
ECE 532 Homework 1

Qihong Lu

November 22, 2015

QUESTION 1: MATRIX MULTIPLICATION

a) The matrix can be constructed in the following way:

$$A = \begin{bmatrix} 3 & 2 \\ 4 & 3 \\ 1 & 2 \end{bmatrix}$$

Each row represents the amount of one required ingredients for widget and gizmo. For example, the first row represents the materials needed for widget and gizmo.

On the other hand, each column represents the amount of different ingredients needed for making a widget or a gizmo. For example, the first column represents the the amount of materials, parts and labor needed for making a widget.

b) The vector can be constructed in the following way:

$$x_1 = \begin{bmatrix} 1 \\ 10 \\ 100 \end{bmatrix}$$

Then $x_1^T A$ gives the total cost of making a widget and a gizmo. Here's the computation:

$$x_1^T A = \begin{bmatrix} 1 & 10 & 100 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 143 & 232 \end{bmatrix}$$

c) The vector can be constructed in the following way:

$$x_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Then Ax_2 gives the total materials, parts and labors of making 3 widget and a gizmo. Here's the computation:

$$Ax_2 = \begin{bmatrix} 3 & 2 \\ 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 17 \\ 24 \\ 11 \end{bmatrix}$$

d) the total cost of making 3 widgets and 4 gizmos is given by $x_1^T Ax_2$.

$$x_1^T Ax_2 = \begin{bmatrix} 1 & 10 & 100 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 4 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1357 \end{bmatrix}$$

e) Replicate the computationa in matlab.

```
% set up the matrix and vectors
clear all; clc;
A = [3 4 1; 2 3 2]';
x1 = [1; 10; 100];
x2 = [3;4];

%% 1. cost of making one widget and one gizmo
% linear combination of rows
result1 = zeros(1,size(A,2));
for i = 1 : length(x1)
    % multiply coefficient in x with row of A
    for j = 1 : length(A(i,:))
        result1(j) = result1(j) + x1(i) * A(i,j);
    end
end
% check my answer
if any(result1 ~= x1' * A)
    warning ('wrong!')
end

%% materials, parts and labors needed for making 3 widgets and 4 gizmos
% consider it as a linear combination of columns of A
result2 = zeros(size(A,1),size(x2,2));
for j = 1:length(x2)
    for i = 1 : length(A(:,j))
        result2(i) = result2(i) + A(i,j) * x2(j);
    end
end
% check my answer
if any(result2 ~= A * x2)
    warning ('wrong!')
end

%% total costs of making 3 widgets and 4 gizmos
result3 = 0;
for i = 1:size(x2)
    result3 = result3 + result1(i) * x2(i);
end
% check my answer
if any(result3 ~= x1' * A * x2)
    warning ('wrong!')
end
```

QUESTION2

a) If I can compute this matrix multiplication in the form of a sum of column times row, then it can be seen as a sum of rank-1 matrices.

A column times row is always rank one. Consider vw^T , where v, w are column vector of dimension n . If we denote the component of w in the following form:

A column times row is always rank one. Consider vw^T , where v, w are column vector of dimension n . If we denote the component of w in the following form:

$$w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

Then we can represent vw^T as the following:

$$vw^T = [w_1 v \quad w_2 v \quad \cdots \quad w_n v]$$

Then w_1, w_2, \dots, w_n can be seen as constants, so if I pick any non-zero column, all other columns can be represented that column times a appropriate constant. Therefore, there is only one linearly independent column, and this shows that column times a row is always rank 1.

Now, to express $\frac{X^T X}{n}$ as a sum of column times row, I just multiply the columns in X^T with the rows in X as follows:

$$\frac{X^T X}{n} = \frac{1}{n} \begin{bmatrix} x_{1\bullet}^T & x_{2\bullet}^T & \cdots & x_{p\bullet}^T \end{bmatrix} \begin{bmatrix} x_{1\bullet} \\ x_{2\bullet} \\ \vdots \\ x_{p\bullet} \end{bmatrix} = \frac{1}{n} \sum_{i=1}^n x_{i\bullet}^T x_{i\bullet}$$

In particular, each term $x_{i\bullet}^T x_{i\bullet}$ is a rank-1 matrix, because $x_{i\bullet}^T$ is a column and $x_{i\bullet}$ is a row.

b) By the assumption x_1, x_2, \dots, x_n are linearly independent, we know that $p \geq n$. Because if $p < n$ then those n vectors cannot be linearly independent.

Therefore, $\text{rank}(C) = \text{rank}(X^T X) = n$. Because if all columns X are linearly independent, the same holds true for $X^T X$. (This is proved on the next page)

Proof: if all columns of A are linearly independent, then all columns of $A^T A$ are also linearly independent.

