EML 6934

optimal estimation, Fall '09

1.1 (i) Let $x, y \in S$, $\Rightarrow$ $x_1 - x_2 = 0$
$y_1 - y_2 = 0$

consider $z := x + y$ then $z_1 - z_2 = (x_1 + y_1) - (x_2 + y_2) = 0$

consider $z = \alpha x$, α is a scalar. Then $z_1 - z_2 = \alpha x_1 - \alpha x_2 = 0$

$\therefore$ S is closed under addition and scalar multiplication.

$\Rightarrow$ S is a vector space.

1.1 (ii) Let $x, y \in S$ $\Rightarrow$ $x_2 - x_1 = 1$, $y_2 - y_1 = 1$

Consider $z = \alpha x$ $\Rightarrow$ $z_2 - z_1 = \alpha(x_2 - x_1) = \alpha \neq 1$ unless $\alpha = 1$

$\therefore$ S is not closed under scalar multiplication, so S is not a vector space. [key: A subspace must contain the 0 vector!]

1.2 Let f and g be two real valued continuous functions defined in the interval $[0, 1]$

To show closure under addition, we have to show that the sum of two continuous functions is continuous.

To do so, start from the definition of a continuous function. A function η is continuous at a point, say, x_0, if, for every $\epsilon > 0$, there exists $\delta > 0$ such that $|\eta(x) - \eta(x_0)| < \epsilon$ whenever $|x - x_0| < \delta$.

By using this definition, we can say the following about f and g at every point $x_0 \in [0, 1]$:

$\forall \epsilon > 0, \exists \delta_{2,1} > 0$ so that $|f(x) - f(x_0)| < \epsilon/2$ whenever $|x - x_0| < \delta_1$

and $|g(x) - g(x_0)| < \epsilon/2$ whenever $|x - x_0| < \delta_2$

$$\Rightarrow |(f+g)(x) - (f+g)(x_0)| = |f(x) + g(x) - f(x_0) - g(x_0)|$$

$$\leq |f(x) - f(x_0)| + |g(x) - g(x_0)| \quad \text{(triangle inequality)}$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad \text{whenever } |x - x_0| < \min(\delta_1, \delta_2)$$

$\Rightarrow$ $f+g$ is a continuous function at x_0. The same argument can be repeated at every point $x_0 \in [0, 1]$, which shows $f+g$ is continuous in the interval $[0, 1]$.

This shows closure under addition. A similar argument, which we are not going to repeat, proves closure under scalar multiplication.

Problem 2

$$Ax = \begin{bmatrix} 1 & 2 \\ 5 & 4 \\ 13 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 5x_1 + 4x_2 \\ 13x_1 + 6x_2 \end{pmatrix} = \begin{bmatrix} 1 \\ 5 \\ 13 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} x_2$$

Problem 3 ① if $y_1, y_2 \in R(A)$, there exist vectors x_1, x_2 s.t.

$$y_1 = Ax_1, \quad y_2 = Ax_2$$

$$\Rightarrow y_1 + y_2 = Ax_1 + Ax_2 = A(x_1 + x_2) = Az, \quad z \triangleq x_1 + x_2$$

$$\therefore y_1 + y_2 \in R(A)$$

Similarly, $\alpha y_1 = \alpha A x_1 = A(\alpha x_1) = Az \Rightarrow z = \alpha x$

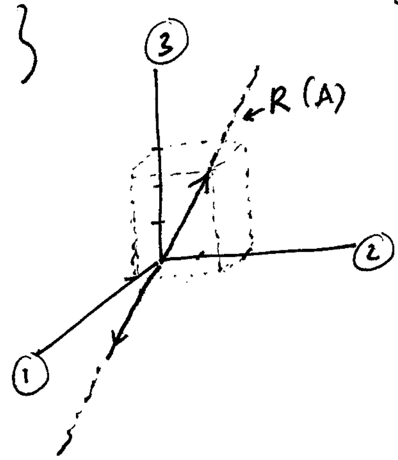
$$\Rightarrow \alpha y_1 \in R(A)$$

This shows $R(A)$ is closed under addition and scalar multiplication, so it is a vector space.

