Kalman Filter

process noise

consider an LTI system: $\begin{cases} \dot{x} = Ax + Bu + Gw \\ y = Cx + D(t) \end{cases}$ → measorment noise

- In this system, first of all we do not know $x(0)$
- secondly, the output $y(t)$ is available but is affected by process noise and measorment noise.

⇒ simple observer doesnot work ⇒ Kalman filter

since noise is a random signal ⇒ we need to review random processes

State estimation

Problem formulation:

Plant: $\begin{cases} \dot{x} = Ax + Bu + Gw(t) \\ y = Cx + v(t) \end{cases}$

where:
- $v(t)$, $w(t)$, $x(0)$ are mutually orthognal. and are unrelated.
- $x(0) \approx N(\bar{x}_0, P_0)$
- $w(t)$, $v(t)$ are white noises with
  - Zero mean: $E\{w(t)\} = 0$, $E\{v(t)\} = 0$
  - $R_{ww}(\tau) = E\{w(t)\, w^T(t+\tau)\} = Q\,\delta(\tau)$
  - $R_{vv}(\tau) = E\{v(t)\, v^T(t+\tau)\} = R\,\delta(\tau)$ $\hookrightarrow$ R non singular

    $E\{w(t)\, w^T(\tau)\} = Q\,\delta(t-\tau)$

Objective: is to design a state estimator to estimate the state $x(t)$ by $\hat{x}(t)$ s.t. the estimation error covariance is minimized

Analysis of this estimator:

plant: $\begin{cases} \dot{x} = Ax + Bu + Gw \\ y = Cx + v \end{cases}$

observer: $\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) \\ \hat{y} = C\hat{x} \end{cases}$

$\Rightarrow$ define $\tilde{x} = x - \hat{x} \Rightarrow \dot{\tilde{x}} = \dot{x} - \dot{\hat{x}} = (Ax + \cancel{Bu} + Gw) - [A\hat{x} + \cancel{Bu} + L(\cancel{y} - \hat{y})]$

$$= (A - LC)x - (A - LC)\hat{x} + Gw - Lv$$

$$= (A - \underline{\underline{L}}C)\tilde{x} + Gw + \underline{\underline{L}}v$$

Remark The dynamic of the estimation error explicitly shows the conflict between estimation speed and the attenuation of the meaurment noise's effect.

- Fast state reconstruction requires rapid decay rate
  estimation decay rate: $Re\{eig(A - LC)\}$
  which is typically acheived at Large L. but Large L magnify the effect of v on the estimation process.

- To mitigate the effect of v on $\hat{x}(t)$, we should choose smal L but this reduce the estimation speed.

$\Rightarrow$ we should find an optimal balance between these two conflicting problems $\Rightarrow$ kalman filter

Optimization

$$\dot{\tilde{X}} = \underbrace{(A - LC)}_{\bar{A}} \tilde{x} + \underbrace{Gw + LV}_{d(t)}$$

Define $P(t) = E\{\tilde{x}(t)\, \tilde{x}^T(t)\}$

$$\dot{P}(t) = \frac{d}{dt} E\{\tilde{x}(t)\, \tilde{x}^T(t)\} = E\{\frac{d}{dt} \tilde{x}(t)\tilde{x}^T(t)\} =$$

$$= E\{\dot{\tilde{x}}(t)\tilde{x}^T(t)\} + E\{\tilde{x}(t)\dot{\tilde{x}}^T(t)\} =$$

$$= E\{(\bar{A}\tilde{x}(t) + d(t))\tilde{x}^T(t)\} + E\{\tilde{x}(t)[\bar{A}\tilde{x}(t) + d(t)]^T\} =$$

$$= \bar{A}\, \underbrace{E\{\tilde{x}(t)\tilde{x}^T(t)\}}_{P(t)} + \underbrace{E\{\tilde{x}(t)\tilde{x}^T(t)\}}_{P(t)}\bar{A}^T + E\{d(t)\tilde{x}^T(t)\} + E\{\tilde{x}(t)\, d(t)\}^T$$

