Review for Exam 2

The exam will cover the material from HW4, 5, 6 and the corresponding Course Notes. First we need to remember the properties of matrix multiplication.

Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix then the $m \times p$ product matrix AB satisfies

- $(i,j)^{th}$ entry of $AB = (i^{th}$ row $A)(j^{th}$ col $B)$
- $(i^{th}$ row of $AB) = (i^{th}$ row $A)\,B$
- $(j^{th}$ col of $AB) = A\,(j^{th}$ col $B)$.

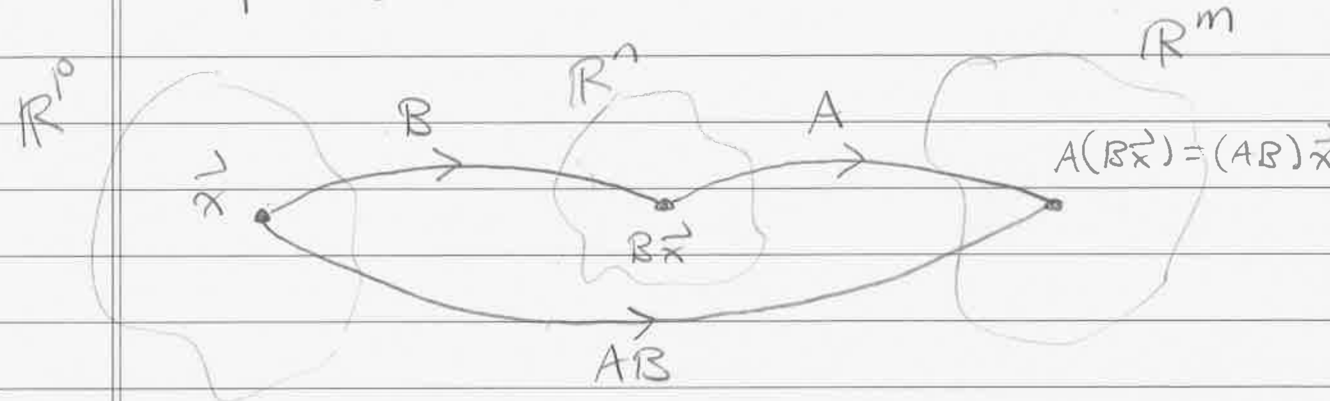
[Note that (i^{th} row A) is a $1 \times n$ matrix and (j^{th} col B) is an $n \times 1$ matrix. Their matrix product is the same as the good old dot product of vectors:

$(i\text{th row } A)(j\text{th col } B) = (i\text{th row } A)^T \cdot (j\text{th col } B)$

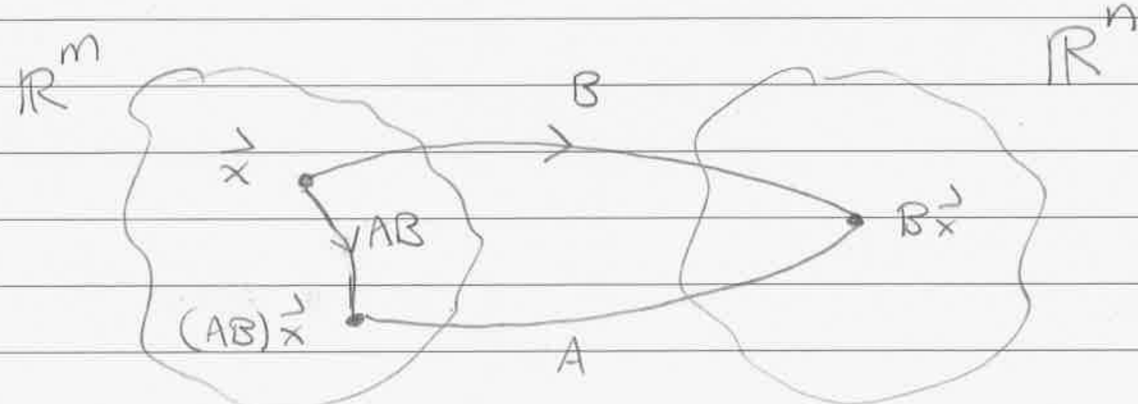
$(1\times n \text{ matrix})(n\times 1 \text{ matrix}) = (\text{vector})\cdot(\text{vector})$