Let v be a vector in the nullspace of $A^T A$, then

$$\begin{aligned}A^T A v &= 0 \\v^T A^T A v &= 0 \\(Av)^T Av &= 0 \\Av \bullet Av &= 0 \\\|Av\| &= 0 \\A\|v\| &= 0 \\v &= 0\end{aligned}$$

This shows that if a vector is in the null space of $A^T A$, it has to be the zero vector. In other words, its null space only has zero vector, which is equivalent to say that all columns in $A^T A$ are linearly independent.

QUESTION3

Let $x, y \in \mathbb{R}^n, \alpha \in \mathbb{R}$. I will show $\Phi(x)$ is a norm by showing the following four properties that characterize a norm:

(i) $\Phi(x) \geq 0, \forall x$

By definition, $\Phi(x) = \sum_{i=1}^n \sum_{j=i+1}^n \max(|x_i|, |x_j|)$, and this is the sum of many terms in the form of $\max(|x_i|, |x_j|)$. which is always non-negative. Therefore, the sum is also going to be always non-negative.

(ii) $\Phi(x) = 0 \iff x = 0$

If x is the zero vector, $\Phi(x) = 0$ will hold true, because there is nothing else to add besides zero.

In addition, if there is any non-zero element in x , $\Phi(x)$ is not going to be zero. This is because the absolute value of any non-zero value is greater than zero, and it will be selected by the max operator.

(iii) $\Phi(\alpha x) = \alpha \Phi(x), \alpha \in \mathbb{R}$

$$\begin{aligned} \Phi(\alpha x) &= \sum_{i=1}^n \sum_{j=i+1}^n \max(|\alpha x_i|, |\alpha x_j|) \\ &= \sum_{i=1}^n \sum_{j=i+1}^n \max(\alpha |x_i|, \alpha |x_j|) \\ &= \alpha \sum_{i=1}^n \sum_{j=i+1}^n \max(|x_i|, |x_j|) \\ &= \alpha \Phi(x) \end{aligned}$$

(iv) $\Phi(x + y) \leq \Phi(x) + \Phi(y)$

By definition, $\Phi(x + y) = \sum_{i=1}^n \sum_{j=i+1}^n \max(|x_i + y_i|, |x_j + y_j|)$

To show that this is equal to $\sum_{i=1}^n \sum_{j=i+1}^n \max(|x_i|, |x_j|) + \sum_{i=1}^n \sum_{j=i+1}^n \max(|y_i|, |y_j|)$, it is enough to test the triangle inequality on one term of this sum.

$$\max(|x_i + y_i|, |x_j + y_j|) \leq \max(|x_i| + |y_i|, |x_j| + |y_j|) \leq \max(|x_i|, |x_j|) + \max(|y_i|, |y_j|)$$

Therefore, the triangle inequality holds.

By (i), (ii), (iii) & (iv), we conclude that $\Phi(x)$ is a norm.

QUESTION4: NORM EQUIVALENCE

(Sorry I am not sure how to use Matlab to plot the unit balls, so the figures are not attached.)

(i) $\alpha \|x\|_1 \leq \|x\|_2 \leq \beta \|x\|_1$

Answer: α is 1 and β is \sqrt{n} .

Explanation:

On 2 dimension, I can draw a the 2 norm unit ball, which is the unit circle. And then I can draw the largest inscribed 1 norm ball, which happens to be the unit 1 norm ball, it is able to touch the 2 norm ball at the corners of the 1 norm ball. So α is 1.

If I draw the smallest circumscribed 1 norm ball. The 1 norm of the point where the 1 norm ball and 2 norm unit ball touches is going to be $\sqrt{2}$. And in general, β is going to be \sqrt{n} .

(ii) $\alpha \|x\|_1 \leq \|x\|_\infty \leq \beta \|x\|_1$

Answer: α is 1 and β is n.

Explanation:

The idea is similar to the idea used in (i). On 2 dimension, I can draw a the infinity norm unit ball, which is the unit square. And then I can draw the largest inscribed 1 norm ball, which happens to be the unit 1 norm ball, it is able to touch the infinity norm ball at the corners of the 1 norm ball. So α is 1.

If I draw the smallest circumscribed 1 norm ball. The 1 norm of the point where the 1 norm ball and 2 norm unit ball touches is going to be 2. And in general, β is going to be n.

(iii) $\alpha \|x\|_1 \leq \Phi(x) \leq \beta \|x\|_1$

Answer: α is $\frac{n-1}{2}$ and β is n-1.

