Joseph High
9E1FDC

EN.553.633 Homework # 13

A.2 (c.)

Pf: $f(x|y) = \frac{f(x,y)}{f(y)} = \frac{1}{f(y)} c_1 \exp\left\{-\frac{1}{2}(x^T-\mu_1^T \; y^T-\mu_2^T)\Sigma^{-1}\begin{pmatrix}x-\mu_1\\ y-\mu_2\end{pmatrix}\right\}$

(*) $$= \frac{c_1}{f(y)} \cdot \exp\left\{-\frac{1}{2}[(x-\mu_1)^T \; (y-\mu_2)^T]\Sigma^{-1}\begin{pmatrix}x-\mu_1\\ y-\mu_2\end{pmatrix}\right\}$$

and by part (b) (from HW # 12) we have the following for any vectors u and v:

$$(u^T \; v^T)\Sigma^{-1}\begin{pmatrix}u\\ v\end{pmatrix} = (u^T - v^T S^T)\tilde{\Sigma}^{-1}(u - Sv) + v^T \Sigma_{22}^{-1} v$$

Using this identity on the exponent above we have

$$[(x-\mu_1)^T \; (y-\mu_2)^T]\Sigma^{-1}\begin{bmatrix}x-\mu_1\\ y-\mu_1\end{bmatrix} = [(x-\mu_1)^T - (y-\mu_2)^T S^T]\tilde{\Sigma}^{-1}[(x-\mu_1) - S(y-\mu_2)] + (y-\mu_2)^T\Sigma_{22}^{-1}(y-\mu_2)$$

$$= [x^T - (\mu_1^T + (y-\mu_2)^T S^T]\tilde{\Sigma}^{-1}[x - (\mu_1 + S(y-\mu_2))] + (y-\mu_2)^T\Sigma_{22}^{-1}(y-\mu_2)$$

$$= [x^T - (\mu_1 + S(y-\mu_2))^T]\tilde{\Sigma}^{-1}[x - (\mu_1 + S(y-\mu_2))] + (y-\mu_2)^T\Sigma_{22}^{-1}(y-\mu_2)$$

$$= [x^T - \tilde{\mu}^T]\tilde{\Sigma}^{-1}[x - \tilde{\mu}] + (y-\mu_2)^T\Sigma_{22}^{-1}(y-\mu_2)$$

(where $\tilde{\mu} = \mu_1 + S(y-\mu_2)$)

Thus, by (*) above

$$f(x|y) = \frac{c_1}{f(y)}\exp\left\{-\frac{1}{2}[(x-\mu_1)^T, (y-\mu_2)^T]\Sigma^{-1}\begin{bmatrix}x-\mu_1\\ y-\mu_2\end{bmatrix}\right\} =$$

$$= \frac{c_1}{f(y)}\exp\left\{-\frac{1}{2}(x^T-\tilde{\mu}^T)\tilde{\Sigma}^{-1}(x-\tilde{\mu})\right\}\exp\left\{-\frac{1}{2}(y-\mu_2)^T\Sigma_{22}^{-1}(y-\mu_2)\right\}$$

$$= c_2(y)\exp\left\{-\frac{1}{2}(x^T-\tilde{\mu}^T)\tilde{\Sigma}^{-1}(x-\tilde{\mu})\right\}$$

where $c_2(y) = \frac{c_1(y)}{f(y)}\exp\left\{-\frac{1}{2}(y-\mu_2)^T\Sigma_{22}^{-1}(y-\mu_2)\right\}$

$//$ $\square$

A.) $x_{k+1} = f_k(x_k) + w_k \quad ; \quad w_k \sim N(0, Q_k)$
$z_k = H_k x_k + v_k \quad ; \quad v_k \sim N(0, R_k)$

and $x_k \mid x_{k-1}, z_k \sim N(a_k, \Sigma_k)$

$$\left[\begin{array}{c} \underline{\text{WTS:}}\ \Sigma_k = Q_{k-1} - Q_{k-1} H_k^T S_k^{-1} H_k Q_{k-1} \\ \text{where } S_k = H_k Q_{k-1} H_k^T + R_k \end{array}\right]$$

<u>Proof</u>: W.l.o.g. suppose $x_k \in \mathbb{R}^n$ and $z_k \in \mathbb{R}^m$, $m \leq n$.

