HW # 06 Solutions

Problem 1:

$$v^1 = y^1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \|v^1\|^2 = 6.$$

$$v^2 = y^2 - a_{21} v^1.$$

$$a_{21} = \frac{< y^2, v^1 >}{\|v^1\|^2} = \frac{\begin{bmatrix} 4 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}{6},$$

$$= \frac{3}{6} = \frac{1}{2}.$$

$$\therefore v^2 = \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3\frac{1}{2} \\ 1 \\ -1\frac{1}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix}.$$

$$\|v^2\|^2 = \frac{1}{4}(49 + 4 + 9) = \frac{62}{4} = \frac{31}{2}.$$

$$v^3 = y^3 - a_{31} v^1 - a_{32} v^2.$$

$$a_{31} = \frac{< y^3, v^1 >}{\|v^1\|^2} = \frac{\begin{bmatrix} -2 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}{6},$$

$$= -\frac{3}{6} = -\frac{1}{2}.$$

$$a_{32} = \frac{< y^3, v^2 >}{\|v^2\|^2} = \frac{\begin{bmatrix} -2 & 2 & 3 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix} \left(\frac{1}{2}\right)}{3\frac{1}{2}},$$

$$= -\frac{19}{31}.$$

$$v^3 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{19}{31} \begin{bmatrix} 3\frac{1}{2} \\ 1 \\ -1\frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 40 \\ 100 \\ 160 \end{bmatrix} (\frac{1}{62}) \approx \begin{bmatrix} 0.65 \\ 1.61 \\ 2.58 \end{bmatrix}.$$

Problem 2:

(a) The naive estimate is plotted in Fig. 1

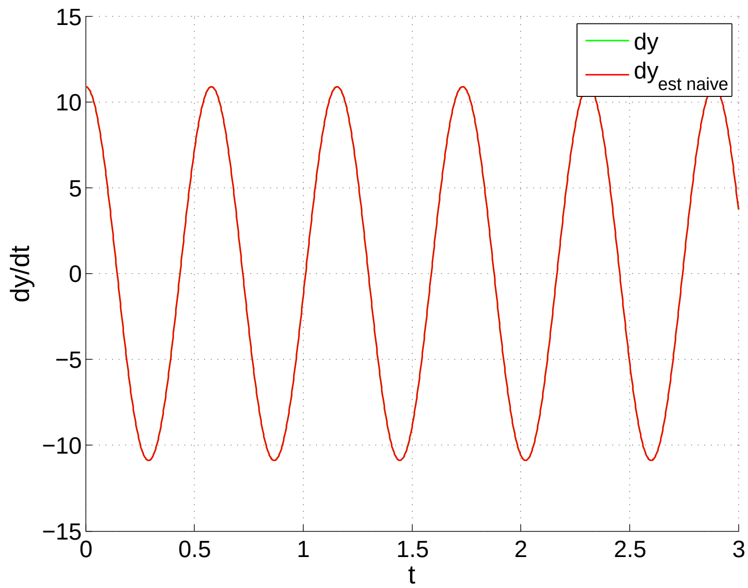

Figure 1: Naive Estimate

(b) We define Y_k so that it contains $M \geq 2$ of the "most recent" data points

$$Y_k = \begin{bmatrix} y[k-M+1] \\ \vdots \\ y[k] \end{bmatrix},$$

where $y[k] = y(k\Delta T)$. For basis functions, we take the monomials, but you can use any set of independent functions for which you can compute the derivative. We let $\varphi_i(t) = t^i$, where $\varphi_0(t) = 1$.

Suppose that at time $t_k = k\Delta T$, we regress the data against $\{\varphi_0(t), \cdots, \varphi_N(t)\}$, in other words,

$$y(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_N t^N.$$

We then have

$$Y_k = A_k \alpha$$

where

$$A_k = \begin{bmatrix} 1 & (k-M+1)\Delta T & \cdots & ((k-M+1)\Delta T)^N \\ 1 & (k-M+2)\Delta T & \cdots & ((k-M+2)\Delta T)^N \\ \vdots & \vdots & \cdots & \vdots \\ 1 & (k-2)\Delta T & \cdots & ((k-2)\Delta T)^N \\ 1 & (k-1)\Delta T & \cdots & ((k-1)\Delta T)^N \\ 1 & k\Delta T & \cdots & (k\Delta T)^N \end{bmatrix}$$