$R(A)$, $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$

$Ax = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = (x_1 + 2x_2) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, so Ax is a vector along $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$\Rightarrow R(A) = \left\{ y \mid y = \alpha \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ \alpha \text{ any scalar} \right\}$



3.2 Let $x_1, x_2 \in N(A)$

$\Rightarrow Ax_1 = 0, \ Ax_2 = 0$

Now, $A(x_1 + x_2) = Ax_1 + Ax_2 = 0$

which shows $x_1 + x_2 \in N(A)$

similarly, $\alpha x_1 \in N(A)$, for every scalar α

$\Rightarrow$ $N(A)$ is closed under addition and scalar multiplication.

and therefore a vector space.

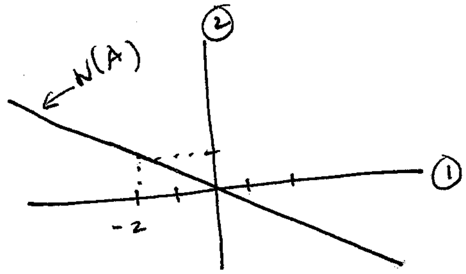
to find $N\left(\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}\right)$, examine $Ax = 0$

$\begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow (x_1 + 2x_2) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 0 \Rightarrow x_1 + 2x_2 = 0$

$\Rightarrow x_1 = -2x_2$

$\Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} x_2$ for x to be in $N(A)$

$\Rightarrow N(A) = \left\{ \alpha \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \ \alpha \text{ any scalar} \right\}$



Problem 4 $\quad span(L) = R(A)$. reason :

$$Ax = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$ (using the useful fact from Problem 2)

$$= \sum_{i=1}^{n} x_i a_i$$

as x takes all possible values, Ax creates all the possible vectors in the set $span(L) = \{ y \mid y = \sum_{i=1}^{n} x_i a_i \}$,

where $L = \{ a_1, a_2, \cdots, a_n \}$

Problem 5 (5.1) S is a set of vectors, $S^\perp = \{ y \mid y^T x = 0 \quad \forall x \in S \}$

Let $y_1, y_2 \in S \Rightarrow y_1^T x = y_2^T x = 0 \quad \forall x \in S$

then $(y_1 + y_2)^T x = y_1^T x + y_2^T x = 0 + 0 = 0 \quad \forall x \in S$

$\Rightarrow y_1 + y_2 \in S^\perp$

Similarly, $(\alpha y_1)^T x = \alpha y_1^T x = 0 \quad \forall x \in S$

$\Rightarrow \alpha y_1 \in S^\perp$ as well, where α is any scalar.

This shows that $S^\perp$ is closed under addition and scalar multiplication, ~~&~~ which shows $S^\perp$ is a vector space. (Note that S need not be a vector space for $S^\perp$ to be one).

(5.2) No, since the vector $v = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in$ "x-y plane" and also "y-z plane"

but $v^T v \neq 0$

5.3

$$R(A)^{\perp} = \{x \mid x^T y = 0 \quad \forall y \in R(A)\}$$

$$= \{x \mid x^T (Az) = 0 \quad \underline{\forall z \in \mathbb{R}^n}\}$$

(Assume, for concreteness, that $A: \mathbb{R}^n \to \mathbb{R}^m$
so $x \in \mathbb{R}^m$)

$$= \{x \mid z^T A^T x = 0 \quad \forall z \in \mathbb{R}^n\}$$

$$= \{x \mid A^T x = 0\}$$

(if $z^T (A^T x) = 0 \ \forall z$, then $A^T x$ must be the 0 vector,

$\because$ we can pick $z = A^T x$, which gives us

$(A^T x)^T (A^T x) = 0 \iff y^T y = 0 \ (y \triangleq A^T x) \iff y_1^2 + y_2^2 + \cdots + y_n^2 = 0$

$\iff y_i = 0 \ \forall i \iff A^T x = 0$)

$$= N(A) \quad \text{— Q.E.D.}$$

(or, think of it this way,

if $x \in R(A)^{\perp}$, $x \perp y \ \forall y \in R(A) \iff x^T (Az) = 0 \quad \forall z \in \mathbb{R}^n$

$\iff z^T A^T x = 0 \quad \forall z \in \mathbb{R}^n$

$\iff A^T x = 0$

by the arguments above

$\Rightarrow x \in R(A)^{\perp} \iff A^T x = 0$, which means $R(A)^{\perp} = N(A^T)$)

(Problem 5.3 contd.)

