ACM 104 Problem Set 1

Amitesh Anand Pandey

October 2024

Problem 1

If $AB = C$, then from the problem statement we have

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$$

For any column j of C, we have that

$$c_j = \sum_{k=1}^{p} a_k b^{kj}$$

Then, since $C = [c_1, c_2, \dots, c_n]$, substituting the expressions for the columns we get

$$C = \left[\sum_{k=1}^{p} a_k b^{k1}, \sum_{k=1}^{p} a_k b^{k2}, \dots, \sum_{k=1}^{p} a_k b^{kn} \right]$$

Since the summation for each column follows the same form, we can represent C as follows:

$$C = \sum_{k=1}^{p} \left[a_k b^{k1}, a_k b^{k2}, \dots, a_k b^{kn} \right]$$

By definition of matrix multiplication, we can now write

$$C = AB = \sum_{k=1}^{p} a_k b^k$$

Problem 2

(a) Notice that for $a, b \in \mathbb{R}$, we can rewrite the expression as

$$a^{-1} - b^{-1} = a^{-1}(b - a)b^{-1}$$

This leads us to the guess that for matrices, the identity is

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

To prove this, see that $A^{-1}(B - A)B^{-1} = (A^{-1}B - I)B^{-1} = (A^{-1}I - IB^{-1}) = A^{-1} - B^{-1}$ as desired.

(b) First, from the definition of the derivative, we have

$$\frac{\mathrm{d}A^{-1}}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{(A + \Delta A)^{-1} - A^{-1}}{\Delta t}$$

Upon employing the identity derived in 2(a), we obtain

$$\frac{\mathrm{d}A^{-1}}{\mathrm{d}t} = \lim_{\Delta t \to 0} \left[(A + \Delta A)^{-1} \left(-\frac{\Delta A}{\Delta t} \right) A^{-1} \right]$$

Now as Δt goes to 0, trivially ΔA does too. The first coefficient becomes A^{-1} itself, thus finally we have

$$\frac{\mathrm{d}A^{-1}}{\mathrm{d}t} = -A^{-1}A'(t)A^{-1}$$

For the special case where A is a family of functions, recall that the inverse of a scalar output of a function can simply be written by $1/A(t)$. Thus, from the identity above we have

$$\frac{\mathrm{d}}{\mathrm{d}t} A^{-1}(t) = -\frac{1}{A(t)} \cdot A'(t) \cdot \frac{1}{A(t)} = -\frac{1}{A^2} A'(t)$$

On the other hand, from the power and chain rules for differentiation, we also have

$$\frac{\mathrm{d}A^{-1}}{\mathrm{d}t} = \frac{\mathrm{d}A^{-1}}{\mathrm{d}A} \cdot \frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}A} \left(\frac{1}{A(t)} \right) \cdot A'(t) = -\frac{1}{A^2} A'(t)$$

Thus our identity works for the special case of functions.

(c) To find an identity for the derivative of A^2 with respect to t, we once again begin with the definition of the derivative as follows:

$$\frac{\mathrm{d}A^2}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{(A(t + \Delta t))^2 - A(t)^2}{\Delta t}$$

For the sake of notational simplicity, assume that Δt brings about a corresponding ΔA change in the matrix $A(t)$. So now let $A(t + \Delta t) = A(t) + \Delta A$. Note that since A depends smoothly on t, we also have $A(t) + \Delta A \to A$ as $\Delta t \to 0$. Then replacing this expression in the derivative, we have

$$\frac{\mathrm{d}A^2}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{(A + \Delta A)^2 - A^2}{\Delta t} = \lim_{\Delta t \to 0} \frac{A^2 + A \cdot \Delta A + \Delta A \cdot A + \Delta A^2 - A^2}{\Delta t}$$

Since ΔA^2 surely vanishes as $\Delta t \to 0$, we obtain

$$\frac{\mathrm{d}A^2}{\mathrm{d}t} = \lim_{\Delta t \to 0} A\frac{\Delta A}{\Delta t} + \frac{\Delta A}{\Delta t} A = AA'(t) + A'(t)A$$

Problem 3

From running the code, we hypothesize that $P_n = I_n$,

$$l_{ij} = \begin{cases} (1-i)/i, & \text{for } i = j+1 \\ 1, & \text{for } i = j \\ 0, & \text{for elsewhere} \end{cases}$$

$$u_{ij} = \begin{cases} (j+1)/j, & \text{for } i = j \\ -1, & \text{for } i = j-1 \\ 0, & \text{for elsewhere} \end{cases}$$

The above rules indicate that P, L, and U must follow the following patterns:

$$P = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \dots & \\ 0 & \dots & & 1 \end{pmatrix}, \ L = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -\frac{1}{2} & 1 & \dots & 0 \\ \vdots & -\frac{2}{3} & 1 & \vdots \\ 0 & \dots & \frac{1-n}{n} & 1 \end{pmatrix}, \text{ and } U = \begin{pmatrix} \frac{2}{1} & -1 & \dots & 0 \\ 0 & \frac{3}{2} & & \vdots \\ \vdots & & \frac{4}{3} & -1 \\ 0 & \dots & & \frac{n+1}{n} \end{pmatrix}$$

Now, we will proceed with the proof that the above described definitions of P_n, L_n and U_n satisfy $P_nA_n = L_nU_n$.