which depends on k, and thus changes step-to-step. We need $M \geq N+1$ for the columns of the matrix to be linearly independent. At the k-th step we have

$$\alpha = (A_k^\top A_k)^{-1} A_k^\top Y_k$$

We plug these coefficients back into

$$y(t) = \alpha_0 + \alpha_1(t) + \cdots + \alpha_N(t)^N,$$

we differentiate it, evaluate it at whatever time we desire, and use that for our estimate of $\dot{y}(t)$. This is an acceptable solution, but a much more practical solution is available to us.

Suppose instead that at time t_k, we regress the data against $\{\varphi_0(t-t_k), \cdots, \varphi_N(t-t_k)\}$, in other words,

$$y(t) = \alpha_0 + \alpha_1(t-t_k) + \cdots + \alpha_N(t-t_k)^N.$$

All we are doing is shifting the time origin to t_k. By doing this, we end up with

$$Y_k = A\alpha$$

where

$$A = \begin{bmatrix} 1 & (-M+1)\Delta T & \cdots & ((-M+1)\Delta T)^N \\ 1 & (-M+2)\Delta T & \cdots & ((-M+2)\Delta T)^N \\ \vdots & \vdots & \cdots & \vdots \\ 1 & -2\Delta T & \cdots & (-2\Delta T)^N \\ 1 & -\Delta T & \cdots & (-\Delta T)^N \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

which does not change from one time step to the next. We still need $M \geq N+1$ for the columns of the matrix to be linearly independent. At the k-th step we have

$$\alpha = (A^\top A)^{-1} A^\top Y_k$$

and we only need to compute $(A^\top A)^{-1} A^\top$ once. This is what I do on my robots. The calculation of the inverse is done off-line and stored.

We now compute

$$\dot{y}(t) = \alpha_1 + 2\alpha_2(t-t_k) + \cdots + N\alpha_N(t-t_k)^{(N-1)},$$

and thus

$$\dot{y}(t) = \left[0, 1, 2(t-t_k), \cdots, N(t-t_k)^{(N-1)}\right]\alpha$$

Setting $t = t_k$, we obtain

$$\widehat{\dot{y}}_k = [0, 1, 0, \cdots, 0]\,\alpha,$$

in other words,

$$\widehat{\dot{y}}_k = RY_k$$

where

$$R = [0, 1, 0, \cdots, 0]\,(A^\top A)^{-1} A^\top.$$

Choosing $M = 4$ and $N = 2$, we obtain the plot of the derivative given in Fig. 2 It looks exactly that same as the naive derivative, so we are disappointed that we worked so hard! **Remark: If you take $M = 2$ and $N = 1$ you get exactly the naive derivative.**

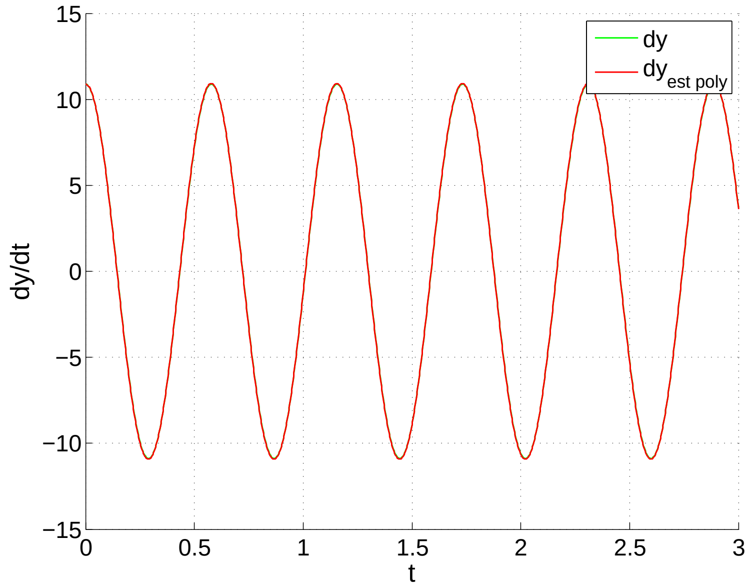

Figure 2: Regression

Problem 3:

(a) The naive estimate is plotted in Fig. 3. Computing the error gives

$$\frac{||\dot{y}_k - \widehat{\dot{y}}_k||}{\text{Length of the data vector}} = 0.093,$$

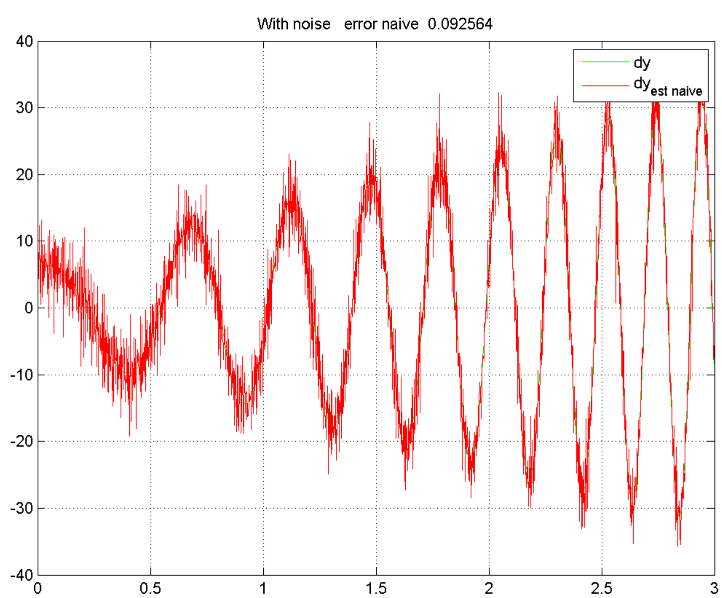

Figure 3: Naive Estimate

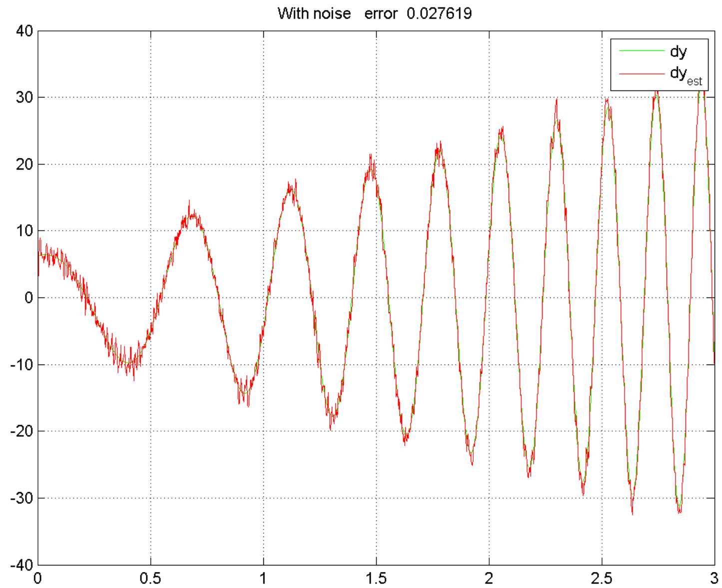

Figure 4: Regression

(b) We do the same regression analysis as in Problem 2. We play around with the parameters a bit and settle on $M = 10$ and $N = 2$. We obtain

$$\frac{||\dot{y}_k - \widehat{\dot{y}}_k||}{\text{Length of the data vector}} = 0.027,$$

and a plot of the derivative is given in Fig. 4

Problem 4: We apply the normal equations:

$$\hat{x} = \alpha_1 y^1 + \alpha_2 y^2,$$

where $G^\top \alpha = b$ and

$$G = \begin{bmatrix} < y^1, y^1 > & < y^2, y^1 > \\ < y^1, y^2 > & < y^2, y^2 > \end{bmatrix},$$
$$b = \begin{bmatrix} < x, y^1 > \\ < x, y^2 > \end{bmatrix}.$$