We have shown $R(A)^{\perp} = N(A^T)$ — ①

to show $N(A)^{\perp} = R(A^T)$, take '$\perp$' of both sides of ①.

$$\left(R(A)^{\perp}\right)^{\perp} = \left(N(A^T)\right)^{\perp}$$

$$\Leftrightarrow \quad R(A) = N(A^T)^{\perp}$$

replace A by A^T to get

$$R(A^T) = N(A)^{\perp} \quad \text{— Q.E.D.}$$

Problem 6:
6.1

$$Ax = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 2 & 4 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

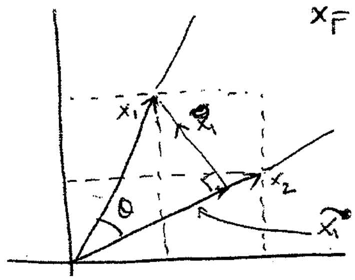
$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix} x_1 + \begin{bmatrix} 2 \\ 4 \end{bmatrix} x_2 + \begin{bmatrix} 2 \\ 1 \end{bmatrix} x_3 + \begin{bmatrix} 4 \\ 2 \end{bmatrix} x_4 \quad \text{— ①}$$

$$= (x_1 + 2x_2)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}(x_3 + 2x_4) \quad \text{— ②}$$

① shows, (we have seen it before) that $R(A)$ is spanned by the four vectors that are the columns of A. This is true in general, $R(A)$ is spanned by the columns of A. But in this example, not all the column of A are linearly independent.

(problem 6 contd...) ② shows that $R(A)$ is spanned by $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\1\end{bmatrix}$, which are linearly independent and hence form a basis. So, a basis of $R(A)$ is $\left\{ \begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix} \right\}$

6.2 :



$x_1 = \begin{bmatrix}1\\2\end{bmatrix}$ $\quad x_2 = \begin{bmatrix}2\\1\end{bmatrix}$

$$x_1 = x_1^0 + \tilde{x}_1$$

$\tilde{x}_1$ = projection of x_1 onto x_2

$$= \frac{(x_1^T x_2)}{\|x_2\|} \hat{x}_2 \quad \text{← unit vector along } x_2$$

$$= (x_1^T x_2) \frac{x_2}{\|x_2\|^2}$$

$$\left(\begin{aligned} &= \|x_1\| \cdot \|x_2\| |\cos\theta| = |x_1^T x_2| \\ &\text{proj} = \|x_1\| \cos\theta \end{aligned} \right)$$

$$\Rightarrow x_1^0 = x_1 - \tilde{x}_1 = x_1 - (x_1^T x_2) \frac{x_2}{\|x_2\|^2}$$

$$= \begin{bmatrix}1\\2\end{bmatrix} - \frac{4}{(\sqrt{5})^2}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}1 - \frac{8}{5}\\ 2 - \frac{4}{5}\end{bmatrix} = \begin{bmatrix}-\frac{3}{5}\\ +\frac{6}{5}\end{bmatrix}$$

normalize: $\quad y_1 = \dfrac{x_1^0}{\|x_1^0\|} = \dfrac{1}{\sqrt{5}}\begin{bmatrix}-1\\2\end{bmatrix}$

$$y_2 = \frac{x_2}{\|x_2\|} = \frac{1}{\sqrt{5}}\begin{bmatrix}2\\1\end{bmatrix}$$

6.3 Let $x_1, x_2, \cdots, x_n$ be a given basis of $\mathbb{R}^n$.

We "project x_1 onto span$\{x_2, \cdots x_n\}$ orthogonally,"

so $x_1 = x_1^o + \tilde{x}_1$ where $\tilde{x}_1 \in \text{span}\{x_2, \cdots x_n\}$

and $x_1^o \perp \text{span}(x_2, \cdots x_n)$

Assign $y_1 \triangleq \dfrac{x_1^o}{\|x_1\|}$

Now pick x_2 and project it orthogonally onto span$\{x_3, \cdots x_n\}$ so that $x_2 = x_2^o + \tilde{x}_2$, where $\tilde{x}_2 \in \text{span}\{x_3, \cdots x_n\}$ and $x_2^o \perp \text{span}\{x_3, \cdots x_n\}$

Note that $x_2^o \perp y_1$ by construction.

Now assign $y_2 \triangleq \dfrac{x_2^o}{\|x_2^o\|}$.

repeat this procedure till you are done with all the n - vectors. By construction, $y_i \perp y_j$ $i \neq j$ and $\|y_i\| = 1$ $\forall i$, and $\{y_1, \cdots y_n\}$ span $\mathbb{R}^n$.