Appendix 1 $\Rightarrow \dot{P}(t) = \bar{A}P(t) + P(t)\bar{A}^T + \frac{1}{2}\nabla + \frac{1}{2}\nabla^T = \bar{A}P(t) + P(t)\bar{A}^T + \nabla$

where $E\{d(t)d^T(\tau)\} = \nabla\delta(t-\tau)$, $\nabla = LRL^T + GQG^T$

objective: our purpose is to minimize the cost function

How $E\{\tilde{x}(t)\, W\, \tilde{x}(t)\}$

$$\min_{L} J = E\{\tilde{x}^T(t)\,\tilde{x}(t)\} = \text{trace}(\underbrace{E\{\tilde{x}(t)\,\tilde{x}^T(t)\}}_{P})$$

estimator gain $\leftarrow L$

$$\begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix}[\cdot\ \cdot\ \cdot] = \begin{bmatrix} * & & \\ & * & \\ & & * \end{bmatrix}$$

subject to $\dot{P}(t) = \bar{A}P(t) + P(t)\bar{A}^T + \nabla$

$\Rightarrow$ introduce a Lagrange multiplier Λ

$$\bar{J} = \text{trace}(P + \Lambda(-\dot{P} + \bar{A}P + P\bar{A}^T + \nabla))$$

$$= \text{trace}(P - \Lambda\dot{P} + \Lambda(A - LC)P + \Lambda P(A^T - C^TL^T) + \Lambda GQG^T + \Lambda LRL^T)$$

$$= \text{trace}(P - \Lambda\dot{P} + \Lambda AP + \Lambda PA^T - \Lambda LCP - \Lambda PC^TL^T + \Lambda GQG^T + \Lambda LRL^T)$$

The external happens at $\frac{\partial \bar{J}}{\partial L} = 0$

$$\frac{\partial}{\partial X}\text{trace}(AXB) = \frac{\partial}{\partial X}\text{trace}(B^TX^TA^T) = A^TB^T$$

$$\frac{\partial}{\partial x}\text{trace}(AXBX^T) = A^TXB + AXB$$

$$\frac{\partial}{\partial x}\text{trace}(AX^T) = A$$

$$\Rightarrow \frac{\partial \bar{J}}{\partial L} = -\Lambda^TP^TC^T - \Lambda PC^T + \Lambda^TLR + \Lambda LR$$

$=$

See Appendix 5

Since Λ could be arbitrary, choose $\Lambda = \Lambda^T \Rightarrow \frac{\partial \bar{J}}{\partial L} = -2\Lambda PC^T + 2\Lambda LR = 0$

$(P = E\{x(t)\,\tilde{x}(t)^T\} \Rightarrow P^T(t) = P(t))$

$$\Rightarrow \Lambda(-PC^T + LR) = 0$$

$$\Rightarrow \boxed{L = +PC^TR^{-1}}$$

$\Rightarrow$ substituting $L = PC^TR^{-1}$ in

$$\dot{P} = \bar{A}P + P\bar{A}^T + \bar{V} = (A-LC)P + P(A-LC)^T + GQG^T + LRL^T$$

$$= (A - PC^TR^{-1}C)P + P(A - PC^TR^{-1}C)^T + GQG^T + PC^TR^{-1}CP$$

$$= AP + PA^T - 2PC^TR^{-1}CP + GQG^T + PC^TR^{-1}CP$$

$$= AP + PA^T + GQG^T - PC^TR^{-1}CP$$

$$\boxed{\dot{P}(t) = AP + PA^T + GQG^T - PC^TR^{-1}CP}$$

$\hookrightarrow$ which is a matrix DRE.

