

Assignment 1 - Background review

Due: October 2, in class
No late assignments accepted

Issued: October 1, 2013

Important:

- Give complete answers: Do not only give mathematical formulae, but explain what you are doing. Conversely, do not leave out critical intermediate steps in mathematical derivations.
- Write your **name** as well as your **Sunet ID** on your assignment. **Please staple pages together.** Points will be docked otherwise.
- Questions preceded by * are harder and/or more involved.

Problem 1

Indicate whether the following statements are TRUE or FALSE and motivate your answers clearly. To show a statement false, it is sufficient to give one counter example. To prove a statement true, prove a general proof.

- (a) If $A^2 + A = I$ then $A^{-1} = I + A$
- (b) If all diagonal entries of A are zero, then A is singular (not invertible).

(a) True. We have

$$A^2 + A = I$$

$$A(A + I) = I$$

Similarly,

$$A^2 + A = I$$

$$A + A^2 = I$$

$$(I + A)A = I$$

Thus,

$$A^{-1} = I + A$$

(b) False. A counterexample is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is invertible.

Problem 2

The product of two n by n lower triangular matrices is again lower triangular (all its entries above the main diagonal are zero). Prove it in general and confirm this with a 3 by 3 example.

Suppose $AB = C$ and A and B are lower-triangular. Recall that, by property of matrix multiplication

$$c_{ij} = \vec{r}_i^T \vec{b}_j$$

Consider the i, j element of C , only need to show that this element is necessarily 0 if $i < j$, i.e. above the diagonal, as this is the definition of lower-triangular matrix. Note that for a lower-triangular matrix, the i -th row necessarily has zero after the i -th element and the j -th column also necessarily has zero until the j -th element. Schematically, we have

$$c_{ij} = \underbrace{[a_1 \ a_2 \ \dots \ a_i]}_{i \text{ non-zeros}} \underbrace{[0 \ 0 \ \dots \ 0]}_{n-i \text{ zeros}}^T \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b_{j+1} \\ b_{j+2} \\ \vdots \\ b_n \end{bmatrix} \begin{matrix} \left. \begin{matrix} \dots \\ \dots \\ \dots \end{matrix} \right\} j-1 \text{ zeros} \\ \left. \begin{matrix} \dots \\ \dots \\ \dots \end{matrix} \right\} n-j+1 \text{ nonzeros} \end{matrix} = 0$$

because $j-1 \geq i$ and $n-j+1 \leq n-i$.

Problem 3

If $A = A^T$ and $B = B^T$, which of these matrices are certainly symmetric? A counter-example would be sufficient to prove that the statement is false, but to prove it is true

- (a) $A^2 - B^2$
- (b) $(A + B)(A - B)$
- (c) ABA
- (d) $ABAB$

(a) True

$$(A^2 - B^2)^T = (AA)^T - (BB)^T \quad (1)$$

$$= A^T A^T - B^T B^T \quad (2)$$

$$= A^2 - B^2 \quad (3)$$

(b) False,

$$(A + B)(A - B) = A^2 + BA - AB - B^2 \quad (4)$$

$$[(A + B)(A - B)]^T = [A^2 + BA - AB - B^2]^T \quad (5)$$

$$= (A^2)^T + (BA)^T - (AB)^T - (B^2)^T \quad (6)$$

$$= (A^T)^2 + A^T B^T - B^T A^T - B^T B^T \quad (7)$$

$$= A^2 + AB - BA - B^2 \quad (8)$$

$$(9)$$

But

$AB \neq BA$ in general, (although they can be equivalent, in which case we say A and B commute)

(c) True

$$(ABA)^T = A^T B^T A^T = ABA \quad (10)$$

(d) False

$$(ABAB)^T = B^T A^T B^T A^T = BABA \neq ABAB \text{ because matrices do not commute in general} \quad (11)$$

Problem 4

A *skew-symmetric* matrix is a matrix that satisfies $A^T = -A$. Prove that if A is $n \times n$ and skew-symmetric, then for any \vec{x} , we must have that $\vec{x}^T A \vec{x} = 0$.

