18.06 **Hints and Answers to Problem Set 5**

1. First, we introduce some notation. We write the matrix A as

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \mathbf{a}^{(1)} & \mathbf{a}^{(2)} & \cdots & \mathbf{a}^{(n)} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{(1)} \\ \mathbf{a}_{(2)} \\ \vdots \\ \mathbf{a}_{(m)} \end{pmatrix}.$$

Thus, we denote the columns of A by $\mathbf{a}^{(i)}$ and the rows by $\mathbf{a}_{(i)}$ and the notation is analogous for B.

(a) Let $\mathbf{v} = (v_1, v_2, \cdots, v_n)$. Then

$$\begin{aligned} \mathbf{v}B &= (v_1 b_{11} + v_2 b_{21} + \cdots + v_n b_{n1}, \cdots, v_1 b_{1p} + v_2 b_{2p} + \cdots + v_n b_{np}) \\ &= v_1(b_{11}, b_{12}, \cdots, b_{1p}) + \cdots + v_n(b_{n1}, b_{n2}, \cdots, b_{np}) \\ &= v_1 \mathbf{b}_{(1)} + \cdots + v_n \mathbf{b}_{(n)}. \end{aligned}$$

Thus, we see that $\mathbf{v}B$ is a simple linear combination of rows of B. Now consider,

$$\begin{aligned} A\mathbf{v} &= \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{pmatrix} = v_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + v_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= v_1 \mathbf{a}^{(1)} + \cdots + v_n \mathbf{a}^{(n)} \end{aligned}$$

Hence, we see that $A\mathbf{v}$ is a linear combination of columns of A.

(b) Let us write the product AB as follows,

$$AB = \begin{pmatrix} \mathbf{a}_{(1)} \\ \mathbf{a}_{(2)} \\ \vdots \\ \mathbf{a}_{(m)} \end{pmatrix} \begin{pmatrix} \mathbf{b}_{(1)} \\ \mathbf{b}_{(2)} \\ \vdots \\ \mathbf{b}_{(m)} \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{(1)}B \\ \mathbf{a}_{(2)}B \\ \vdots \\ \mathbf{a}_{(m)}B \end{pmatrix}.$$

So each row of AB is a vector product of a row vector times B. From part (a), we know that such a vector product is just a linear combination of rows of B. Hence, the rows of AB are linear combinations of rows of B. So every vector which is a linear combination of rows of AB is also a linear combination of rows of B. Hence, the row space of AB is contained in the row space of B.

(c) The rank of a matrix is the dimension of its row space. We need to show that if row space $(AB) \subset$ row space of B, then dim row space $(AB) \leq$ dim row space (B). Pick a basis for row space (AB). This is a set of linearly independent vectors, which span row space (AB). Say there are p of them. Now, since the row space of AB is contained in that of B, all of these p vectors are also in the row space of B. Furthermore, they are linearly independent. Since the dimension of a space is equal to the maximum number of linearly independent vectors in that space, and there are at least p linearly independent vectors in the row space of B, the dimension of the row space of B must be at least p, dim row space $(B) \geq p =$ dim row space (AB). Thus, rank $(AB) \leq$rank (B).

(d) Since A is an $m \times n$ matrix, A^T is $n \times m$, and $A^T A$ is $n \times n$. We are given that $m \geq n$. The statement that $A^T A$ is nonsingular is thus equivalent to saying that $A^T A$ has rank n. The maximum rank that A can have is n, because it only has n columns. Using the result of part (c), we have that

$$\mathrm{Rk}(A^T A) \leq \mathrm{Rk}(A) \leq n.$$

Hence, if $\mathrm{Rk}(A^T A) = n$, the inequality implies that $\mathrm{Rk}(A) = n$. For the proof that $\mathrm{Rk}(A) = n$ implies that $\mathrm{Rk}(A^T A) = n$ also, we can borrow the argument from the proof of **4G** on page 180 of Strang's book. We know that since $\mathrm{Rk}(A) = n$, the dimension of the nullspace of A is 0. We only need to show that an element of the nullspace of $A^T A$ is also an element of the null space of A, to conclude that the only element in the null space of $A^T A$ is $\mathbf{0}$ ($\mathbf{0}$ is always an element of the nullspace). So suppose that $A^T A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x}$. Multiplying both sides of this equation by $\mathbf{x}^T$, we have

$$(\mathbf{x}^T)A^T A\mathbf{x} = 0 \Leftrightarrow (A\mathbf{x})^T(A\mathbf{x}) = 0 \Leftrightarrow ||A\mathbf{x}||^2 = 0.$$

So the vector $A\mathbf{x}$ has length 0, and hence $A\mathbf{x} = \mathbf{0}$.

5.

```
>> A=[-1 1 0 0 0 0;-1 0 1 0 0 0;-1 0 0 1 0 0;-1 0 0 0 1 0;-1 0 0 0 0 1]

A =

    -1     1     0     0     0     0
    -1     0     1     0     0     0
    -1     0     0     1     0     0
    -1     0     0     0     1     0
    -1     0     0     0     0     1

>> a=null(A,'r')

a =

     1
     1
     1
     1
     1
     1

>> P = a*(inv(a'*a))*a'
```

```
P =

    0.1667    0.1667    0.1667    0.1667    0.1667    0.1667
    0.1667    0.1667    0.1667    0.1667    0.1667    0.1667
    0.1667    0.1667    0.1667    0.1667    0.1667    0.1667
    0.1667    0.1667    0.1667    0.1667    0.1667    0.1667
    0.1667    0.1667    0.1667    0.1667    0.1667    0.1667
    0.1667    0.1667    0.1667    0.1667    0.1667    0.1667

>> b = [1 2 3 4 5 6]'

b =

     1
     2
     3
     4
     5
     6

>> p = P*b

p =

    3.5000
    3.5000
    3.5000
    3.5000
    3.5000
    3.5000

>> e = p-b

e =

    2.5000
    1.5000
    0.5000
   -0.5000
   -1.5000
   -2.5000

>> e'*a

ans =
```

```
     0

>> diary off
```