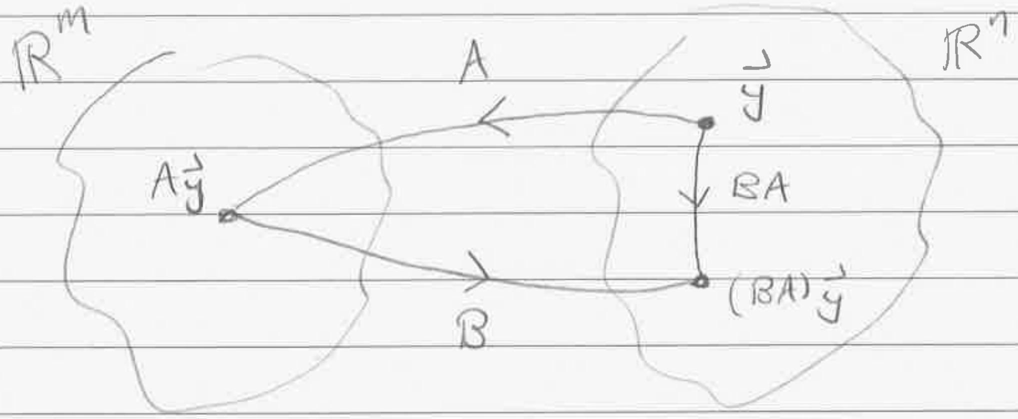
$1\times 1 \text{ matrix} = \text{number}$

The matrix AB was originally <u>defined</u> by composing two functions:



In other words we have $(AB)\vec{x} = A(B\vec{x})$ for all $p\times 1$ vectors $\vec{x}$. Now if we have $p = m$ then we can also define the matrix BA and we obtain two pictures:





If both of AB & BA are "do nothing" functions, i.e., if

$$AB = I_m \quad \& \quad BA = I_n$$

then we will say that A & B are <u>inverses</u> of each other.

But I discussed in class that this is <u>impossible</u> when $m \neq n$. [Idea: If $m < n$ then RREF(A) has a non-pivot column which implies that A has a non-trivial row relation. But then if $BA = I_n$ then the formula

$$(j^{th} \text{ col } I_n) = B(j^{th} \text{ col } A)$$

implies that I_n has a non-trivial column relation, which is <u>impossible</u>.]

If $m=n$ then then the $n \times n$ matrix A might have an inverse. Suppose it does, i.e., suppose that there exists an $n \times n$ matrix B such that

$$AB = I_n \quad \& \quad BA = I_n.$$

To compute this matrix B we will solve the linear systems

$$A(j^{th} \text{ col } B) = (j^{th} \text{ col } I_n)$$

to get the columns of B. All of these systems can be solved simultaneously with a trick:

$$(*) \quad (A \mid I_n) \xrightarrow{RREF} (I_n \mid B).$$

[Remark: The matrix B is unique so we call it "the" inverse of A and we write $B =$ "A^{-1}". Indeed, if we also have $AC = I_n$ & $CA = I_n$ then it follows that

$$C = CI_n = C(AB) = (CA)B = I_nB = B.$$]

If we try to compute A^{-1} using (*) and it fails this means that A is not invertible. The reason (*) is because

$$\mathrm{RREF}(A) \neq I_n$$

and there are many things that can cause this, e.g.,

- A has a non-trivial column relation
- A has a non-trivial row relation
- $\det(A) = 0$.

You do not need to know the "Fundamental Theorem of Linear Algebra" for the exam. However, you do need to know the basic properties of matrix arithmetic, e.g.,

$$A(xB + yC) = xAB + yAC$$

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

etc.

