Nonlinear Filtering Concepts and Engineering Applications 1st edition Raol Solutions Manual

SKU: 9781498745178-SOLUTIONS

Solutions of the Exercises for Section II
(Chapters 5-7)
(For the CRC Book - Nonlinear Filtering: Concepts and Engineering Applications)

II.1 Let $\hat{z} = \hat{y}$; Then,

$$\text{cov}(z - \hat{z}) = \text{cov}(y + v - \hat{y})$$

$$E\{(z - \hat{z})(z - \hat{z})^T\} = E\{(y + v - \hat{y})\ (y + v - \hat{y})^T\}$$

$$= E\{(y - \hat{y})\ (y - \hat{y})^T\} + E\{vv^T\}$$

Here, we assume that the measurement residuals $(y - \hat{y})$ and measurement noise v are uncorrelated. Then, we get $\text{cov}(z - \hat{z}) = \text{cov}(y - \hat{y}) + R$

II.2

$$\phi = e^{A\Delta t} = I + A\Delta t + \frac{A^2\Delta t^2}{2!}; \quad \phi = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \Delta t \\ 0 & -a\Delta t \end{bmatrix} + \begin{bmatrix} 0 & -a\Delta t^2/2 \\ 0 & a^2\Delta t^2/2 \end{bmatrix}; \quad \phi = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 - a\Delta t \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \Delta t - a\Delta t^2/2 \\ 0 & 1 - a\Delta t + a^2\Delta t^2/2 \end{bmatrix}$$

II.3 Since w is unknown,

$$\tilde{x}(k+1) = \phi\, \hat{x}(k) + bu$$

$$\tilde{\sigma}_x^2 = \phi\, \hat{\sigma}_x^2 \phi^T + g^2\sigma_w^2$$

Since 'u' is a deterministic input, it does not appear in covariance equation of the state error. The measurement update equations are:

$$r(k+1) = z(k+1) - c\tilde{x}(k+1)$$

$$K = \frac{\tilde{\sigma}_x^2 c}{(c^2\tilde{\sigma}_x^2 + \sigma_v^2)}$$

$$\hat{\sigma}_x^2 = (1 - Kc)\tilde{\sigma}_x^2$$

II.4 We have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

Since a_{ij} are unknown parameters, we consider them as extra states:

$\dot{x}_1 = a_{11}x_1 + a_{12}x_2 + w_1$

$\dot{x}_2 = a_{21}x_1 + a_{22}x_2 + w_2$

$\dot{x}_3 = 0$

$\dot{x}_4 = 0$

$\dot{x}_5 = 0$

$\dot{x}_6 = 0$

with $x_3 = a_{11}$, $x_4 = a_{12}$, $x_5 = a_{21}$ and $x_6 = a_{22}$

We finally get,

$\dot{x}_1 = x_1x_3 + x_2x_4 + w_1$

$\dot{x}_2 = x_1x_5 + x_2x_6 + w_2$

$\dot{x}_3 = 0$

$\dot{x}_4 = 0$

$\dot{x}_5 = 0$

$\dot{x}_6 = 0$

Then $\dot{x} = f(x) + w$, where f is a nonlinear vector valued function.

II.5 Let the linear model be given by

$\dot{x} = A_1x + Gw_1$

$z = Hx + v$

By putting the equations for x and v together, we get

$\dot{x} = A_1x + Gw_1$

$\dot{v} = A_2v + w_2$

We define joint vector $\begin{bmatrix} x \\ v \end{bmatrix}$ to get

$$\begin{bmatrix} \dot{x} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} + \begin{bmatrix} G & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \text{and} \quad z = \begin{bmatrix} H & I \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}$$

We see that the vector v, which is correlated noise, is now augmented to the state vector x and hence, there is no measurement noise term in the measurement equation. This amounts to the situation that the measurement noise in the composite equation is zero, leading to $R^{-1} \to \infty$, and hence the Kalman gain will be ill-conditioned. Thus, this formulation is not directly suitable in KF.

II.6 The residual error is the general term arising from say $z - \hat{z}$ (see Chapter 2).

Prediction error: Consider $\tilde{x}(k+1) = \phi\,\hat{x}(k)$; Then, $z(k+1) - H\tilde{x}(k+1)$ is the prediction error, since $\hat{z} = H\tilde{x}(k+1)$ is the predicted measurement based on the estimate $\tilde{x}$.

