(4) A is a $n \times m$ matrix with linearly independent columns.

In the question no. (3) we have shown that, if $A_{n \times m}$ has ~~with~~ linearly independent columns, $A^T A$ is an invertible matrix.

So there is a least square solution $\bar{x}$ and its associated normal system,

$$A^T A \bar{x} = A^T b$$

As $A^T A$ is invertible,

$$(A^T A)^{-1} A^T A \bar{x} = (A^T A)^{-1} A^T b$$

$$I \bar{x} = (A^T A)^{-1} A^T b$$

$$\bar{x} = (A^T A)^{-1} A^T b$$

As inverse of a matrix is unique,

$\bar{x}$ is the unique solution to the system of equation $Ax = b$. [Showed]

(2) $p(w|t) = \mathcal{N}(w \mid m_N, S_N)$ .... (3.49)

$m_N = S_N (S_0^{-1} m_0 + \beta \Phi t)$ ... (3.50)

$S_N^{-1} = S_0^{-1} + \beta \Phi^T \Phi$ ... (3.51)

From Bayes's theorem, $p(w|t) \propto p(t|w)\, p(w)$

$\Rightarrow \cancel{p(w|t)} \propto \left[\prod_{n=1}^{N} \mathcal{N}(t_n \mid w^T \phi(x_n), \beta^{-1})\right] \mathcal{N}(w \mid m_0, S_0)$

$\propto \exp\left(-\frac{\beta}{2}(t - \Phi w)^T (t - \phi w)\right) \exp\left(-\frac{1}{2}(w - m_0)^T S_0^{-1} (w - m_0)\right)$

$\propto \exp\left(+\frac{\beta}{\cancel{2}}\left(t^T t - 2 w^T \Phi^T\right)(t - \phi w)\right\} + \cancel{1}\,(w^T - m_0^T) S_0^{-1} (w - M_0)\Big)$

[omitting $-\frac{1}{2}$ & expanding transpose]

$\propto \exp\Big(\beta t^T t - 2\beta w^T \Phi^T t + \beta w^T \Phi^T \Phi w + w^T S_0^{-1} w$

$\cancel{-2m_0} - 2 w^T S_0^{-1} m_0 + m_0^T S_0^{-1} m_0\Big)$

$\propto \exp\Big(\beta w^T \Phi^T \Phi w + w^T S_0^{-1} w - \cancel{2}\big(2 w^T (\beta \Phi^T t + S_0^{-1} m_0\big) +$

$\underbrace{\beta t^T t + m_0^T S_0^{-1} m_0}_{\text{constant}}\Big)$

$\propto \exp\Big(w^T (\beta \Phi^T \Phi + S_0^{-1}) w - 2 w^T . S_N^{-1} S_N (S_0^{-1} m_0 + \beta \Phi^T t)$

$+ \text{constant}\Big)$

$\propto \exp\left(w^T S_N^{-1} w - 2 w^T S_N^{-1} m_N + \text{constant}\right)$

$\propto \exp\left((w - m_N)^T S_N^{-1} (w - m_N)\right)$ [$m_N^T S_N^{-1} m_N$ = constant]