Doing the calculations, we have

$$< y^1, y^1 >= \text{tr}\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \right) = \text{tr}\left(\begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix} \right) = 5.$$

$$< y^1, y^2 >=< y^2, y^1 >= \text{tr}\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) = \text{tr}\left(\begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix} \right) = 3.$$

$$< y^2, y^2 >= \text{tr}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) = \text{tr}\left(\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \right) = 4.$$

$$< x, y^1 >= \text{tr}\left(\begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \right) = \text{tr}\left(\begin{bmatrix} 4 & 0 \\ -1 & 0 \end{bmatrix} \right) = 4.$$

$$< x, y^2 >= \operatorname{tr}\left(\begin{bmatrix} 0 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) = \operatorname{tr}\left(\begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix} \right) = 1.$$

$$\therefore \begin{bmatrix} 5 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

$$\therefore \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 4 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 13 \\ -7 \end{bmatrix} = \begin{bmatrix} 1.18 \\ -0.64 \end{bmatrix}.$$

$$\hat{x} = \frac{13}{11} \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} - \frac{7}{11} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{6}{11} & -\frac{7}{11} \\ \frac{19}{11} & -\frac{7}{11} \end{bmatrix}.$$

Problem 5: Let $\gamma := d(x, M)$, and suppose that $m_1, m_2 \in M$ satisfy $\|x - m_i\| = \gamma$.
<u>To Show</u> $m_1 = m_2$ when the norm is strict.
Because M is a subspace, $\frac{m_1+m_2}{2} \in M$.
Hence,

$$\begin{aligned} \gamma = \inf_{y \in M} ||x - y|| &\leq \left\| x - \frac{m_1 + m_2}{2} \right\| \\ &= \left\| \frac{x - m_1}{2} + \frac{x - m_2}{2} \right\| \\ &\leq \frac{1}{2} ||x - m_1|| + \frac{1}{2} ||x - m_2|| \\ &= \frac{\gamma}{2} + \frac{\gamma}{2} \\ &= \gamma. \end{aligned}$$

Hence, $||(x - m_1) + (x - m_2)|| = ||x - m_1|| + ||x - m_2||$.
Because the norm is strict, $\exists \alpha \geq 0$ such that either

(i) $(x - m_1) = \alpha(x - m_2)$ or

(ii) $(x - m_2) = \alpha(x - m_1)$.

In either case, we deduce from $\gamma = ||x - m_1|| = ||x - m_2||$, that $\gamma = \alpha\gamma$, and, because $\gamma \neq 0$, we have $\alpha = 1$.
With $\alpha = 1$, we have $x - m_1 = x - m_2$, and thus $m_1 = m_2$. $\square$

Problem 6:

(a) Not strictly normed. Let

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then $||x+y||_1 = 2 = ||x||_1 + ||y||_1$, but there does not exist any $\alpha \geq 0$ such that either $x = \alpha y$ or $y = \alpha x$.

(c) Not strictly normed. Let

$$x = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{and} \quad y = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

Then $6 = ||x+y||_\infty = ||x||_\infty + ||y||_\infty$, but there does not exist any $\alpha \geq 0$ such that either $x = \alpha y$ or $y = \alpha x$.

(b) Strictly normed. The result is true for any norm induced by an inner product. Hence we give the proof for $||x|| =< x, x >^{1/2}$.

Let $x, y \in X$

Case 1 Either x or y is zero. Then $||x+y|| = ||x|| + ||y||$ is always true and either $x = 0 \cdot y$ or $y = 0 \cdot x$ holds. □

Case 2 Both $x \neq 0$ and $y \neq 0$, but $\{x, y\}$ is linearly dependent. Then $x = \alpha y$ for some $\alpha \in \mathbb{R}$. It follows that $||x+y|| = ||1+\alpha|| \cdot ||y||$ and $||x|| + ||y|| = (1+|\alpha|)||y||$. Because $y \neq 0$, we have

$$||x+y|| = ||x|| + ||y|| \iff |1+\alpha| = 1 + |\alpha| \iff \alpha \geq 0.$$

Hence, $||x+y|| = ||x|| + ||y|| \iff x = \alpha y, \alpha \geq 0.$ □

Case 3 $\{x, y\}$ is linearly independent. By the Gram-Schmidt procedure, there exists $v \in X$ such that $x \perp v$ and $\text{span}\{x, y\} = \text{span}\{x, v\}$. Write $y = \alpha_1 x + \alpha_2 v$, so that

$$\begin{aligned} &x + y = (1+\alpha_1)x + \alpha_2 v \ , \quad ||x+y|| = ||x|| + ||y|| \\ &\text{if, and only if,} ||x+y||^2 = ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| \end{aligned}$$