Filtering error: Assume that we have already obtained the estimate of the state after incorporating the measurement data $\hat{x}(k+1) = \tilde{x}(k+1) + K(z(k+1) - H\tilde{x}(k+1))$ Then, the quantity can be considered as filtering error $z(k+1) - H\hat{x}(k+1)$, since the error is obtained after using the $\hat{x}(k+1)$, the filtered state estimate.

II.7 The main reason is that the measurement data occurring at arbitrary intervals can be easily incorporated in the Kalman filtering algorithm.

II.8 The quantity '*S*' is the theoretical (prediction) covariance of the residuals, whereas the $\text{cov}(rr^T)$ is the actual computed covariance of the residuals. For proper tuning of KF, both should match. In fact the computed residuals should lie within the theoretical bounds predicted by '*S*'.

II.9

Let
$$x(k+1) = \phi\, x(k) + gw(t)$$
$$z(k) = cx(k) + v(k)$$

Then,
$$\tilde{p} = \phi\, \hat{p}\, \phi^T + g^2\sigma_w^2$$
$$\hat{p} = (1 - Kc)\tilde{p}$$

Also $$K = \tilde{p}c(c^2\tilde{p} + \sigma_v^2)^{-1} = \frac{\tilde{p}c}{\tilde{p}c^2 + \sigma_v^2}$$

and hence $$\hat{p} = \left(1 - \frac{\tilde{p}c^2}{\tilde{p}c^2 + \sigma_v^2}\right)\tilde{p} = \frac{\tilde{p}\sigma_v^2}{c^2\tilde{p} + \sigma_v^2} = \frac{\tilde{p}}{1 + \frac{c^2\tilde{p}}{\sigma_v^2}}$$

If σ_v^2 is low, then $\hat{p}$ is low, meaning thereby, we have more confidence in the estimates.

We can also rearrange $\hat{p}$ as $\hat{p} = \frac{\sigma_v^2}{c^2 + \frac{\sigma_v^2}{\tilde{p}}}$; then if $\tilde{p}$ is low, then $\hat{p}$ is low. If the observation model is strong, then also $\hat{p}$ is low.

II.10

$$\sigma_x^2 = E\{(x - E\{x\})^2\}$$
$$= E\{x^2 - 2xE\{x\} + (E\{x\})^2\}$$
$$= E\{x^2\} + (E\{x\})^2 - 2E\{x\}E\{x\}$$
$$\sigma_x^2 = E\{x^2\} - (E\{x\})^2$$

II.11 Std.$= \sqrt{\sigma_x^2} = \sigma_x$ = RMS if the random variable has zero mean.

II.12 The residual is given as $r(k) = z(k) - H\tilde{x}(k)$, where $\tilde{x}(k)$ is the time propagated estimates of KF. We see that z(k) is the current measurement and the term $H\tilde{x}(k)$ is the effect of past or old information derived from the past measurements. Thus, the term $r(k)$ generates new information and, hence, it is called 'innovations' process.

II.13 It is based on the principle that any arbitrary pdf can be approximated by a finite sum of weighted Gaussian pdfs.

II.14 It is based on the expansion of the nonlinear function f and h, in terms of Fourier-Hermite series.

II.15 A complete expansion can be written as $p(x) = \sum_{i=0}^{\infty} a_i \varphi_i(x)$ which can be separated into two parts as

$$\begin{aligned} p(x) &= \sum_{i=0}^{\infty} a_i \varphi_i(x) \\ &= \sum_{i=0}^{N} a_i \varphi_i(x) + \sum_{i=N+1}^{\infty} a_i \varphi_i(x) \end{aligned} = \hat{p}(x) + errror_p(x) \approx \hat{p}(x).$$

II.16 $p(v) = \dfrac{p_1}{\sigma_1(v)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v-m_1(v)}{\sigma_1(v)}\right)^2} + \dfrac{p_2}{\sigma_2(v)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v-m_2(v)}{\sigma_2(v)}\right)^2}$ with appropriate means and variances of the individual component of v(.).