Since A is $n \times n$, \vec{x} must be $n \times 1$, which is consistent with our view of all vectors as *column* vectors. Then the product $A\vec{x}$ must also be a column vector of the same size as \vec{x} . Call this column vector \vec{y} . Then

$$\vec{x}^T A \vec{x} = \vec{x}^T (A\vec{x}) = \vec{x}^T \vec{y}$$

But we recognize $\vec{x}^T \vec{y}$ as a scalar product, and so $\vec{x}^T A \vec{x}$ is a scalar quantity. Now, a nice property of scalars is that they equal their own transpose, so for any \vec{x} , we have that

$$\vec{x}^T A \vec{x} = (\vec{x}^T A \vec{x})^T$$

By skew symmetry of A , we then have that

$$\vec{x}^T A \vec{x} = (\vec{x}^T A \vec{x})^T = \vec{x}^T A^T \vec{x} = \vec{x}^T (-A) \vec{x} = -(\vec{x}^T A \vec{x})$$

The only scalar that satisfies this equation is 0.

Problem 5

Suppose A is an $n \times n$ invertible matrix, and you exchange its first two rows to create a new matrix B . Is the new matrix B necessarily invertible? If so, how could you find B^{-1} from A^{-1} ? If not, why not?

As we know, swapping rows does not change the represented system of equations. Swapping the first and second row is equivalent to premultiplying A by an elementary matrix $E = E_{12}$ to get B , where E_{12} is the identity with its first two rows swapped, i.e.,

$$E_{12} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

.

So we write

$$B = E_{12}A,$$

which implies

$$B^{-1} = (E_{12}A)^{-1} = A^{-1}E_{12}^{-1}$$

Now, we know that E_{12} is invertible, because multiplying E_{12} by itself returns the identity matrix. So, $(E_{12})^{-1} = E_{12}$. But now

$$B^{-1} = A^{-1}E_{12}^{-1} = A^{-1}E_{12}.$$

Since we are *post*-multiplying A^{-1} by E_{12} , we are swapping the first two *columns* of A^{-1} to get B^{-1} .

Problem 6

Let A be an invertible n -by- n matrix. Prove that A^m is also invertible and that

$$(A^m)^{-1} = (A^{-1})^m$$

for $m = 1, 2, 3, \dots$

We want to show that for any integer value of m , the inverse of A^m is the inverse of A raised to the m -th power. In other words, we want to show that

$$A^m(A^{-1})^m = (A^{-1})^m A^m = I$$

We prove this by induction on m .

(i) First, we look at the base case, i.e., the case $m = 1$:

$$(A^1)^{-1} = A^{-1} = (A^{-1})^1$$

.

(ii) Next, we *assume* that the result holds for some arbitrary value $m = k$:

$$\text{Assume } (A^k)^{-1} = (A^{-1})^k$$

(iii) Now, using steps i) and ii), we *show* the result holds for $m = k + 1$:

$$(A^{-1})^{k+1} A^{k+1} = A^{-1}(A^{-1})^k A^k A = A^{-1}(A^k)^{-1} A^k A = A^{-1} I A = I$$

and

$$A^{k+1}(A^{-1})^{k+1} = A A^k (A^{-1})^k A^{-1} = A A^k (A^k)^{-1} A^{-1} = A I A^{-1} = I$$

where we have used the result in part (ii). Thus, the general result follows by the induction principle.

Problem 7

Let A be a 2×2 matrix $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ with $a_{11} \neq 0$ and let $\alpha = a_{21}/a_{11}$. Show that A can be factored into a product of the form

$$\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ 0 & b \end{pmatrix}$$

What is the value of b ?

If $\alpha = a_{21}/a_{11}$, then multiplying the matrices in the factorization gives

$$\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ 0 & b \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ \alpha a_{11} & \alpha a_{12} + b \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & \alpha a_{12} + b \end{pmatrix}.$$

This product will equal the original A provided

$$\alpha a_{12} + b = a_{22}$$

Thus, we must have that

$$b = a_{22} - \alpha a_{12} = a_{22} - \frac{a_{21}a_{12}}{a_{11}}$$

Problem 8*

(a) Show that if the $n \times n$ matrices A and B are invertible, and if the matrix $A + B$ is also invertible, then the matrix $B^{-1} + A^{-1}$ is also invertible.

- (b) Assume that C is a skew-symmetric matrix and that D is a matrix defined as

$$D = (I + C)(I - C)^{-1}$$

Prove that $D^T D = DD^T = I$.