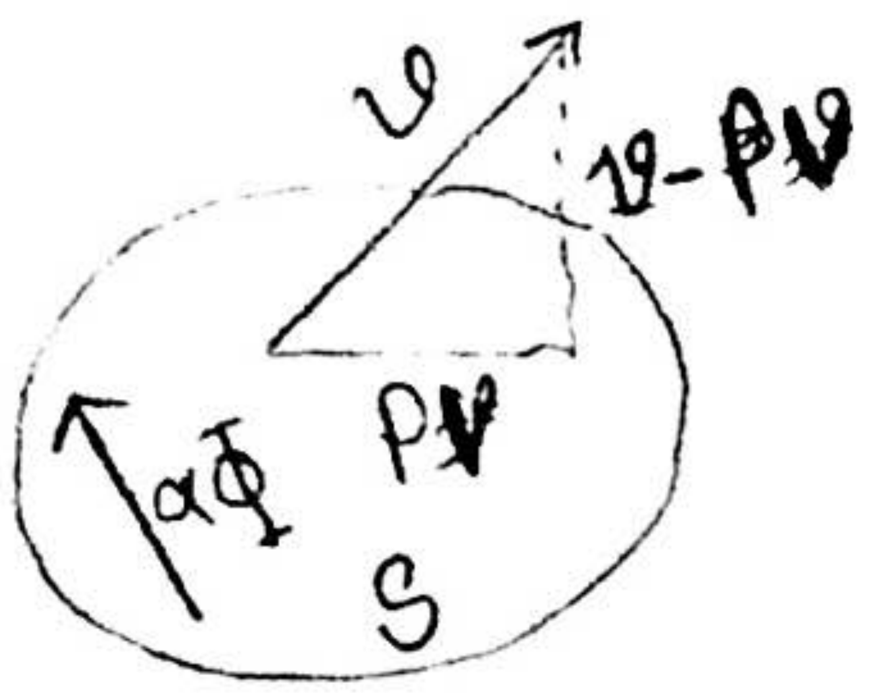
(8)

$$\propto \mathcal{N}(\omega \mid m_N, S_N) = P(\omega \mid \mathcal{D})$$

[Showed]

Linear Regression

(1)



Here, $P = \Phi(\Phi^T\Phi)^{-1}\Phi^T$

$$Pv = \Phi(\Phi^T\Phi)^{-1}\Phi^T v$$

$$v - Pv = v(I - P) = v(I - \Phi(\Phi^T\Phi)^{-1}\Phi^T)$$

Let an arbitrary vector $\alpha\Phi$ lies in the subspace S.

Now,

$$(v - Pv) . \alpha\Phi$$

$$= \alpha . v(I - \Phi(\Phi^T\Phi)^{-1}\Phi^T) . \Phi$$

$$= \alpha . v(\cancel{\Phi}\, I . \Phi - \Phi \underbrace{(\Phi^T\Phi)^{-1} . \Phi^T . \Phi}_{I})$$

$$= \alpha . v(\Phi - \Phi . I)$$

$$= \alpha . v(\Phi - \Phi)$$

$$= 0$$

As their dot product is 0, so matrix P projects v onto the space S. [Showed]

P.T.O.

Now, $w_{ML} = (\Phi^T \Phi)^{-1} \Phi^T t$

w_{ML} is a vector. Because,

let Φ is $n \times m$ matrix & t is a $n \times 1$ vector.

$$\text{then, } w_{ML} = \underbrace{\underbrace{\underbrace{(\underbrace{\Phi^T \; \Phi}_{m\times n \;\; n\times m})^{-1}}_{m\times m} \; \underbrace{\Phi^T}_{m\times n}}_{m\times n} \; \underbrace{t}_{\substack{n\times 1 \\ n\times 1}}}_{m\times 1}$$

From

$y = \Phi w^{*}_{ML} = \Phi (\Phi^T \Phi)^{-1} \Phi^T t$

So This is an orthogonal projection of t

onto the manifold S, as,

Φw_{ML} evaluates to a vector~~ta~~ & we can

(Matrix. Vector $\rightarrow$ vector)

think of it as a linear combination of

the columns of Φ. So, according to the

first part, w_{ML} corresponds to an

orthogonal projection of t on the S.

[Showed

(3) Using (3·3), (3·8) & (3·49), we can rewrite (3·5) as

$$P(t|x,\mathbf{t},\alpha,\beta) = \int P(t|x,w,\beta)\, P(w|\mathbf{t},\alpha,\beta)\, dw$$

$$= \int \underbrace{\mathcal{N}(t|\phi(x)^T w, \beta^{-1})}_{\text{matching with (2·114)}} \underbrace{\mathcal{N}(w|m_N, S_N)}_{\text{matching with (2·113)}} dw$$

matching with (2·114)

$\Downarrow$

$t = y$

$\phi(x)^T = A$

$w = x$

$\beta^{-1} = L^{-1}$

matching with (2·113)

