Orthogonalization & Steepest Descent Algorithm

Sunday, November 17, 2024 8:10 AM

Orthogonal Matrix.

$$A A^H = I$$

$$A = \begin{bmatrix} | & | & | & & | \\ | & | & | & \cdots & | \\ \underline{e}_1 & \underline{e}_2 & \underline{e}_3 & & \underline{e}_N \end{bmatrix}$$

$\underline{e}_1 \underline{e}_2^H = \underline{0}$

$\underline{e}_1 \underline{e}_1^H = \underline{I}$

e_1, e_3

Inner product / dot

$\langle \underline{e}_1, \underline{e}_1 \rangle = \| e_1^2 \|$

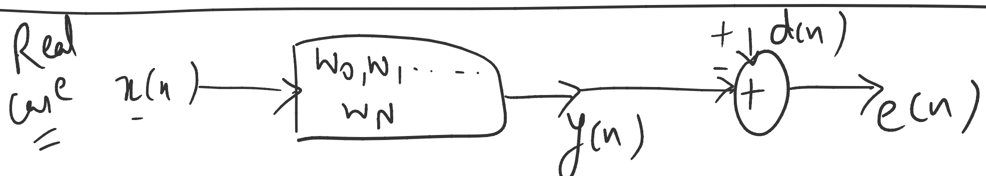
$\langle \underline{e}_1, \underline{e}_2 \rangle = 0$

orthogonal

SF

Basis function $\rightarrow$ inner product

ML



H: Conjugate + Transpose

$$y(n) = \underline{w}^H \underline{x}(n)$$
$$= \underline{w}^t \underline{x}(n)$$

Least square solution (ML)

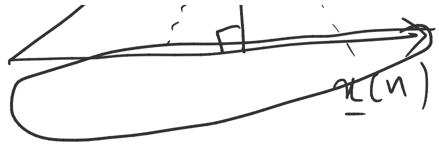
$$W_{opt} = \underline{R}^{-1} \underline{P}$$
$$\underline{R} = E[\underline{x}(n) \underline{x}^H(n)]$$
$$\underline{P} = E[\underline{x}(n) d(n)]$$

$\underline{\text{error}} = \boxed{d(n) - y(n)}$

$\rightarrow$ is orthogonal to the space span by $\underline{x}(n)$ when we use W_{opt} values in finding $y(n)$.

$\rightarrow$ (Linear Algebra)

$\underline{x}(n) = \{ x(n), x(n-1), \ldots, x(n-N) \}$



$\underline{x}(n) = \{ x(n),\ x(n-1) \ldots x(n-N) \}$

error $= \| d(n) - w_0 x(n) + w_1 x(n-1) + w_2 x(n-2) + \ldots + w_N x(n-N) \|$

$\perp \underline{x}(n)$

are optimal values

Weights: $N \times 1$ $\quad \{ w_0^{N+1} + w_1 x(n) + w_2 (x(n-1)) + \ldots + w_{N+1} x(n-N) + \text{bias} \}$

After bias weights: $N+1 \times 1$ $\quad w_0 \times x(n+1) + w_1 x(n) + w_2 x(n-1) + \ldots$

bias

bias & variance trade-off (ML)

$w_{opt} = \underline{R}^{-1} \underline{p}$

$y(n) = \underline{w}^t \underline{x}(n)$

$= (\underline{R}^{-1} \underline{p})^t \underline{x}(n)$

$(\underline{A}\underline{B})^t = \underline{B}^t \underline{A}^t$

For example

$w = \underline{R}\underline{p}$

$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$

$= \underline{p}^t (\underline{R}^{-1})^t \underline{x}(n)$

$= \underline{p}^t (\underline{R}^t)^{-1} \underline{x}(n)$

$y(n) = p^t R^{-1} x(n)$

$(R^{-1})^t = (R^t)^{-1}$

Proof

R: Hermition

$R = R^t$

$$y(n) = \underline{p}^t \, \underline{R}^{-1} \, \underline{x}(n)$$

$$\text{error} = e(n) = d(n) - y(n)$$

$$e(n) = d(n) - \underline{p}^t \, \underline{R}^{-1} \, \underline{x}(n)$$

$$E\left[e(n) \, \underbrace{x(n-k)}\right] = 0, \quad k = 0, 1, 2 \dots N$$

$$E\left[e(n) \, \underline{x}^t(n)\right] = \underline{0}$$

$$= E\left[\left(d(n) - \underline{p}^t \, \underline{R}^{-1} \, \underline{x}(n)\right)\left(\underline{x}^t(n)\right)\right]$$

$$= E\left[d(n) \, \underline{x}^t(n) - \underline{p}^t \, \underline{R}^{-1} \, \underline{x}(n) \, \underline{x}^t(n)\right]$$

$$= \underbrace{E\left[d(n) \, \underline{x}^t(n)\right]} - E\left[\underline{p}^t \, \underline{R}^{-1} \, \underline{x}(n) \, \underline{x}^t(n)\right]$$