Optimal Kalman Filter

Kalman-Bucy Filter – linear quadratic estimator (LQE)

plant : $\begin{cases} \dot{x} = Ax + Bu + G\,w(t) \\ y = Cx + v(t) \end{cases}$

where $v(t), w(t), x(0)$ are mutually orthogonal and unrelated

$x(0) = N(\bar{x}_0, P_0)$

$E\{w(t)\} = 0 \quad , \quad E\{v(t)\} = 0$

$R_{ww}(\tau) = E\{w(t)\, w^T(t+\tau)\} = Q\,\delta(\tau)$

$R_{vv}(\tau) = E\{v(t)\, v^T(t+\tau)\} = R\,\delta(\tau)$

observe : $\begin{cases} \dot{\hat{x}} = A\hat{x} + L(y - \hat{y}) \\ \hat{y} = C\hat{x} \end{cases}$ with initial guess $\hat{x}(0)$

P.I. $\quad J = E\{[x(t) - \hat{x}(t)]^T [x(t) - \hat{x}(t)]\} = \text{trace}\Big(\underbrace{E\{(x(t)-\hat{x}(t))(x(t)-\hat{x}(t))^T}_{P}\Big)$

objective: $\min\limits_{L} J$

solution : $L(t) = P(t)\, C^T R^{-1}$ where $P(t) \geq 0$ is the solution of following matrix DRE:

$$\begin{cases} \dot{P}(t) = AP(t) + P(t)A^T + GQG^T - P(t)C^T R^{-1} C^T P(t) \\ P(0) = E\{(x(0) - \hat{x}(0))(x(0) - \hat{x}(0))^T\} \end{cases}$$

$$= E\{x(0)^2\} - 2\hat{x}(0)\, E\{x(0)\} - \hat{x}(0)^2$$

$$= P_0 - 2\hat{x}(0)\,\bar{x}_0 - \hat{x}(0)^2$$

This DRE should be solved forward in time.

Remarks

- $\uparrow P(t) \approx \uparrow$ uncertainty in state estimation
- increase in process noise $\equiv \uparrow Q \Rightarrow$ increases GQG^T in $\dot{P}$
- increase in measurment noise $\equiv \uparrow R \Rightarrow$ decreases $PC^T R^{-1} C P$

Theorem

If (A, C) is detectable, then for every $P(0)$ there exists a bounded limiting P_∞ to the DRE. Furthermore P_∞ is a solution to the ARE too!

(A, C) detectable $\Rightarrow$

in $\dot{\tilde{x}} = (A - LC)\tilde{x} + d(t)$

there exists L s.t. $\bar{A} = A - LC$ is stable and $\tilde{x}(t)$ becomes asymptotically stable. $\Rightarrow$ From Appendix 4 $P(t) = e^{\bar{A}t} P(0) e^{\bar{A}^T t} + \int_0^t e^{\bar{A}\eta} \nabla e^{\bar{A}^T \eta} d\eta \Rightarrow$

$$\lim_{t \to \infty} P(t) = \lim \underbrace{e^{\bar{A}t} P(0) e^{\bar{A}^T t}}_{\bar{A}\text{ stable}}{}^{\to 0} + \int_0^t e^{\bar{A}\eta} \nabla e^{\bar{A}^T \eta} d\eta = \int_0^\infty e^{\bar{A}\eta} \nabla e^{\bar{A}^T \eta} d\eta = P_\infty$$

Moreover, since $P(t) \xrightarrow[t \to \infty]{} P_\infty \Rightarrow \dot{P} = 0 \Rightarrow P_\infty$ is a solution of ARE:

$$AP + PA^T + GQG^T - PC^T R^{-1} C P = 0$$

Theorem:

If (A, C) is detectable and if $(A, G\sqrt{Q})$ is stabilizable

for $Q \geq 0$, $R \geq 0$, the observer gain $L = P_\infty C^T R^{-1}$ results

in stable ~~$\bar{A}$~~ estimation matrix $\bar{A} = A - LC$

$L = P_\infty C^T R^{-1}$ where P_∞ is the solution of ARE:

$$P_\infty A^T + A P_\infty - P_\infty C^T R^{-1} C P_\infty + GQG^T = 0$$

$$P_\infty A^T - P_\infty C^T R^{-1} C P_\infty + A P_\infty \underbrace{- P_\infty C^T R^{-1} C P_\infty + P_\infty C^T R^{-1} C P_\infty}_{0} + GQG^T = 0$$

$$P_\infty \underbrace{(A^T - C^T R^{-1} C P_\infty)}_{\bar{A}^T} + \underbrace{(A - P_\infty C^T R^{-1} C)}_{\bar{A}} P_\infty + P_\infty C^T R^{-1} C P_\infty + GQG^T = 0$$