We discussed "orthogonal projection" and its applications to "least squares regression".

Suppose we want to find the line $C + tD = b$ that is <u>closest</u> to the data points

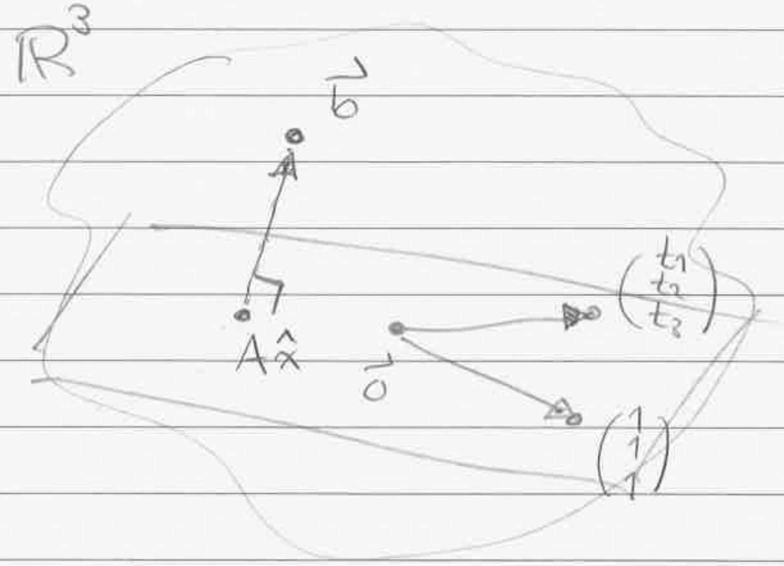
$$\begin{pmatrix} t_1 \\ b_1 \end{pmatrix}, \begin{pmatrix} t_2 \\ b_2 \end{pmatrix}, \begin{pmatrix} t_3 \\ b_3 \end{pmatrix}.$$

First we write down the silly equations

$$\begin{cases} C + t_1 D = b_1 \\ C + t_2 D = b_2 \\ C + t_3 D = b_3 \end{cases} \rightsquigarrow \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ 1 & t_3 \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

$$\text{“} A\vec{x} = \vec{b} \text{”}.$$

This equation probably has no solution because the three points probably don't lie on a line. Alternatively, this means that the point $\vec{b}$ in $\mathbb{R}^3$ does not lie in the plane $A\vec{x} = C(1,1,1) + D(t_1, t_2, t_3)$ which is the "column space" of the matrix A:



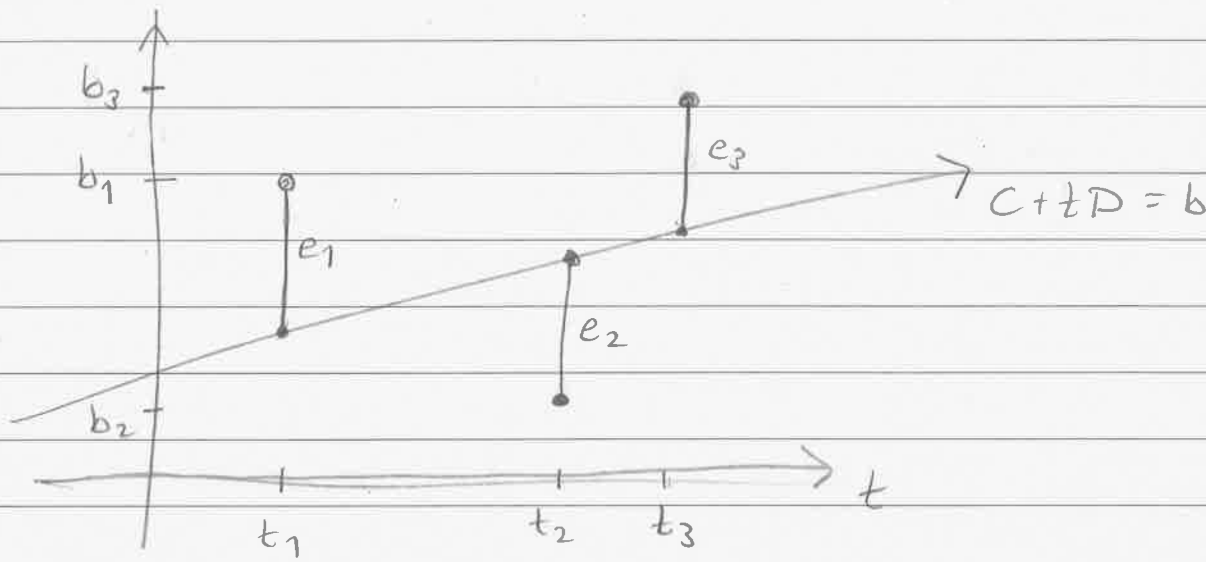
Gauss' idea is to replace $\vec{b}$ by the closest point $A\hat{x}$ in the column space. This will be accomplished when the error vector $\vec{e} := \vec{b} - A\hat{x}$ is perpendicular to all of the columns of A (the plane in the picture). We can express this condition with one matrix equation