$\Downarrow$

$w = x$

$m_N = \mu$

$S_N = \Lambda^{-1}$

Using (2·115), $[P(y) = \mathcal{N}(y|A\mu + b, L^{-1} + A\Lambda^{-1}A^T)$

$$P(t|x,\mathbf{t},\alpha,\beta) = \mathcal{N}\left(t \,|\, \phi(x)^T m_N,\ \beta^{-1} + \phi(x)^T S_N \phi(x)\right)$$

$$= \mathcal{N}\left(t \,|\, m_N^T \phi(x)^T,\ \frac{1}{\beta} + \phi(x)^T S_N \phi(x)\right) \quad \text{--- (3·58)}$$

and, $$\sigma_N^2(x) = \frac{1}{\beta} + \phi(x)^T S_N \phi(x)$$

4 $$\left(M + vv^T\right)^{-1} = M^{-1} - \frac{(M^{-1}v)(v^TM^{-1})}{1+v^TM^{-1}v} \quad \dots (3 \cdot 110)$$

from (3·59)

$$\sigma^2_{N+1} = \frac{1}{\beta} + \phi(x)^T S_{N+1} \phi(x) \dots (3 \cdot 59)$$

Using P (r.f. Question 3·8):

$$S^{-1}_{N+1} = S_N^{-1} + \beta \phi_{N+1} \phi^T_{N+1} \quad \dots (P)$$

Using (3·110) & ~~(3·59)~~ (P)

$$S_{N+1} = \left[S_N^{-1} + \cancel{\beta\phi(x_{N+1}} \beta\phi_{N+1}\phi^T_{N+1} \right]^{-1}$$

$$= S_N - \frac{(\sqrt{\beta}\, S_N \phi_{N+1}) \cdot (\sqrt{\beta}\, \phi^T_{N+1} S_N)}{1 + \beta \phi^T_{N+1} S_N \phi_{N+1}}$$

$$= S_N - \frac{\beta S_N \phi_{N+1} \phi^T_{N+1} S_N}{1 + \beta \phi^T_{N+1} S_N \phi_{N+1}}$$

Incorporating ~~(P)~~ it into (3·59),

$$\sigma^2_{N+1}(x) = \frac{1}{\beta} + \phi(x)^T S_{N+1} \phi(x)$$

$$= \cancel{\sigma^2_N(x)} + \frac{1}{\beta} + \phi(x)^T \left(S_N - \frac{\beta S_N \phi_{N+1} \phi^T_{N+1} S_N}{1+\beta\phi^T_{N+1} S_N \phi_{N+1}} \right) \phi(x)$$

$$\cancel{= \sigma^2_N(x) - \beta\phi(x)^T S_N \phi(x_{N+1}) \phi(x_{N+1})^T}$$

$$= \sigma^2_N(x) - \frac{\beta\, \phi(x)^T S_N \phi(x_{N+1}) \phi(x_{N+1})^T S_N \phi(x)}{1 + \beta \phi(x_{N+1})^T S_N \phi(x_{N+1})}$$

$$\sigma^2_{N+1}(x) = \sigma^2_N(x) - \frac{\phi(x)^T S_N \phi(x_{N+1}) \phi(x_{N+1})^T S_N \phi(x)}{\frac{1}{\beta} + \phi(x_{N+1})^T S_N \phi(x_{N+1})}$$

$$= \sigma^2_N(x) - \underbrace{\frac{\cancel{\phi(x)^T \phi S_N \phi(x_{N+1})}\left(\phi(x)^T S_N \phi(x_{N+1})\right)^2}{\frac{1}{\beta} + \phi(x_{N+1})^T S_N \phi(x_{N+1})}}$$

→ this part is larger or equal to (≥ 0) ~~than~~ 0, because ~~&~~ the numerator ≥ 0 [square] & denominator is greater than 0. and S_N is positive definite matrix.

So, $\sigma^2_{N+1}(x) \leq \sigma^2_N(x)$