ECE 532 - FALL 2017 - EXTRA CREDIT

3. Projection matrices

P is square
$P^2 = P$

(a) Assume $I-P$ is a projection matrix. Then $(I-P)^2 = I-P$.

$$(I-P)^2 = (I-P)\cdot(I-P)$$
$$= I^2 - 2IP + P^2 = I - 2P + P^2$$
$$= I - 2P + P = I - P$$

Therefore $I-P$ is a projection matrix.

(b) $$(A(A^TA)^{-1}A^T)^2 = (A(A^TA)^{-1}A^T)\cdot(A(A^TA)^{-1}A^T)$$
$$= A\cdot(A^TA)^{-1}\cancel{(A^TA)}\cancel{(A^TA)^{-1}}A^T$$
$$= A(A^TA)^{-1}A^T$$

A has linearly independent columns, so it is full rank, so it is a square matrix.

Therefore, $(A(A^TA)^{-1}A^T)^2$ is a projection matrix.

(c) Let $U = [\underline{u}_1 \quad \underline{u}_2 \quad \cdots \quad \underline{u}_K]$. Since $\underline{u}_i \; \forall i$ are orthonormal vectors, U is an orthogonal matrix.
Then:

$$U\cdot U^T = [\underline{u}_1 \quad \underline{u}_2 \quad \cdots \quad \underline{u}_K]\cdot\begin{bmatrix}\underline{u}_1\\ \underline{u}_2\\ \vdots\\ \underline{u}_K\end{bmatrix} = \underline{u}_1\underline{u}_1^T + \underline{u}_2\cdot\underline{u}_2^T + \cdots + \underline{u}_K\underline{u}_K^T$$

and $$(UU^T)^2 = (U\underbrace{\cancel{U^T})(\cancel{U}}_{=I}U^T)$$
$$= UU^T \quad \therefore \text{ projection matrix}$$

4. (a) Using the Vandermonde matrix:

$$\underbrace{\begin{bmatrix} x_1^2 & x_1 & 1\\ x_2^2 & x_2 & 1\\ \vdots & \vdots & \vdots\\ x_n^2 & x_n & 1\end{bmatrix}}_{X}\cdot\underbrace{\begin{bmatrix}p\\ q\\ r\end{bmatrix}}_{\underline{w}} \approx \underbrace{\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}}_{\underline{y}} \Rightarrow X\underline{w} \approx \underline{y}$$

So we want to minimize the difference between $X\underline{w}$ and $\underline{y}$: $\min_{\underline{w}} \|\underline{y} - X\underline{w}\|_2^2$.
Since the Vandermonde matrix is always invertible: $\hat{\underline{w}} = (X^TX)^{-1}X^T\cdot\underline{y}$.

(b) If our vertical offset is 1, that specifies our choice of r : $r = 1$.

$$\begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \cdot \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \approx \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow \underbrace{\begin{bmatrix} x_1^2 & x_1 \\ x_2^2 & x_2 \\ \vdots & \vdots \\ x_n^2 & x_n \end{bmatrix}}_{X} \cdot \underbrace{\begin{bmatrix} p \\ q \end{bmatrix}}_{\underline{w}} \approx \underbrace{\begin{bmatrix} y_1 - 1 \\ y_2 - 1 \\ \vdots \\ y_n - 1 \end{bmatrix}}_{\underline{y}} \Rightarrow \hat{\underline{w}} = \min_{w} \| X\underline{w} - \underline{y} \|_2^2$$

$$= (X^T X)^{-1} X^T \cdot \underline{y}$$

(c)
$y = px^2 + qx + 1$
$y' = 2px + q$

$y'(0) = 2p(0) + q = 1$ (since it's an angle of $45°$ at launch, $t=0$)
$\Rightarrow q = 1$

$$\begin{bmatrix} x_1^2 & x_1 \\ x_2^2 & x_2 \\ \vdots & \vdots \\ x_n^2 & x_n \end{bmatrix} \cdot \begin{bmatrix} p \\ 1 \end{bmatrix} \approx \begin{bmatrix} y_1 - 1 \\ y_2 - 1 \\ \vdots \\ y_n - 1 \end{bmatrix} \Rightarrow \underbrace{\begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{bmatrix}}_{X} \cdot \underbrace{[p]}_{\underline{w}} \approx \underbrace{\begin{bmatrix} y_1 - x_1 - 1 \\ y_2 - x_2 - 1 \\ \vdots \\ y_n - x_n - 1 \end{bmatrix}}_{\underline{y}} \Rightarrow \hat{\underline{w}} = \min_{w} \| X\underline{w} - \underline{y} \|_2^2$$