$$(*) \quad A^T \vec{e} = \vec{0}$$

After substituting $\vec{e} = \vec{b} - A\hat{x}$, this $(*)$ becomes the "normal equation"

$$A^T A \hat{x} = A^T \vec{b}.$$

Now we can solve this to find $\hat{x} = (C, D)$ and hence the best fit line:



The vertical errors are the entries of the error vector:

$$\vec{e} = \vec{b} - A\hat{x} = \begin{pmatrix} b_1 - (C + t_1 D) \\ b_2 - (C + t_2 D) \\ b_3 - (C + t_3 D) \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

and the best fit line is "best" in the sense that

$$\|\vec{e}\|^2 = e_1^2 + e_2^2 + e_3^2$$

is as small as possible.

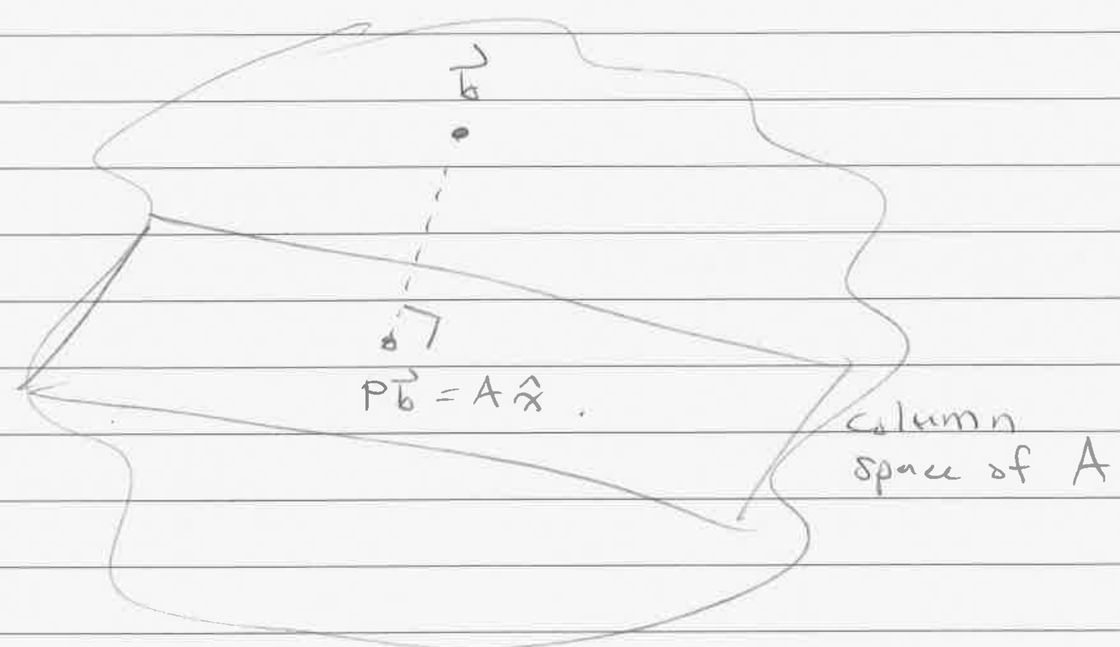
Least squares regression applies to a much broader range of problems. Given (almost) any linear system $A\vec{x} = \vec{b}$ with no solution, you can find the "best approximate solution" $\hat{x}$ by solving