This is called the Gram - Schmidt procedure.

To do it numerically, you'll have to construct matrices called orthogonal projectors. You can read about them in Carl Meyer's book.

<u>Problem 7:</u> statement $p \equiv$ " H is a full column rank matrix and P is a symmetric positive definite matrix."

Statement $q \equiv$ " $H^T P H$ is symmetric positive definite "

We have to show $p \Rightarrow q$

ie, starting with p, we have to show $x^T(H^T P H)x > 0 \quad \forall x \neq 0$

pick an arbitrary x, Now $x^T(H^T P H)x = (Hx)^T P (Hx) \triangleq J$

Since P is sym pos. def., $J \geq 0$, and $J = 0$ if and only if $Hx = 0$, but that occurs if and only if $x = 0$, since H is full column rank.

$\Rightarrow x^T(H^T P H)x > 0 \quad \forall x \neq 0 \Rightarrow H^T P H$ is positive definite

$(H^T P H)^T = H^T P^T H = H^T P H \Longrightarrow$ it is symmetric. So we have $p \Rightarrow q$.

<u>reverse:</u> To show it is an if and only if statement, we have to show $q \Rightarrow p$ as well.

That is, if " $H^T P H$ is symmetric +ve definite", then " H is a full column rank matrix and $P = P^T > 0$ ". As you can see, it is asking quite a lot. In fact, this is ~~true~~ not true.

<u>Counter example:</u> $H = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$. $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ H is full column rank, but P is not +ve definite, so ~~P~~ this example violates "p".

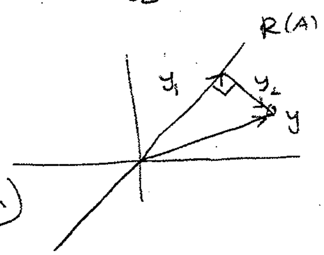
evaluate $H^T P H = \begin{bmatrix} 5 & 14 \\ 14 & 41 \end{bmatrix}$,

which is a symmetric +ve definite matrix (how do you verify ?)

so, we have an example of "q, but not p", which shows that '$q \Rightarrow p$' is false. Hence the answer is <u>no</u>

Problem 8: decompose y as:

$y = y_1 + y_2$ where $y_1 \in R(A)$, $y_2 \in R(A)^{\perp}$



Since $y_2 \in R(A)^{\perp}$ and $R(A)^{\perp} = N(A^T)$,

(by the orthogonal decomposition theorem)

$\therefore y_2 \in N(A^T)$

$\Rightarrow A^T y_2 = 0$

Now multiply both sides of the equation $y = y_1 + y_2$ by A^T:

$$A^T y = A^T(y_1 + y_2) = A^T y_1 + A^T y_2 = A^T y_1$$

Since $y_1 \in R(A)$, $\exists$ a vector, call it x_0 such that

$Ax_0 = y_1$

$$\Rightarrow A^T y = A^T(y_1) = A^T(Ax_0) = A^T A x_0$$

(problem 8 contd...)

Therefore, ~~we have found~~ there exists an x_0 which satisfies ~~the eq~~

$A^TA\,x = A^Ty$, (given y and A)

$\Rightarrow$ Normal equations are always consistent.

Note: When people say "$Ax=y$ does not have a solution" or even "consider $Ax=y$", ~~they do not mean~~ the equality between Ax and y is merely a convinient notation. The data given is A and y, and the problem is to find x such that Ax becomes equal to y. So "$Ax=y$" does not in general mean $\exists x$ and y such that Ax is equal to y. !

Problem 9: $\|x\|_2 \triangleq \sqrt{x_1^2+\cdots+x_n^2} = \sqrt{x^Tx}$ (typo in the H.W.)

start with $\|\alpha x - y\|_2^2 \geq 0$, $\alpha \in \mathbb{R}$, which is true since the 2-norm of a vector is always non-negative..