Proof. First note, if $P_n = I_n$, we have $P_nA_n = I_nA_n = A_n$. So we must show that $A_n = L_nU_n$. Let's try to find L_nU_n. Assume $B_n = L_nU_n$. Then we have

$$b_{ij} = \sum_{k=1}^{n} l_{ik}u_{kj}$$

Consider the follows cases encompassing all b_{ij}:

1. When $i, j = 1$ (main diagonal), we have
$$b_{11} = \sum_{k=1}^{n} l_{11}u_{11} = 2$$
2. When $i = j-1$ (super diagonal), we have
$$b_{ij} = \sum_{k=1}^{n} l_{ik}u_{kj} = -1$$
3. When $i = j$, and $i, j \neq 1$ (main diagonal), we have
$$b_{ij} = \sum_{k=1}^{n} l_{ik}u_{kj} = -\left(\frac{1-i}{i}\right) + \frac{j+1}{j} = \frac{2i}{i} = 2$$
4. When $i = j+1$ (sub diagonal), we have
$$b_{ij} = \left(\frac{1-i}{i}\right) \cdot \left(\frac{j+1}{j}\right) = -1$$
5. Finally, in all other cases, the sum only has '0' terms, thus $b_{ij} = 0$.

Now, note that Cases (1) and (3) ensure that B_n has the same main diagonal as A_n. Cases (2) and (4) ensure that B_n has the same super and sub diagonals as A_n. Case (5) ensures that B_n has all other elements '0,' just like A_n. Thus the product matrix $B_n = A_n$. □

Problem 4

(a) From the definition of an orthogonal matrix, we have that any $P \in \mathbb{M}_{n\times n}$ is orthogonal if $P^T = P^{-1}$. Now notice that this implies that P is orthogonal if and only if $PP^T = PP^{-1} = I_n$. Thus it is sufficient to show that $PP^T = I_n$.

$$(PP^T)_{ij} = \sum_{k=1}^{n} p_{ik}p^T_{kj}$$

Now observe that $p^T_{kj} = p_{jk}$, so we have

$$(PP^T)_{ij} = \sum_{k=1}^{n} p_{ik}p_{jk}$$

Intuitively, we can imagine that $(PP^T)_{ij}$ is filled by iterating through all columns of both rows i and j and taking the sum of pair-wise products of corresponding cells in both of the rows. We know that a single column cannot have a repeat '1' in a permutation matrix, thus $(PP^T)_{ij} = 1$ only when i and j are the same row, so that the '1' cell is multiplied by itself *once*. Then this implies that

$$(PP^T)_{ij} = \begin{cases} 1, & \text{when } i = j \\ 0, & \text{everywhere else} \end{cases}$$

This is exactly the definition of the identity matrix I_n. Thus $P^T = P^{-1}$ and P is orthogonal.

(b) We will prove that an orthogonal matrix need not be a permutation matrix by method of counterexample. Consider the matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Obviously A is not a permutation matrix since it contains values other than 0 and 1. Note that

$$A^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = A$$

Since $A^T = A$, obviously $A^T A^{-1} = AA^{-1} = I$, thus A is orthogonal but not permutational.

Problem 5

Proof. We will have

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

Now, we will show that $\frac{1}{2}(A + A^T)$ is symmetric and that the matrix $\frac{1}{2}(A - A^T)$ is skew-symmetric. Observe that

$$\frac{1}{2}(A + A^T)_{ij} = \frac{1}{2}(A_{ij} + A^T_{ij}) = \frac{1}{2}(A^T_{ji} + A_{ji}) = \frac{1}{2}(A + A^T)_{ji}$$

Thus $1/2(A + A^T) = 1/2(A + A^T)^T$, thus the first term is symmetric. Now notice that

$$\frac{1}{2}(A - A^T)_{ij} = \frac{1}{2}(A_{ij} - A^T_{ij}) = \frac{1}{2}(A^T_{ji} - A_{ji}) = -\frac{1}{2}(A - A^T)_{ji}$$