$$A^T A \hat{x} = A^T \vec{b}.$$

The geometry behind this is to project $\vec{b}$ orthogonally onto the column space of A by using the projection matrix

$$P = A(A^T A)^{-1} A^T.$$

Picture:



More Review

① Projection / Least Squares

② Eigenvalues / vectors.

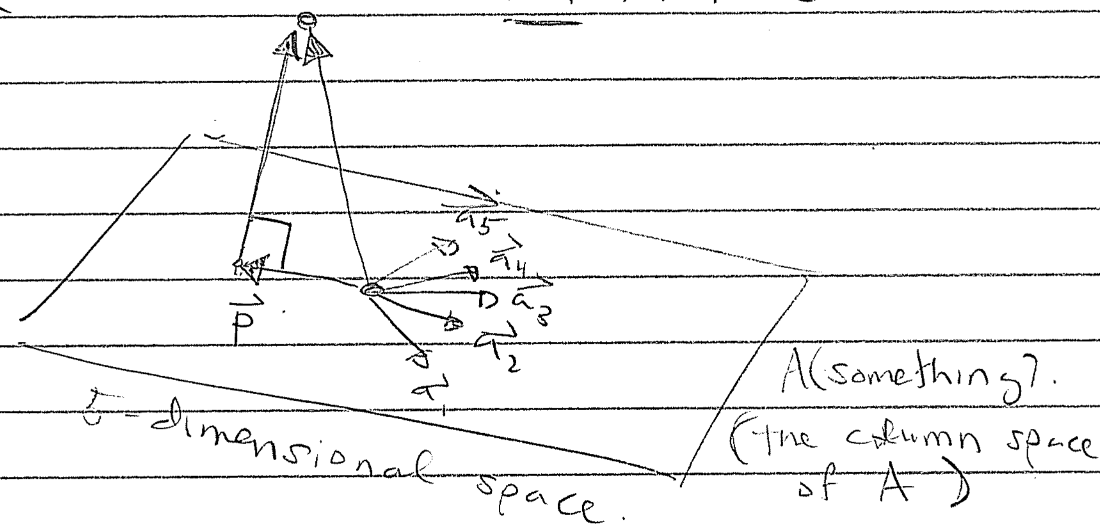
① Start with $A\vec{x} = \vec{b}$.

Suppose it has no solution $\vec{x}$, i.e. $\vec{b} \neq A(\text{something})$.

What does A(something) look like.
If $A = (\vec{a}_1\ \vec{a}_2\ \cdots\ \vec{a}_5)$ (5 columns).
Then

$$A\vec{x} = A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_5 \end{pmatrix} = \underbrace{x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_5\vec{a}_5}_{\text{linear combination of the columns.}}$$

Picture: $\vec{b}$ is not in there



Find the point $\vec{p}$ in the space, closest to $\vec{b}$.

(i) Since $\vec{p}$ is in the space we have

$$\vec{p} = A(\text{something}) = A\hat{x}.$$

Goal: Find $\hat{x}$.