$$\|\alpha x - y\|_2^2 \geq 0 \Leftrightarrow (\alpha x - y)^T(\alpha x - y) \geq 0 \Leftrightarrow \alpha^2 x^Tx + y^Ty \geq 2\alpha\, x^Ty \quad \text{—(1)}$$

choose $\alpha = \dfrac{x^Ty}{x^Tx}$ $(\in \mathbb{R})$

$$\therefore\ \alpha^2 x^Tx + y^Ty \geq 2\alpha\, x^Ty \Rightarrow \frac{(x^Ty)^2}{x^Tx} + y^Ty > 2\frac{(x^Ty)^2}{x^Tx} \Rightarrow (x^Ty)^2 \leq (y^Ty)(x^Tx)$$

$$\Rightarrow |x^Ty| \leq \sqrt{y^Ty}\cdot\sqrt{x^Tx} = \|x\|\,\|y\|$$

Problem 10 1 and 2 are trivial.

to prove 3,

$$\|x+y\|^2 = (x+y)^T(x+y) = x^Tx + y^Ty + 2x^Ty \quad (\because x^Ty = y^Tx)$$

by the CBS inequality,

$$x^Ty \le |x^Ty| \le \sqrt{(x^Tx)(y^Ty)}$$

$$\Rightarrow \|x+y\|^2 \le \|x\|_2^2 + \|y\|_2^2 + 2\sqrt{\|x\|_2^2\,\|y\|_2^2}$$

$$= \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2\|y\|_2$$

$$= (\|x\|_2 + \|y\|_2)^2$$

$$\Rightarrow \|x+y\| \le \big|\|x\| + \|y\|\big| = \|x\| + \|y\|$$

Alt solution to 9

$$(\alpha x+y)^T(\alpha x+y) \ge 0 \quad \forall \alpha \in \mathbb{R}$$

$$\Rightarrow \alpha^2 x^Tx + 2\alpha x^Ty + y^Ty \ge 0$$

this is a quadratic (in α) inequality. For this to hold for all α, no real roots of $\alpha^2(x^Tx) + \alpha(2x^Ty) + y^Ty = 0$ are possible, $\Leftrightarrow$ $b^2 - 4ac \le 0$

$$\Leftrightarrow (2x^Ty)^2 - 4(x^Tx)(y^Ty) \le 0$$

$$\Leftrightarrow (x^Ty)^2 \le (x^Tx)(y^Ty) = \|x\|^2\|y\|^2$$

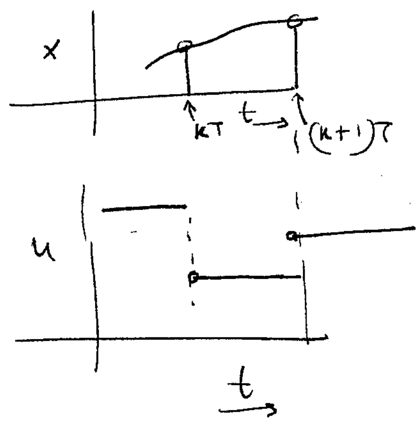
$$\Rightarrow |x^Ty| \le \|x\|\,\|y\|.$$

Problem 11

$$x(t) = e^{A(t-t_0)} x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)} B u(\tau) d\tau$$

$$x_k := x(kT)$$

$$x_{k+1} := x((k+1)T), \qquad u(t) = u_k \quad \text{for } kT \le t < (k+1)T$$



so integrate between

$t_0 = kT$ and $t = (k+1)T$

$$x((k+1)T) = e^{A((k+1)T - kT)} x(kT) + \int_{kT}^{(k+1)T} e^{A((k+1)T - \tau)} B u_k \, d\tau$$

$$\Rightarrow \quad x_{k+1} = e^{AT} x_k + \int_{kT}^{(k+1)T} e^{A(kT+T-\tau)} B u_k \, d\tau$$

define $\quad \eta = (k+1)T - \tau$

$$\Rightarrow d\eta = -d\tau, \quad \eta : T \longrightarrow 0$$

as $\tau : kT \longrightarrow (k+1)T$

$$= e^{AT} x_k + \int_{T}^{0} e^{A\eta} (-d\eta) . B u_k$$

$$= e^{AT} x_k + \left(\int_0^T e^{A\eta} d\eta \; B \right) u_k$$

$$\Rightarrow \underline{A_d = e^{AT}}, \quad \underline{B_d = \left(\int_0^T e^{A\eta} d\eta \right) B}.$$

[note that e^A, B, u. etc are matrices and vectors, their order of multiplication cannot be changed]

$$y(kT) = C \, x(kT)$$

$$\Rightarrow y_k = C \, x_k \quad \Rightarrow \quad \underline{C_d = C}$$