$$=$$

$$= \underline{p}^t - \underline{p}^t \, \underline{R}^{-1} \, \underline{R}$$

$$= \underline{0}$$

$$R = \underline{R}^t$$

$$R R^{-1} = I$$

$$(R R^{-1})^t = (\underline{I})^t$$

$$(\underline{R}^{-1})^t \, \underline{R}^t = \underline{I}$$

$$(R^{-1})^t = (R^t)^{-1}$$

$$\underline{R} = E\left[\underline{x}(n) \, \underline{x}^t(n)\right]$$

$$\underline{p} = E\left[\underline{x}(n) \, d^t(n)\right]$$

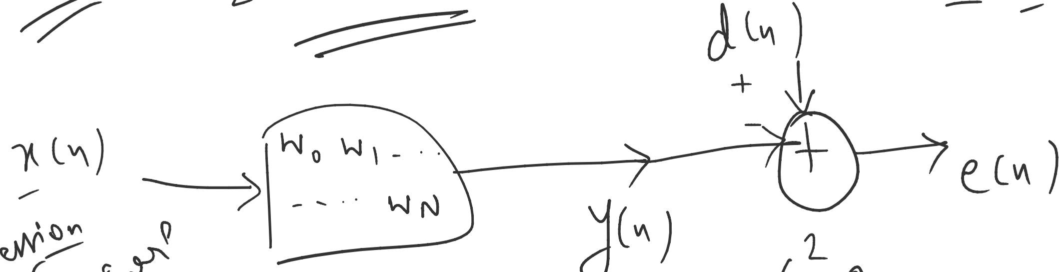
$$\underline{p}^t = E\left[d(n) \, \underline{x}^t(n)\right]$$

Steepest Descent Procedure

Steepest Descent Procedure

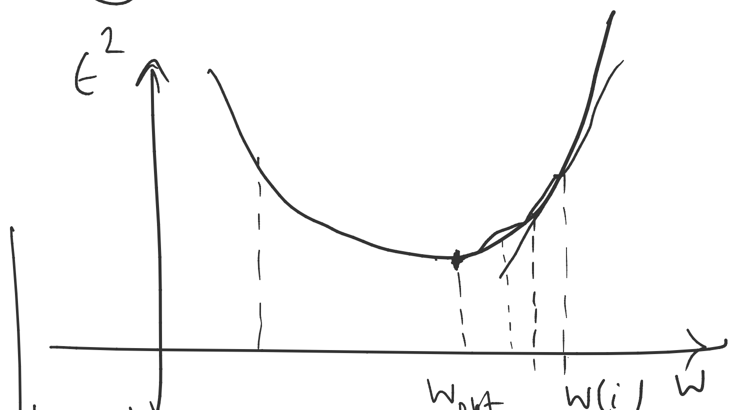
$10^5 \times 10^5$ $\quad$ 2048×2048

$w_{opt} = \underline{R}^{-1}\underline{p}$ closed form



Regression
Error ↳ Mean Squared error
↳ cosine sim.
↳ KL divergence

$$\epsilon^2 = E\left[e^2(n)\right] = E\left[(d(n) - y(n))^2\right]$$

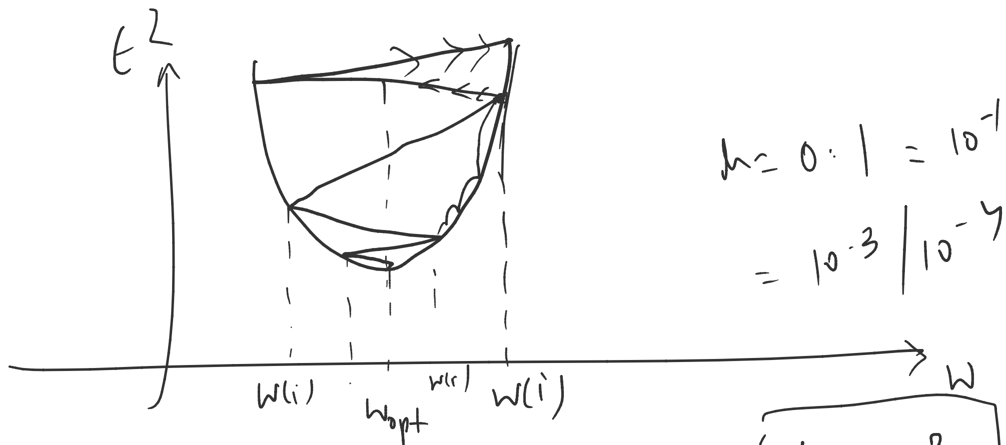


$$w(i+1) = w(i) - \frac{\mu}{2} \nabla_w \epsilon^2 \Big|_{w=w(i)}$$

$$= w(i) - \frac{\mu}{2}\left[2\underline{R}w - 2\underline{p}\right]_{w=w(i)}$$

$$= w(i) - \mu\left[Rw(i) - p\right]$$

$$= w(i) + \mu p - \mu \underline{R} w(i)$$



$\mu = 0.1 = 10^{-1}$

$= 10^{-3} / 10^{-4}$

$\|w_{pr} - w_{opt}\| < \eta$

$\boxed{\mu = 10^3 \text{ +ve value}}$