$$P_\infty \bar{A}^T + \bar{A} P = \underbrace{- P_\infty C^T R^{-1} C P_\infty - GQG^T}_{\nabla = LRL^T + GQG^T} = -\nabla$$

$$\left.\begin{array}{l} R > 0, Q \geq 0 \Rightarrow P_\infty \bar{A}^T + \bar{A} P_\infty = -\nabla < 0 \\ (A, G\sqrt{Q}) \text{ stabilizable} \end{array}\right\} \Rightarrow \bar{A} = A - LC \text{ is asymptotically stable.}$$

Theorem

If (A, C) is detectable, then P_∞ is the unique solution.

of ARE: $PA^T + AP + GQG - PC^T R^{-1} C P = 0 \equiv P\bar{A}^T + \bar{A}P + \nabla = 0$

Proof: by contradiction assume that, the solution of ARE is not unique and $\hat{P}$ also satisfies the ARE. Since (A, C) is detectable, $\exists L$ s.t. $\bar{A} = A - LC$ is stable

$\Rightarrow$ Let $\dot{z} = \bar{A}^T z \Rightarrow z(t) = e^{\bar{A}^T t} z(0) \Rightarrow z(\infty) = 0$

$\hat{P}$ is the solution of ARE $\Rightarrow \hat{P}\bar{A}^T + \bar{A}\hat{P} = -\nabla \Rightarrow z^T[\hat{P}\bar{A}^T + \bar{A}\hat{P}]z = -z^T \nabla z$

$\Rightarrow z^T \hat{P} \dot{z} + \dot{z}^T \hat{P} z = -z^T \nabla z \Rightarrow \frac{d}{dt}(z^T \hat{P} z) = -z^T \nabla z$

$\Rightarrow \int_0^\infty z^T \nabla z \, dt = \int_0^\infty z^T(0) e^{\bar{A}t} \nabla e^{\bar{A}^T t} z(0) dt = z^T(0) \int_0^\infty e^{\bar{A}t} \nabla e^{\bar{A}^T t} dt \, z(0) = z^T(0) P_\infty z(0)$

$\int_0^\infty \frac{d}{dt}(z^T \hat{P} z) dt = z^T \hat{P} z \Big]_0^\infty = z^T(\infty) \hat{P} \cancel{z(\infty)}^{0} - z^T(0) \hat{P} z(0) = -\int_0^\infty z^T \nabla z \, dt$

$\Rightarrow z^T(0) \hat{P} z(0) = z^T(0) P_\infty z(0)$ for any $z(0) \Rightarrow \hat{P} = P_\infty$ %.

Steady state case

summary

plant: $$\begin{cases} \dot{x} = Ax + Bu + Gw(t) \\ y = Cx + v(t) \end{cases}$$

where $v(t), w(t), x(0)$ are mutually orthogonal and are unrelated

$x(0) = N(\bar{x}_0, P_0)$

$w(t)$ and $v(t)$ are white noise with

- zero mean: $E\{v(t)\} = 0$ $E\{w(t)\} = 0$

- $R_{ww}(\tau) = E\{w(t)\, w^T(t+\tau)\} = Q\,\delta(t-\tau)$ $Q = Q^T \geq 0$
- $R_{vv}(\tau) = E\{v(t)\, v^T(t+\tau)\} = R\,\delta(t-\tau)$ $R = R^T > 0$

the measurment noise corrupts all the measurment channels

- (A, C) detectable
- $(A, G\sqrt{Q})$ stabilizable

observer: $$\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) \\ \hat{y} = C\hat{x} \end{cases} \qquad \hat{x}(0) = \bar{x}_0$$