Thus $1/2(A - A^T) = -1/2(A - A^T)^T$ and the second term is skew symmetric. Hence, A, an arbitrary square matrix, can be expressed as a sum of appropriate symmetric and skew symmeteric matrices. $\square$

Problem 6

(a) Recall that the rank of a matrix is invariant of the elementary row operations applied to it. Thus, we will prove that the number of linearly independent rows spanning B is 2, thereby proving $\text{rank}(B) = 2$ using a sequence of row operations as follows: $R_k : R_k - R_{k-1}$ for $1 < k \leq n$. Then we have

$$B' = \begin{pmatrix} 1 & 2 & \dots & n \\ n & n & \dots & n \\ \vdots & \vdots & & \vdots \\ n & n & \dots & n \end{pmatrix}$$

Clearly that $R_2, R_3, \dots, R_n$ are all identical rows. Then it follows that we have only 2 linearly independent rows spanning B' (or B). So then the $\text{rank}(B) = 2$.

(b) *Code attached in the end*, the non-zero elements of x are $x_{99} = 1.37 \cdot 10^{-14}$ and $x_{100} = 0.01$.

ACM/IDS 104 - Problem Set 1 - MATLAB Problems

Before writing your MATLAB code, it is always good practice to get rid of any leftover variables and figures from previous scripts.

```
clc; clear; close all;
```

NOTE: As this is the first problem set (and many of you might be unfamiliar with MATLAB) we will provide some helper code. As the term progresses (and you become more experienced) we will omit this.

Problem 6 (10 points) Solving Linear Systems

We have the matrix:

$$B = \begin{pmatrix} 1 & 2 & \cdots & n \\ n+1 & n+2 & \cdots & 2n \\ \vdots & \vdots & & \vdots \\ n^2-n+1 & n^2-n+2 & \cdots & n^2 \end{pmatrix}$$

Part (a) (5 points)

In this part, your task is to find $\mathrm{rank}(B)$. As mentioned in the problem set, MATLAB is not needed to obtain the answer. However, we can use MATLAB to make a right guess and check our answer. To do this, we first need to construct matrix B in MATLAB:

NOTE: Although you can check your answer here, you still need to justify and show your reasoning to obtain full credit :)

```
n = 100; % set n as specified in Part (b)
B = 1 : n;
for i = 2 : n
    B(i, :) = B(i-1, :) + n;

end
r = rank(B); % check your answer here
```

Part (b) (5 points)

Set $n = 100$ and consider the system of linear equations $Bx = c$ where $c = (1 \quad 2 \quad \cdots \quad n)^T$. Find a solution x such that its first $[n - \mathrm{rank}(B)]$ components are zero. What are the non-zero components of x?

HINT: The backslash operator $B\backslash c$ issues a warning if B is nearly singular and raises an error condition if it detects exact singularity. In that case, use `pinv(B)*c` for finding a particular solution of $Bx = c$. The function `pinv(B)` returns the "pseudoinverse" of B (will discuss the Moore-Penrose pseudoinverse in lecture 16). Also, the following built-in function may be useful: `null`.

```
%{
Let us start by defining the column vector c as specified above.
Remember that in MATLAB we can use ' to transpose a vector.
```

```
%}

c = (1:n)';
%{
Now, we obtain a particular solution, x_0, as described above
%}
x_0 = pinv(B)*c;

%{
Use null(B) to define the matrix V, whose columns form an orthonormal
basis in the vector space of all solution of the homogeneous
system Bx=0.
%}
V = null(B);

%{
Use the rank, r, found in part (a) to define k = n - rank(B).
This is the number of free variables / dimension of the vector
space.
%}
k = n - r;
```

This is a good point to review what we have done so far. Recall that the desired solution of the system is:

$$x = x_0 + V\alpha \quad (\star)$$

where α is a $k \times 1$ vector. We can obtain α by solving the system:

$$x_0 + V\alpha = 0 \quad (\star\star)$$

```
%{
Find alpha by solving the described system (**).
-> Hint1: Remeber that alpha is a k*1 vector. Hence, you need to
restrict the sizes of x_0 and V
-> Hint2: Use backslash
%}
a = V(1:k, :)\-x_0(1:k);
%{
Finally, put everything together and find x using (*)
Use disp(x) to display your solution.
%}
x = x_0 + V*a
```

```
x = 100×1
   -0.0000
    0.0000
   -0.0000
   -0.0000
    0.0000
   -0.0000
    0.0000
    0.0000
   -0.0000
```

```
   -0.0000
     ⋮
```

```
%{
Now, let us see how x compares to the actual solution.
Un-comment the following 2 lines of code once you reach this part.
%}
error = norm(B*x - c);
disp(error);
```

```
   1.0274e-12
```

Don't forget to report the non-zero components of x!