By <u>Bayes' Rule</u> we have $p(x_k \mid x_{k-1}, z_k) = \dfrac{p(x_k, x_{k-1}, z_k)}{p(x_{k-1}, z_k)}$
where

$$p(x_k, x_{k-1}, z_k) = p(x_k, x_{k-1})\, p(z_k \mid x_k, x_{k-1})$$
$$= p(x_{k-1})\, p(x_k \mid x_{k-1})\, p(z_k \mid x_k, x_{k-1})$$

and $p(x_{k-1}, z_k) = p(x_{k-1})\, p(z_k \mid x_{k-1})$

$$\Rightarrow p(x_k \mid x_{k-1}, z_k) = \frac{\cancel{p(x_{k-1})}\, p(x_k \mid x_{k-1})\, p(z_k \mid x_k, x_{k-1})}{\cancel{p(x_{k-1})}\, p(z_k \mid x_{k-1})}$$

$$= \frac{p(x_k \mid x_{k-1})\, p(z_k \mid x_k)}{p(z_k \mid x_{k-1})} \propto p(x_k \mid x_{k-1})\, p(z_k \mid x_k)$$

(where $p(z_k \mid x_k, x_{k-1}) = p(z_k \mid x_k)$ by Markov property)

Now, from $\begin{cases} x_k = f_{k-1}(x_{k-1}) + w_{k-1}, & w_{k-1} \sim N(0, Q_{k-1}) \\ z_k = H_k x_k + v_k, & v_{k-1} \sim N(0, R_k) \end{cases}$

<u>we have</u>,

$$\mathbb{E}[x_k \mid x_{k-1}] = \mathbb{E}[f_{k-1}(x_{k-1}) \mid x_{k-1}] + \underbrace{\mathbb{E}[w_{k-1} \mid x_{k-1}]}_{=0}$$
$$= f_{k-1}(x_{k-1})$$

$$\mathrm{Var}[x_k \mid x_{k-1}] = \mathrm{Var}[f_{k-1}(x_{k-1}) + w_{k-1} \mid x_{k-1}] = \mathrm{Var}(w_{k-1}) = Q_{k-1}$$

Then since w_k is normally distributed and since we are conditioning on x_{k-1} $\Rightarrow$ $x_k \mid x_{k-1} \sim N(f_{k-1}(x_{k-1}), Q_{k-1})$

$\Rightarrow$

Joseph High

Now,

$$\mathbb{E}[z_k \mid x_k] = \mathbb{E}[H_k x_k + v_k \mid x_k] = \mathbb{E}[H_k x_k \mid x_k] + \underbrace{\mathbb{E}[v_k \mid x_k]}_{=0}$$

$$= H_k x_k$$

$$Var[z_k \mid x_k] = Var[H_k x_k + v_k \mid x_k]$$

$$= \underbrace{Var(H_k x_k \mid x_k)}_{=0} + Var(v_k \mid x_k) = R_k$$

and since v_k normally distributed and $H_k x_k$ fixed when conditioning on x_k $\Longrightarrow$ $z_k \mid x_k \sim N(H_k x_k, R_k)$.

We now have that $x_k \mid x_{k-1} \sim N(f_{k-1}(x_{k-1}), Q_{k-1})$
$z_k \mid x_k \sim N(H_k x_k, R_k)$

Then,

$$p(x_k \mid x_{k-1}, z_k) \propto p(x_k \mid x_{k-1})\, p(z_k \mid x_k) \propto$$

$$\propto \exp\left\{-\tfrac{1}{2}(x_k - f_{k-1}(x_{k-1}))^T Q_{k-1}^{-1}(x_k - f_{k-1}(x_{k-1}))\right\} \exp\left\{-\tfrac{1}{2}(z_k - H_k x_k)^T R_k^{-1}(z_k - H_k x_k)\right\}$$

$$= \exp\left\{-\tfrac{1}{2}\left[(x_k - f_{k-1}(x_{k-1}))^T Q_{k-1}^{-1}(x_k - f_{k-1}(x_{k-1})) + (z_k - H_k x_k)^T R_k^{-1}(z_k - H_k x_k)\right]\right\}$$

Focusing on only the exponent we have

$$(x_k - f_{k-1}(x_{k-1}))^T Q_{k-1}^{-1}(x_k - f_{k-1}(x_{k-1})) + (z_k - H_k x_k)^T R_k^{-1}(z_k - H_k x_k) =$$

$$= x_k^T Q_{k-1}^{-1} x_k - 2 x_k^T Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + f_{k-1}^T(x_{k-1}) Q_{k-1}^{-1} f_{k-1}(x_{k-1}) +$$

$$+ z_k^T R_k^{-1} z_k - 2 x_k^T H_k^T R_k^{-1} z_k + x_k^T H_k^T R_k^{-1} H_k x_k$$

$$= x_k^T \left[Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k\right] x_k - 2 x_k^T \left[Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + H_k^T R_k^{-1} z_k\right]$$

$$+ f_{k-1}^T(x_{k-1}) Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + z_k^T R_k^{-1} z_k$$

$\Longrightarrow$

cont'd...