$$= (X^T X)^{-1} X^T \underline{y}$$

5. Visualizing least-squares

(a) $\hat{\underline{x}} = (A^T A)^{-1} A^T \cdot \underline{b}$

$= \left(\begin{bmatrix} 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

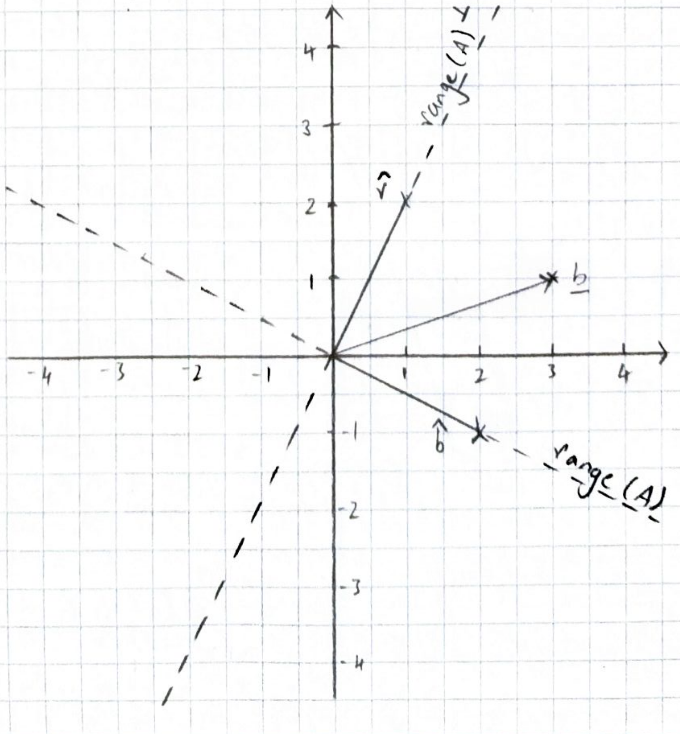
$= (5)^{-1} \cdot (5)$

$= 1$

$\therefore \hat{\underline{b}} = A \cdot \hat{\underline{x}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

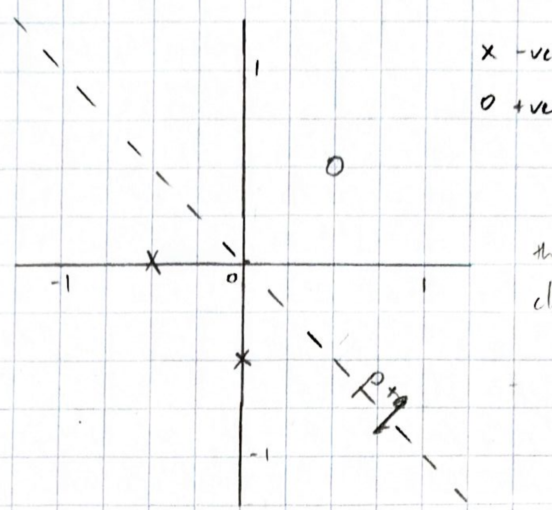
$\hat{\underline{r}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

(b)



6. A Microbiology Experiment

(a)



if we let $w_1 = 1$ and $w_2 = 1$, the boundary that forms successfully classifies our data

(b) Our weight vector $\underline{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ has no offset, so all decision boundaries $w_1 p_1 + w_2 q_2 = 0$ pass through the origin. There's no way to arrange the weights such that (p_4, q_4) is classified correctly while leaving the other classifications correct. However, any offset $-0.5 < r < -0.1$ would work:

$$\hat{y} = \text{sign}(1 \cdot p + 1 \cdot q - 0.3).$$

7. Properties of the SVD

(a)
- Σ is a diagonal matrix
- Σ is sorted in descending order
- $V^T V = I$
- $U^T U = I$

(b) Since the last row of Σ is zeros, one possible change could be

$$\tilde{u} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ 2 & -2 & -1 \end{bmatrix}$$

and we still get that $A = \tilde{u} \Sigma V^T$

(c) Since the smallest element of Σ is 15 (i.e. the smallest amplification factor), the matrix x should correspond to the last column of V: $x = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$

(d) Since the largest element of Σ is 30, the best rank-1 approximation of A is

$$X = \left(\frac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}\right) \cdot \left(\overset{2}{\cancel{30}}\right) \cdot \left(\frac{1}{5}\begin{bmatrix} 3 & 4 \end{bmatrix}\right) = \begin{bmatrix} 12 & 16 \\ 6 & 8 \\ 12 & 16 \end{bmatrix}$$

(e) For the least squares solution $\hat{x}$ to result in a zero residual, b must lie in the range of A, which is spanned by the first 2 columns of U.