Explanation:

Let $\alpha, \beta \in \mathbb{R}$. Let $x \in \mathbb{R}^n$. And let x be a "ordered" vector such that $x_1 \geq x_2 \geq x_3 \geq \dots \geq x_n$

Then

$$\Phi(x) = (n-1)|x_1| + (n-2)|x_2| + \dots + 0$$

$$\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots |x_n|$$

If I need to find β such that $\Phi(x) \leq \beta \|x\|_1$, then $\beta = n-1$.

To find the lower bound, I need to let all components in x to be the same value. Therefore,

$$\Phi(x) = (n-1)|x_1| + (n-2)|x_2| + \dots + 0 = \frac{n(n-1)}{2} \|x\|_1$$

$$\|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots |x_n| = \alpha n \|x\|_1$$

When $\Phi(x)$ and $\|x\|_1$ are equal,

$$\alpha n \|x\|_1 = \frac{n(n-1)}{2} \|x\|_1$$

$$\alpha = \frac{n-1}{2}$$

QUESTION5

a) We know that $y = Ax$, and I want to find an expression in terms of A and y .

$$\begin{aligned}y &= Ax \\ A^{-1}y &= A^{-1}Ax \\ A^{-1}y &= x \\ x &= A^{-1}y\end{aligned}$$

b) Bound the 2-norm of x in terms of $\|y\|_2$ and a function of the matrix A .
Here's a upper bound:

$$\begin{aligned}x &= A^{-1}y \\ \|x\|_2 &= \|A^{-1}y\|_2 \leq \|A^{-1}\|_2 \cdot \|y\|_2\end{aligned}$$

Here's a lower bound:

$$\begin{aligned}y &= Ax \\ \|y\|_2 &= \|Ax\|_2 \leq \|A\|_2 \cdot \|x\|_2 \\ \|y\|_2 &\leq \|A\|_2 \cdot \|x\|_2 \\ \|x\|_2 &\geq \frac{\|y\|_2}{\|A\|_2}\end{aligned}$$

To summarize:

$$\frac{\|y\|_2}{\|A\|_2} \leq \|x\|_2 \leq \|A^{-1}\|_2 \cdot \|y\|_2$$

QUESTION6

a) To find the rank of A, I need to find out how many linearly independent columns or rows A has. I begin with doing some row operations, including exchanging rows and row reductions:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

So A can be transformed to a 3 by 3 identity matrix, which form the standard basis and we know they are linearly independent. It follows that A is also linearly independent. Therefore, A is full rank, and $\text{rank}(A) = 3$.

b) I want to derive a formula for x in terms of y, based on $y = Ax$.

$$\begin{aligned} y &= Ax \\ A^{-1}y &= A^{-1}Ax \\ x &= A^{-1}y \end{aligned}$$

So I just need to figure out what is A^{-1} . This can be compute by going from $[A|I]$ to $[I|B]$, then $A^{-1} = B$. Here's the procedure:

$$[A|I_3] = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{bmatrix} = [I_3|B]$$

Therefore,

$$A^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

In conclusion, here's the formula of x in terms of y

$$x = A^{-1}y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} y$$

```
%Here's the matlab code that verify the computations.
A = [1 1 1; 1 1 0; 1 0 0];
rank(A)
inv(A)
```

QUESTION 7

a)

To find the rank, I need to find the number of linearly independent columns.

$$X = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

After exchanging the third and fourth rows, and subtract the first row from the third, X is in row reduced echelon form. The first three columns are clearly independent.

Additionally, from the results of c), we know that there are three linearly independent columns, so $\text{rank}(X) = 3$.

b)

$$XX^T = \begin{bmatrix} 3 & 1 & 0 & 2 \\ 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 1 & 2 & 0 \\ 1 & 3 & 2 & 0 \\ 2 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ignore the zero row and column,

$$\begin{bmatrix} 3 & 1 & 2 \\ 1 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix} \rightarrow \dots \rightarrow \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 1/4 \\ 0 & 0 & 1/2 \end{bmatrix}$$

This is in upper triangular form, it is easy to check that these three vectors are linearly independent. So the rank of the original matrix XX^T is 3.

c) Finding a linearly independent columns in X. Let's call of columns of X by v_1, v_2, \dots, v_5 . Then:

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, v_5 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Then v_4 and v_5 can be expressed in terms of v_1, v_2, v_3 :

$$v_4 = v_2 + v_3 - v_1,$$

$$v_5 = v_2 + v_3$$

And there is no way of expressing any vector among v_1, v_2, v_3 using the rest, so they are three independent column vectors.

```
%Here's the matlab code that verify the computations.  
X = [1 0 1 0 1; 0 1 0 1 1; 0 0 0 0 0; 0 0 1 1 1];  
rank(X)  
rank(X*X')
```