Completing the square we have

$$x_k^T\left[Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k\right] x_k - 2x_k^T\left[Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + H_k^T R_k^{-1} z_k\right]$$

$$\pm \left(\left[Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k\right]^{-1}\left[Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + H_k^T R_k^{-1} z_k\right]\right)^T \left[Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k\right] \cdot$$

$$\cdot \left(\left[Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k\right]^{-1}\left[Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + H_k^T R_k^{-1} z_k\right]\right)$$

$$+ f_{k-1}^T(x_{k-1}) Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + z_k^T R_k^{-1} z_k$$

(*) Let $A_k = Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k$ (going to show $A_k = \Sigma_k^{-1}$)
and $B_k = Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + H_k^T R_k^{-1} z_k$
Then the above is

$$= x_k^T A_k x_k - 2x_k^T B_k + (A_k^{-1} B_k)^T A_k (A_k^{-1} B_k) + C$$

$$= (x_k - B_k)^T A_k (x_k - B_k) + C$$

(where $C = f_{k-1}^T(x_{k-1}) Q_{k-1}^{-1} f_{k-1}(x_{k-1}) + z_k^T R_k^{-1} z_k - (A_k^{-1} B_k)^T A_k (A_k^{-1} B_k)$)

Therefore, we have

$$p(x_k \mid x_{k-1}, z_k) = \frac{1}{p(z_k \mid x_{k-1})} p(x_k \mid x_{k-1})\, p(z_k \mid x_k) =$$

$$= \frac{1}{p(z_k \mid x_{k-1})}\left[\frac{1}{\sqrt{(2\pi)^{n+m} \det(Q_k) \det(R_k)}} \exp\left\{-\tfrac{1}{2}(x_k - B_k)^T A_k (x_k - B_k)\right\} \cdot \exp\left\{-\tfrac{1}{2} C\right\}\right]$$

$$= \frac{D_k}{p(z_k \mid x_{k-1})} \exp\left\{-\tfrac{1}{2}(x_k - B_k)^T A_k (x_k - B_k)\right\}$$

$$\left(\text{where } D_k = \frac{\exp\left\{-\tfrac{1}{2} C\right\}}{\sqrt{(2\pi)^{n+m} \det(Q_k) \det(R_k)}}\right) \Longrightarrow$$

Joseph High

Now,

$$1 = \int_{\mathbb{R}^n} p(x_k \mid x_{k-1}, z_k)\, dx_k = \int_{\mathbb{R}^n} \frac{D_k}{p(z_k \mid x_{k-1})} \exp\left\{-\tfrac{1}{2}(x_k - B_k)^T A_k (x_k - B_k)\right\} dx_k$$

$$\Rightarrow p(z_k \mid x_{k-1}) = D_k \int_{\mathbb{R}^n} \exp\left\{-\tfrac{1}{2}(x_k - B_k)^T A_k (x_k - B_k)\right\} dx_k =$$

$$= D_k \sqrt{(2\pi)^n \det(A_k^{-1})}$$

$$\Rightarrow p(x_k \mid x_{k-1}, z_k) = \frac{p(x_k \mid x_{k-1})\, p(z_k \mid x_k)}{p(z_k \mid x_{k-1})} =$$

$$= \frac{D_k \exp\left\{-\tfrac{1}{2}(x_k - B_k)^T A_k (x_k - B_k)\right\}}{D_k \sqrt{(2\pi)^n \det(A_k^{-1})}}$$

$$= \frac{1}{\sqrt{(2\pi)^n \det(A_k^{-1})}} \exp\left\{-\tfrac{1}{2}(x_k - B_k)^T A_k (x_k - B_k)\right\}$$

Since $x_k \mid x_{k-1}, z_k \sim N(a_k, \Sigma_k)$, then this implies that

$A_k = \Sigma_k^{-1}$ and $a_k = B_k$ and by (*)
$A_k = Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k$

$$\Rightarrow \Sigma_k^{-1} = Q_{k-1}^{-1} + H_k^T R_k^{-1} H_k$$

By the Matrix Inversion Lemma (Sherman-Morrison-Woodberry)

$$\boxed{\begin{aligned} \Sigma_k &= Q_{k-1} - Q_{k-1} H_k^T (H_k Q_{k-1} H_k^T + R_k)^{-1} H_k Q_{k-1} \\ &= Q_{k-1} - Q_{k-1} H_k^T S_k^{-1} H_k Q_{k-1} \\ &\text{where } S_k = H_k Q_{k-1} H_k^T + R_k \end{aligned}}$$

/// $\square$