(ii) Since $\vec{p}$ is closest to $\vec{b}$, the error $\vec{e} = \vec{b} - \vec{p} = \vec{b} - A\hat{x}$ is orthogonal to the space.

i.e. $\vec{e}$ is orthogonal to each of $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_5$

i.e. $$\vec{a}_1^T \vec{e} = 0$$
$$\vec{a}_2^T \vec{e} = 0$$
$$\vdots$$
$$\vec{a}_5^T \vec{e} = 0.$$

i.e. $$A^T \vec{e} = \vec{0}.$$

put (i) and (ii) together

$$A^T \vec{e} = \vec{0}$$
$$A^T(\vec{b} - A\hat{x}) = \vec{0}$$
$$A^T\vec{b} - A^T A\hat{x} = \vec{0}$$
$$A^T\vec{b} = A^T A\hat{x}.$$

Now we can <u>solve</u> for $\hat{x}$:

Since $(A^T A)$ is invertible

$$\Longrightarrow \hat{x} = (A^T A)^{-1} A^T \vec{b}.$$

Now we can solve for the projection $\vec{p}$.

$$\vec{p} = A\hat{x} = \underbrace{A(A^T A)^{-1} A^T}_{\text{the projection matrix.}} \vec{b}.$$

Easiest Case: Project onto a <u>line</u> $t\vec{a}$.

Then $A = \vec{a}$ is just a column vector.

The projection matrix is

$$P = \vec{a}(\vec{a}^T\vec{a})^{-1}\vec{a}^T$$

But $\vec{a}^T\vec{a}$ is a 1×1 matrix (i.e. a number),

so $(\vec{a}^T\vec{a})^{-1} = \dfrac{1}{\vec{a}^T\vec{a}}$. (Easy).

Hence

$$P = \vec{a}(\vec{a}^T\vec{a})^{-1}\vec{a}^T = \underbrace{\frac{1}{\vec{a}^T\vec{a}}}_{\text{number}}\underbrace{(\vec{a}\vec{a}^T)}_{\text{matrix}}.$$

② <u>Eigen</u>.

Let A be square. We say $\vec{x}$ is an eigenvector if

- $\vec{x} \neq \vec{0}$
- $A\vec{x} = \lambda\vec{x}$ for some number λ (the eigenvalue).

Q: Which numbers λ could be eigenvalues?

We need
$$A\vec{x} = \lambda\vec{x}$$
$$A\vec{x} = \lambda I\vec{x}$$
$$A\vec{x} - \lambda I\vec{x} = \vec{0}.$$
$$(A - \lambda I)\vec{x} = \vec{0}.$$

for some $\vec{x} \neq \vec{0}$. In other words, the matrix $A - \lambda I$ must be singular (i.e. non-invertible).

For 2×2 matrices this is easy to say:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$$

is singular $\Leftrightarrow$ the rows are parallel

$$\Leftrightarrow \quad \frac{a-\lambda}{b} = \frac{c}{d-\lambda}$$

$$\Leftrightarrow \quad \underbrace{(a-\lambda)(d-\lambda) - bc = 0}_{\substack{\text{The characteristic equation.} \\ \text{The solutions are the eigenvalues}}}.$$

[Remark: Another language says

$$\det(A - \lambda I) = 0.$$]

Once you get the eigenvalues, the eigenvectors are easy:

Just solve $(A - \lambda I)\vec{x} = \vec{0}$ using your favorite method.

Q: Who Cares?

If you want to solve a linear recurrence

$$\vec{V}_{n+1} = A\vec{V}_n,$$

(i) Find the eigenvectors $A\vec{x}_1 = \lambda_1 \vec{x}_1$

$A\vec{x}_2 = \lambda_2 \vec{x}_2$

(ii) Express $\vec{V}_0 = s\vec{x}_1 + t\vec{x}_2$.

(iii) The solution is $\vec{V}_n = A^n \vec{V}_0$

$$= sA^n\vec{x}_1 + tA^n\vec{x}_2 = s(\lambda_1)^n\vec{x}_1 + t(\lambda_2)^n\vec{x}_2.$$

Enjoy!