By the Pythagorean Theorem,

$$\begin{aligned} ||x+y||^2 &= ||(1+\alpha_1)x + \alpha_2 v||^2 \\ &= (1+\alpha_1)^2 ||x||^2 + (\alpha_2)^2 ||v||^2 \\ &= [1 + 2\alpha_1 + (\alpha_1)^2]||x||^2 + (\alpha_2)^2 ||v||^2. \end{aligned}$$

Furthermore

$$||y||^2 = (\alpha_1)^2 ||x||^2 + (\alpha_2)^2 ||v||^2.$$

Hence,

$$||x+y||^2 = ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| \ \text{ if, and only if, } \ 2\alpha_1 ||x||^2 = 2||x|| \cdot ||y||.$$

Because $||x|| \neq 0$ from $\{x, y\}$ linear independent, we have

$$\alpha_1 ||x|| = ||y||.$$

$\therefore \alpha_1 \geq 0$. Moreover,

$$(\alpha_1)^2 ||x||^2 = ||y||^2 = (\alpha_1)^2 ||x||^2 + (\alpha_2)^2 ||v||^2 \ \text{ if, and only if, } \ \alpha_2 = 0.$$

Hence,

$$||x+y|| = ||x|| + ||y|| \ \text{ if, and only if, } \ y = \alpha_1 x, \alpha_1 \geq 0$$

□

Problem 7: Given $a \in \mathbb{R}^{m\times n}$, the rank-nullity theorem states that:

$$\text{rank}(A) + \text{nullity}(A) = n$$

Proof: Let $\{x_1, \ldots, x_p\}$ be a set of p linearly independent vectors which form a basis for $\mathcal{N}(A)$. In other words, $\dim \mathcal{N}(A) = p$. Since $\mathcal{N}(A)$ is a subspace of $\mathbb{R}^n$, we can complete this basis to a basis for $\mathbb{R}^n$. Let $\{x_1, \ldots, x_p, x_{p+1}, \ldots, x_n\}$ be our completed basis for $\mathbb{R}^n$.

To show: $\text{rank}(A) = n - p$.

$x \in \mathbb{R}^n$ can be written as:

$$\begin{aligned} x &= \alpha_1 x_1 + \cdots + \alpha_p x_p + \alpha_{p+1} x_{p+1} + \cdots + \alpha_n x_n \\ Ax &= \alpha_1 Ax_1 + \cdots + \alpha_p Ax_p + \alpha_{p+1} Ax_{p+1} + \cdots + \alpha_n Ax_n \end{aligned}$$

Since $x_i \in \mathcal{N}(A)$, $1 \leq i \leq p$, then $\alpha_1 Ax_1 = \cdots = \alpha_p Ax_p = 0$, and

$$Ax = \alpha_{p+1} Ax_{p+1} + \cdots + \alpha_n Ax_n$$

Since $Ax \in \mathcal{R}(A)$, this is sufficient to show that $\mathcal{R}(A) = \text{span}\{Ax_{p+1}, \ldots, Ax_n\}$. However, these vectors are only a basis for $\mathcal{R}(A)$ if they are linearly independent! In other words, if we only have the trivial solution, $\beta_i = 0$, $p+1 \leq i \leq n$ to:

$$\begin{aligned} \beta_{p+1} Ax_{p+1} + \cdots + \beta_n Ax_n &= 0 \\ A(\beta_{p+1} x_{p+1} + \cdots + \beta_n x_n) &= 0 \end{aligned}$$

$\beta_{p+1} x_{p+1} + \cdots + \beta_n x_n \in \mathcal{N}(A)$, so we can write this as:

$$\begin{aligned} \beta_{p+1} x_{p+1} + \cdots + \beta_n x_n &= \gamma_1 x_1 + \cdots + \gamma_p x_p \\ -\gamma_1 x_1 - \cdots - \gamma_p x_p + \beta_{p+1} x_{p+1} + \cdots + \beta_n x_n &= 0 \end{aligned}$$

Since $\{x_1, \ldots, x_p, x_{p+1}, \ldots, x_n\}$ is our basis for $\mathbb{R}_n$, the only solution is $\gamma_i = 0$, $1 \leq i \leq p$ and $\beta_j = 0$, $p+1 \leq j \leq n$. Therefore, $\{Ax_{p+1}, \ldots, Ax_n\}$ form a basis for $\mathcal{R}(A) \implies \dim \mathcal{R}(A) = n - p \implies \text{rank}(A) + \text{nullity}(A) = n$.