Estimation P.I: $J(t) = E\{\underbrace{(x(t) - \hat{x}(t))}_{\tilde{x}}^T (x(t) - \hat{x}(t))\} = \text{trace}\,(E\{\tilde{x}(t)\,\tilde{x}(t)^T\})$

solution $L = P_\infty C^T R^{-1}$ where P_∞ is the P.D. soultion of

$$P A^T + AP - PC^T R^{-1} C P + G Q G^T = 0$$

properties:

$$\lim_{t \to \infty} E\{e(t)\} = 0 \qquad e(t) = \tilde{x}(t)$$

$$\lim_{t \to \infty} J(t) = \lim_{t \to \infty} E\{\tilde{x}(t)^T \tilde{x}(t)\} = \text{trace}\, P_\infty$$

- Properties of steady state LQE:

- $\lim_{t \to \infty} J = \lim E\{ \tilde{x}(t)^T \tilde{x}(t) \} = \text{trace}\, P_\infty$

$$\begin{cases} P(t) = E\{ \tilde{x}(t) \tilde{x}(t)^T \} \Rightarrow E\{ \tilde{x}(t)^T \tilde{x}(t) \} = \text{trace}\, P(t) \\ \lim_{t \to \infty} P(t) = P_\infty \end{cases}$$

$$\Rightarrow \lim_{t \to \infty} J = \text{trace}(P_\infty)$$

- $\lim_{t \to \infty} E\{ \tilde{x}(t) \} = 0$

$$\lim_{t \to \infty} E\{ \tilde{x}(t) \} = \lim_{t \to \infty} E\left\{ e^{\bar{A}t} \tilde{x}(0) + \int_0^t e^{\bar{A}(t-\tau)} d(\tau)\, d\tau \right\}$$

$$= e^{\bar{A}t} \cancel{E\{\tilde{x}(0)\}}^{\,0} + \int_0^t e^{\bar{A}(t-\tau)} \cancel{E\{d(\tau)\}}^{\,0} d\tau = 0$$

see appendix 3

$$E\{\tilde{x}(0)\} = E\{ x(0) - \underbrace{\hat{x}(0)}_{\text{deterministic}} \} = E\{x(0)\} - \hat{x}(0) = \bar{x}_0 - \bar{x}_0 = 0$$

Example:

let
$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + w(t) \\ y = \begin{bmatrix} 1 & 0 \end{bmatrix} x + v(t) \end{cases}$$

$$E\{w(t)\, w^T(\tau)\} = Q\, \delta(t-\tau) = \begin{bmatrix} 0.1 & 0 \\ 0 & 0 \end{bmatrix} \delta(t-\tau)$$

$$E\{v(t)\, v^T(\tau)\} = R\, \delta(t-\tau) = 0.2\, \delta(t-\tau)$$

$$Q = \begin{bmatrix} 0.1 & 0 \\ 0 & 0 \end{bmatrix} \qquad R = 0.2 \quad , \quad G = I$$

$$P_e A^T + A P_e - P_e C^T R^{-1} C P_e + Q = 0$$

$$\Rightarrow P_e = \begin{bmatrix} 0.0792 & -0.0343 \\ -0.0343 & 0.0314 \end{bmatrix}$$

$$\Rightarrow L = P_e C^T R^{-1} = \begin{bmatrix} 0.3962 \\ -0.1715 \end{bmatrix}$$

$$\Rightarrow \begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) \\ \hat{y} = C\hat{x} \end{cases}$$

Appendix-1

$$E\{\tilde{x}(t)\, d^T(t)\} = E\{[e^{\bar{A}t}\tilde{x}(0) + \int_0^t e^{\bar{A}(t-\tau)} d(\tau)\, d\tau]\, d^T(t)\}$$

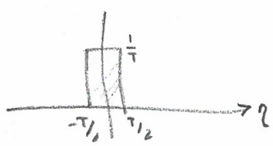
$$= e^{\bar{A}t} \underbrace{\cancel{E\{\tilde{x}(0)\, d^T(t)\}}}_{\text{From Appendix 3}}^{0} + E\{\int_0^t e^{\bar{A}(t-\tau)} d(\tau)\, d(t)\, d\tau\}$$

$$= \int_0^t e^{\bar{A}(t-\tau)} E\{d(\tau)\, d(t)\}\, d\tau = \int_0^t e^{\bar{A}(t-\tau)} \nabla\, \delta(\tau - t)\, d\tau$$