(f) b must then lie in the nullspace of A, so it must ~~be~~ lie in the span of the last column of U:

$$b \in \text{span}\left(\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}\right)$$

8. $f(x) = (x-1)^2 + \lambda x^2$

(a) $f'(x) = 2(x-1)^1 \cdot 1 + 2\lambda x$

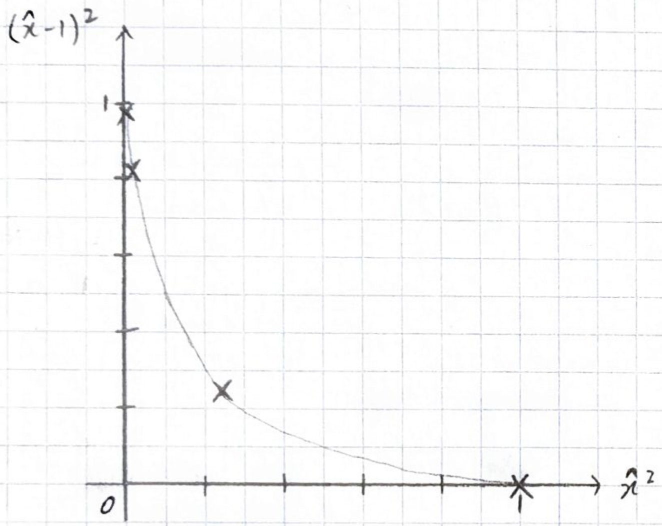
at min point, $f'(x) = 0$

$0 = 2x - 2 + 2\lambda x$

$\cancel{2} = \cancel{2}x(1+\lambda)$

$\Rightarrow \hat{x} = \dfrac{1}{1+\lambda}$

(b)

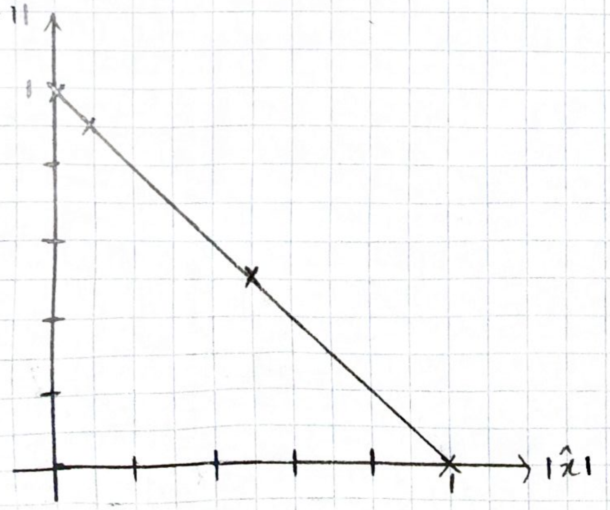


$$\hat{x}^2 = \frac{1}{(\hat{x}+1)^2}$$

$$(\hat{x}-1)^2 = \hat{x}^2 - 2\hat{x} + 1$$
$$= \frac{1}{(1+\lambda)^2} - \frac{2}{1+\lambda} + 1$$
$$= \frac{1 - 2(1+\lambda) + (1+\lambda)^2}{(1+\lambda)^2}$$

λ	$\hat{x}^2$	$(\hat{x}-1)^2$	$\lvert\hat{x}\rvert$	$\lvert\hat{x}-1\rvert$
0	1	0	1	0
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$
10	0.0083	0.8264	$0.\overline{09}$	$0.\overline{90}$
100	0.0001	0.9803	$0.\overline{0099}$	$0.\overline{9900}$

(c)



We are looking for a value of $\hat{x}$ that is close to 1, and in $0 \leq \hat{x} \leq 1$, $x^2 \leq x$. So while this is a straight line, the above curve is quadratic.