Alternate proof using other theorems we know: For some subspace $S \subset \mathcal{X}$, $S \oplus S^\perp = \mathcal{X}$, then $\dim S + \dim S^\perp = \dim \mathcal{X}$. Since $\mathcal{R}(A^\top) \oplus \mathcal{N}(A) = \mathbb{R}^n$ (proven in lecture) and $\dim \mathcal{R}(A) = \dim \mathcal{R}(A^\top)$ (also proven in lecture), then $\dim \mathcal{R}(A) + \dim \mathcal{N}(A) = \dim \mathbb{R}^n = n$.

Problem 8:

```
A=diag([1 0.5 0.5 1 0.5]);
B=[1 0 2 0 3]';
C=0.2; D=B';
AplusBCD=A+B*C*D;
invA=inv(A);
[InvAplusBCD] = MatInvLemma(invA,B,C,D);
test=AplusBCD*InvAplusBCD

test =

 1.0000         0         0         0    0.0000
      0    1.0000         0         0         0
-0.0000         0    1.0000         0    0.0000
      0         0         0    1.0000         0
      0         0    0.0000         0    1.0000
```

```
function [InvAplusBCD] = MatInvLemma(invA,B,C,D)
%
% A, B, C, D are matrices of compatible dimensions so that
% A+BCD makes sense.
%
% inv(A) = invA is assumed to be provided to us because we
% typically use this function in a recursuive process
%
% We also assume (and do check) that C is invertible.
%
if cond(C) < 1e-8
    InvAplusBCD=[];
else
    InvAplusBCD=invA-invA*B*inv(inv(C)+D*invA*B)*D*invA;
end
end
```

Problem 9:
Suppose $\hat{x}$ takes the form $\hat{x} = \alpha_1 y^1 + \alpha_2 y^2 + +\alpha_3 y^3$, we need to solve α_1, α_2, α_3 using normal equation such that $\hat{x} = \arg\min_{y \in M} \|x - y\|$.
We first define the inner products for G:

$$< y^1, y^1 > = \int_{-1}^{1} 1 \cdot 1 \mathrm{d}t = 2, \qquad < y^1, y^2 >= \int_{-1}^{1} 1 \cdot t \mathrm{d}t = 0, \quad < y^1, y^3 >= \int_{-1}^{1} 1 \cdot (3t^2/2 - 1/2) \mathrm{d}t = 0$$
$$< y^2, y^2 > = \int_{-1}^{1} t \cdot t \mathrm{d}t = 2/3, \quad < y^2, y^3 >= \int_{-1}^{1} t \cdot (3t^2/2 - 1/2) \mathrm{d}t = 0, \quad < y^3, y^3 >= \int_{-1}^{1} (3t^2/2 - 1/2)^2 \mathrm{d}t = 2/5$$

Now we only need to calculate:

$$\begin{cases} < x,\, y^1 >= \int_{-1}^{1} e^t \mathrm{d}t = e^t \Big|_{-1}^{1} = e - e^{-1} \\ < x,\, y^2 >= \int_{-1}^{1} t \cdot e^t \mathrm{d}t = e^t(t-1) \Big|_{-1}^{1} = 2e^{-1} \\ < x,\, y^3 >= \int_{-1}^{1} t^2 \cdot e^t \mathrm{d}t = 3/2e^t(t^2 - 2t + 2) - 1/2e^t \Big|_{-1}^{1} = e - 7e^{-1} \end{cases} \Rightarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 2/5 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} e - e^{-1} \\ 2e^{-1} \\ e - 7e^{-1} \end{bmatrix}$$

Solve for α_1, α_2, α_3,we get:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 5/2 \end{bmatrix} \begin{bmatrix} e - e^{-1} \\ 2e^{-1} \\ e - 7e^{-1} \end{bmatrix} = \begin{bmatrix} 1.175 \\ 1.104 \\ 0.358 \end{bmatrix}$$

Then you get $\hat{x} = 1.175 + 1.104t + 0.358(3t^2/2 - 1/2)$.