$$\eta = t - \tau \Rightarrow d\eta = -d\tau$$

$$\Rightarrow E\{\tilde{x}(t)\, d^T(t)\} = -\int_t^0 e^{\bar{A}\eta} \nabla\, \delta(\eta)\, d\eta = \int_0^t e^{\bar{A}\eta} \nabla\, \delta(\eta)\, d\eta$$

$$= \frac{1}{2}\nabla$$



$$\Rightarrow \int_0^t \delta(\eta)\, d\eta = \frac{1}{2}$$

$$\int_0^t f(\eta)\, \delta(\eta)\, d\eta = \frac{1}{2} f(0)$$

Appendix 2

$$E\{d(t)\,d^T(\tau)\} = E\left\{ [G \quad L] \begin{bmatrix} w(t) \\ v(t) \end{bmatrix} [w^T(\tau) \quad v^T(\tau)] \begin{bmatrix} G^T \\ L^T \end{bmatrix} \right\} =$$

$$= [G \quad L]\, E\left\{ \begin{bmatrix} w(t)\,w^T(\tau) & w(t)\,v^T(\tau) \\ v(t)\,w^T(\tau) & v(t)\,v^T(\tau) \end{bmatrix} \right\} \begin{bmatrix} G^T \\ L^T \end{bmatrix} =$$

$$= [G \quad L] \begin{bmatrix} E\{w(t)\,w^T(\tau)\} & \cancel{E\{w(t)\,v^T(\tau)\}}^{\,0} \\ \cancel{E\{v(t)\,w^T(\tau)\}}^{\,0} & E\{v(t)\,v^T(\tau)\} \end{bmatrix} \begin{bmatrix} G^T \\ L^T \end{bmatrix}$$

$$= [G \quad L] \begin{bmatrix} Q\,\delta(t-\tau) & 0 \\ 0 & R\,\delta(t-\tau) \end{bmatrix} \begin{bmatrix} G^T \\ L^T \end{bmatrix}$$

$$= \underbrace{(G Q G^T + L R L^T)}_{V}\, \delta(t-\tau)$$

Appendix 3

$$d(t) = Gw - Lv(t) = \begin{bmatrix} G & -L \end{bmatrix} \begin{bmatrix} w(t) \\ v(t) \end{bmatrix}$$

- $E\{d(t)\} = E\left\{ \begin{bmatrix} G & -L \end{bmatrix} \begin{bmatrix} w(t) \\ v(t) \end{bmatrix} \right\} = \begin{bmatrix} G & -L \end{bmatrix} \begin{bmatrix} E\{w(t)\} \\ E\{v(t)\} \end{bmatrix} = \begin{bmatrix} G & -L \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$

- $E\{\tilde{x}(0)\, d^T(t)\} = E\{\tilde{x}(0)\}\, E\{\cancelto{0}{d^T(t)}\} = 0$

Appendix 4

$$P(t) = E\{\tilde{x}(t)\tilde{x}^T(t)\} = e^{\bar{A}t} P(0) e^{\bar{A}^T t} + \int_0^t e^{\bar{A}\eta} \nabla e^{\bar{A}^T \eta} d\eta$$

proof:

$$\begin{cases} P(t) = E\{\tilde{x}(t)\tilde{x}^T(t)\} \\ \dot{\tilde{x}}(t) = \underbrace{(A - LC)}_{\bar{A}} \tilde{x}(t) + \underbrace{Gw(t) + Lv(t)}_{d(t)} = \bar{A}\tilde{x}(\tau) + d(\tau) \end{cases}$$

$$\Rightarrow \tilde{x}(t) = e^{\bar{A}t} x(0) + \int_0^t e^{A(t-\tau)} d(\tau) d\tau$$

$$\Rightarrow P(t) = E\{[e^{\bar{A}t}\tilde{x}(0) + \int_0^t e^{A(t-\tau)} d(\tau) d\tau][e^{\bar{A}t}\tilde{x}(0) + \int_0^t e^{\bar{A}(t-\tau)} d(\tau) d\tau]^T\}$$