(a) We must find a matrix, call it M , such that

$$(A^{-1} + B^{-1})M = I$$

Using the fact that A, B and $A + B$ are invertible, we will do some algebraic manipulations to simplify to something we are given. We write

$$\begin{aligned} I &= (A^{-1} + B^{-1})M \\ &= (A^{-1} + IB^{-1})M \\ &= (A^{-1} + A^{-1}AB^{-1})M \\ &= A^{-1}(I + AB^{-1})M \\ &= A^{-1}(BB^{-1} + AB^{-1})M \\ &= A^{-1}(B + A)B^{-1}M \end{aligned}$$

From the last equation, we see that

$$(A^{-1}(B + A)B^{-1})M = I$$

Therefore, $M = (A^{-1}(A + B)B^{-1})^{-1} = B(A + B)^{-1}A$

Now we verify that multiplying M on the right by the original matrix also gives the identity:

$$\begin{aligned} M(A^{-1} + B^{-1}) &= B(A + B)^{-1}A(A^{-1} + B^{-1}) \\ &= B(A + B)^{-1}(AA^{-1} + AB^{-1}) \\ &= B(A + B)^{-1}(I + AB^{-1}) \\ &= B(A + B)^{-1}(BB^{-1} + AB^{-1}) \\ &= B(A + B)^{-1}(B + A)B^{-1} \\ &= BB^{-1} \\ &= I \end{aligned}$$

Therefore, we have that, indeed $(B^{-1} + A^{-1})^{-1} = B(A + B)^{-1}A$

(b) The trick here is to observe that $(I + C)$ and $(I - C)$ commute, that is,

$$(I + C)(I - C) = (I - C)(I + C)$$

Then

$$\begin{aligned} D^T D &= [(I + C)(I - C)^{-1}]^T (I + C)(I - C)^{-1} \\ &= [(I - C)^T]^{-1} (I + C)^T (I + C)(I - C)^{-1} \\ &= (I - C^T)^{-1} (I + C^T)(I + C)(I - C)^{-1} \\ &= (I + C)^{-1} (I - C)(I + C)(I - C)^{-1} \\ &= (I + C)^{-1} (I + C)(I - C)(I - C)^{-1} \\ &= I \end{aligned}$$

Going from the first line to the second line, we have used properties of transposes as well as the interchanging of transposition and inversion. Going from the second line to the third line, we have used the property for the transpose of a sum of matrices. Going from the third line to the fourth line, we've used the fact that C is skew-symmetric. Going from the fourth line to the fifth line, we're using the fact that $(I + C)$ and $(I - C)$ commute.

To show $DD^T = I$ we can follow similarly:

$$\begin{aligned} DD^T &= (I + C)(I - C)^{-1}[(I + C)(I - C)^{-1}]^T \\ &= (I + C)(I - C)^{-1}[(I - C)^T]^{-1}(I + C)^T \\ &= (I + C)(I - C)^{-1}(I - C^T)^{-1}(I + C^T) \\ &= (I + C)(I - C)^{-1}(I + C)^{-1}(I - C) \\ &= (I + C)[(I + C)(I - C)]^{-1}(I - C) \\ &= (I + C)[(I - C)(I + C)]^{-1}(I - C) \\ &= (I + C)(I + C)^{-1}(I - C)^{-1}(I - C) \\ &= I \end{aligned}$$

where again we have used the fact that $(I - C)$ and $(I + C)$ commute.

Problem 9*

Given an $n \times n$ matrix A with column vectors $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$, construct a matrix B such that the matrix AB has the columns $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$ with the following properties:

- (i) $\vec{u}_i = \vec{a}_i, \quad i \neq j$
- (ii) $\vec{u}_j = \sum_{k=1}^j \alpha_k \vec{a}_k,$

where j is a *fixed* integer such that $1 \leq j \leq n$.

Writing out the matrices in terms of their columns, we have

$$A = [\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n]$$

$$AB = [\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n]$$

$$B = [\vec{b}_1, \vec{b}_2, \dots, \vec{b}_j, \dots, \vec{b}_n]$$

Note that j is a fixed integer indexing a column \vec{b}_j between 1 and n in the matrix B . Aside from \vec{u}_j , the columns of AB and A are the same. Thus, B looks like the identity matrix *except* for column \vec{b}_j . Writing out the j -th column of the matrix AB in terms of the *columns* of A , we have

$$(A\vec{B})_j = \vec{u}_j = b_{1j}\vec{a}_1 + b_{2j}\vec{a}_2 + \dots + b_{jj}\vec{a}_j + \dots + b_{nj}\vec{a}_n$$

Thus, we see that

$$\alpha_k = b_{kj} \text{ for all } k \text{ such that } 1 \leq k \leq j.$$

Here is a schematic of B :

$$B = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \alpha_1 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \alpha_2 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & 1 & \alpha_{j-1} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & 0 & \alpha_j & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}$$