & 0 \\ 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{pmatrix},$$

and

$$A = SDS^{-1} = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{pmatrix},$$

so by construction $A \in M_n$ is a diagonalizable matrix with non-distinct eigenvalues.

Problem 3. Give an example of two commuting matrices that are not simultaneously diagonalizable. Does this contradict Theorem (1.3.12)?

If

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

then $AB = 0 = BA$. But A is not diagonalizable, because $\sigma(A) = \{0\}$, so $S \in M_n^{-1} \Rightarrow SDS^{-1} = 0 \neq A$ when D is the diagonal matrix consisting of the eigenvalues of A. Likewise, B is not diagonalizable, so A and B commute, but are not simultaneously diagonalizable. This does not contradict Theorem (1.3.12), since A and B are not diagonalizable.

Problem 4. If $A \in M_n$ is diagonalizable, consider the characteristic polynomial $p_A(t)$ and show that $p_A(A)$ is the zero matrix.

Proof. Let $p_A = \sum_{k=0}^{n} \alpha_k x^k \in \mathbf{C}[x]$, and $A \in M_n$ be similar to $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, the diagonal matrix of its eigenvalues, under the similarity transform $S \in M_n^{-1}$. Then $A^0 = I = SD^0S^{-1}$, and if $A^{j-1} = SD^{j-1}S^{-1}$, then

$$A^j = ASD^{j-1}S^{-1} = SDS^{-1}SD^{j-1}S^{-1} = SD^jS^{-1},$$

so by induction $A^j = SD^jS$ for all positive integers j. By linearity,

$$p_A(A) = Sp_A(D)S^{-1} = S\operatorname{diag}(p_A(\lambda_1), \ldots, p_A(\lambda_n))S^{-1} = S0S^{-1} = 0,$$

since $\lambda_1, \ldots, \lambda_n$ are the roots of p_A. □

Problem 5. Show that a left eigenvector Y corresponding to the eigenvalue λ of $A \in M_n$ is a right eigenvector of $A^\star$ corresponding to $\overline{\lambda}$ and $\overline{Y}$ is a right eigenvector of A^t corresponding to λ. Show by example that for $A \in M_n(\mathbf{R})$ right and left eigenvectors need not be the same.

Let Y be a left eigenvector of $A \in M_n$ corresponding to λ, then

$$Y^\star A = \lambda Y^\star \Rightarrow (Y^\star A)^\star = (\lambda Y^\star)^\star \Rightarrow A^\star Y = \overline{\lambda} Y,$$

so Y is a right eigenvector of $A^\star$ corresponding to the eigenvalue $\overline{\lambda}$. Also,

$$Y^\star A = \lambda Y^\star \Rightarrow (Y^\star A)^t = (\lambda Y^\star)^t \Rightarrow A^t \overline{Y} = \lambda \overline{Y},$$

so $\overline{Y}$ is a right eigenvector of A^t corresponding to the eigenvalue λ.
Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$, then $\sigma(A) = 0, 1$ with corresponding right eigenvectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. But by the above, left eigenvectors of A are right eigenvectors of $A^t = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$: $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ corresponding to eigenvalue 1 and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ corresponding to eigenvalue 0. In this case, the left and right eigenvectors for each eigenvalue span different spaces, so left and right eigenvectors for this matrix do not coincide.

Problem 6. If $A^\star = A \in M_n$ and if A has distinct eigenvalues show that there are n pairwise orthogonal eigenvectors of A.

Proof. Let $A \in M_n$ be such that A has n distinct eigenvalues and $A^\star = A$. If $\lambda \in \sigma(A)$ has associated eigenvector x, then

$$\begin{aligned}\langle Ax, Ax\rangle &= \langle \lambda x, \lambda x\rangle = \lambda\overline{\lambda}\langle x, y\rangle \\ &= \langle A^\star Ax, x\rangle = \langle AAx, x\rangle = \lambda^2\langle x, x\rangle,\end{aligned}$$

so $\lambda\overline{\lambda} = \lambda^2 \Rightarrow \lambda = \overline{\lambda}$; consequently $\sigma(A) \subset \mathbf{R}$.

Let $\lambda, \nu \in \sigma(A)$ be distinct with associated eigenvectors x and y respectively,

$$\begin{aligned}\langle Ax, Ay\rangle &= \langle \lambda x, \nu y\rangle = \lambda\nu\langle x, y\rangle \\ &= \langle A^\star Ax, y\rangle = \langle AAx, y\rangle = \lambda^2\langle x, y\rangle \\ &= \langle x, A^\star Ay\rangle = \langle x, AAy\rangle = \nu^2\langle x, y\rangle,\end{aligned}$$

so one of the conditions

$$\lambda\nu = \lambda^2 = \nu^2 \text{ or} \tag{1}$$

$$\langle x, y\rangle = 0 \tag{2}$$

is true. Note (1) does not hold for any $\lambda \neq \nu$, because when either λ or ν is zero, the other is nonzero, so $\lambda^2 \neq \nu^2$ and when neither λ or ν is zero, $\lambda\nu = \lambda^2 \Rightarrow \lambda = \nu$, a contradiction. It follows that (2) holds, and the eigenvectors corresponding to the eigenvalues of A are pairwise orthogonal. □