1 2 3 4

$$= \underbrace{e^{\bar{A}t} E\{\tilde{x}(0)\tilde{x}^T(0)\} e^{\bar{A}^T t}}_{1} + \underbrace{\int_0^t e^{A(t-\tau)} \cancel{E\{d(\tau)\tilde{x}^T(0)\}}^{0} e^{\bar{A}^T t} d\tau}_{2} +$$

is deterministic

$$\underbrace{\int_0^t e^{\bar{A}t} \cancel{E\{\tilde{x}(0) d^T(\tau)\}}^{0} e^{\bar{A}^T t} d\tau}_{3} + \underbrace{\int_0^t \int_0^t e^{A(t-\tau)} \underbrace{E\{d(\tau) d^T(6)\}}_{\nabla \delta(\tau-6)} e^{\bar{A}^T(t-6)} d\tau d6}_{4}$$

see Appendix 2

See Appendix 3

$$= e^{\bar{A}t} P(0) e^{\bar{A}^T t} + \int_0^t e^{A(t-6)} \nabla e^{A(t-6)} d6 = e^{\bar{A}t} P(0) e^{\bar{A}^T t} + \int_0^t e^{\bar{A}\eta} \nabla e^{\bar{A}^T \eta} d\eta$$

$$\eta = t - 6 \Rightarrow d\eta = -d6$$

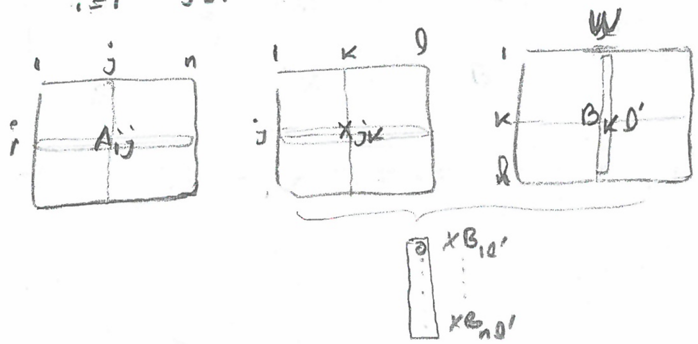
Appendix 5

- $\frac{\partial}{\partial x} \operatorname{trace}(\underset{n\times n}{A}\,\underset{n\times l}{X}\,\underset{l\times n}{B}) = A^T B^T$

the iw'th element

Let $M = AXB$ $\Rightarrow$ $\operatorname{trace}(AXB) = \operatorname{trace}[M_{iw}] = \operatorname{trace}\left[\sum_{j=1}^{n} A_{ij} \sum_{k=1}^{l} X_{jk} B_{k} \right] =$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} \sum_{k=1}^{l} x_{jk} B_{ki}$$



$$\Rightarrow \frac{\partial}{\partial x_{j_0 k_0}} \operatorname{trace}(AXB) = \sum_{i=1}^{n} A_{ij_0} B_{k_0 i} = (BA)_{k_0 j_0} = (BA)^T_{j_0 k_0}$$

$$\Rightarrow \frac{\partial}{\partial x} \operatorname{trace}(AXB) = (BA)^T = A^T B^T$$

- $\frac{\partial}{\partial x} \operatorname{trace}(Ax^T) = A$ $\qquad \operatorname{trace}(Ax^T) = \operatorname{trace}(xA^T)$

$$\frac{\partial}{\partial x} \operatorname{trace}(Ax^T) = \frac{\partial}{\partial x} \operatorname{trace}(xA^T) = \frac{\partial}{\partial x}(I\,x A^T) = A$$

- $\frac{\partial}{\partial x} \operatorname{trace}(AXBX^T) = A^T X B^T + AXB$

Define $\begin{cases} \operatorname{trace}(AXBX^T) = H(W,Y) = \operatorname{trace} AYW^T \\ W = f(X) = XB^T \\ Y = X \end{cases}$

$$\frac{\partial H}{\partial X} = \frac{\partial H}{\partial W}\frac{\partial W}{\partial X} + \frac{\partial H}{\partial Y}\frac{\partial Y}{\partial X} =$$

$$= AYB + A^T W =$$

$$= AXB + A^T X B^T$$