II.17 $m_y = \dfrac{2}{c-b} m_x - \dfrac{b+c}{c-b}; \quad \sigma_y^2 = \left(\dfrac{2}{c-b}\right)^2 \sigma_x^2.$

II.18 It means number of its all positive, negative and zero eigenvalues.

II.19 One simple way is to factorize the matrix as $P = UDU^T$, then since the matrix D is diagonal, the number of positive diagonal elements and the negative diagonal elements gives the number of positive eigenvalues and negative eigenvalues of the matrix P and hence the inertia.

II.20 It is J-square root of $\begin{bmatrix} I & 0 \\ 0 & -\gamma^2 I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & \gamma I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \gamma I \end{bmatrix}$; $J = (I \oplus -I)$, J-unitary transformation.

II.21 No. the reason is that 'd' is deterministic discrepancy (in the model). It is a time-history, which is estimated by IE method. As such, it is not a random variable. We can regard Q^{-1}, perhaps, as some form of information matrix, deriving a hint from the fact that in GLS, W is used and if $W = R^{-1}$, we get the so-called Markov estimates. And since R^{-1} can be regarded as some form of information matrix (R being the covariance matrix), Q^{-1} may be called as information matrix. It is very important tuning parameter for the algorithm.

II.22 The idea is to have correct estimates of the state as the integration of the state's dynamic equation, and simultaneously the correct representation of model error estimation 'd'. In order that both the things happen, the state constraint equation and the cost function equation should be satisfied. The estimate should evolve such that by proper tuning of Q we obtain good estimate of 'd'. In J equation, the second term is also to be minimized thereby saying that just accurate 'd' needs to be obtained by choosing appropriate penalty by Q. Too much or too less d will not obtain correct estimate of x.

II.23 Use of R^{-1} normalizes the cost function, since $E\{(y-\hat{y})(y-\hat{y})^T\}$ is a covariance matrix of residuals and R is the measurement noise covariance matrix. Then $E\{(y-\hat{y})^T R^{-1}(y-\hat{y})\}$ will be normalized sum of squares of residuals.

II.24 In order to determine additional model from 'd', the LS method will be used and the residuals arising from the term will be treated as measurement noise.

II.25 We have the cost function as

$$J = \sum_{k=0}^{N} (z(k)-\hat{x}(k))^2(\sigma^2)^{-1} + \int_{t_0}^{t_f} d^2 Q dt$$

The Hamiltonian is $H = \Psi(x(t),u(t),t) + \lambda^T(t) f(x(t),u(t),t)$; $H = d^2 Q + \lambda^T d$

II.26 The term $\varphi\{x(t(N)),t(N)\}$ in the cost function can be replaced by

$$\sum_{k=0}^{N} \varphi_k\{x(t(k),t(k))\}$$

II.27 We have $\frac{\partial H}{\partial x} = -\lambda^T(t)\frac{\partial f}{\partial x} + \frac{\partial \psi}{\partial x}$; from Pontryagin's necessary condition, we have $\frac{\partial H}{\partial x} = \dot{\lambda}^T$ and hence $\dot{\lambda}^T = -\lambda^T(t)\left(\frac{\partial f}{\partial x}\right) + \left(\frac{\partial \psi}{\partial x}\right)$; which can be rewritten as

$\dot{\lambda} = -\left(\frac{\partial f}{\partial x}\right)^T \lambda^T(t) + \left(\frac{\partial \psi}{\partial x}\right)^T$; $\dot{\lambda}(t) = A\lambda(t) + u(t)$ with appropriate equivalence.

It must be noted that since f_x and ψ_x are matrices evaluated at estimated state $\hat{x}$, we see that the co-state equation has similar structure as the state equation.

II.28 These models are

$p_b(x(k+1)/x(k)) = p_{w(k)}(x(k+1) - f(x(k),b,k))$

$p_b(z(k)/x(k)) = p_{e(k)}(z(k) - h(x(k),b,k))$

.

II.29 i) represent the state error covariances by Gaussian pdfs, ii) generate the (N) samples of x(k) from the Gaussian pdf., iii) generate the samples of the process noise, iv) then obtain the samples at x(k+1), by using the nonlinear system equations, v) then from this obtain the sample mean as the state estimate, vi) then {from iv) and v)} obtain the sample covariance as P(k+1/k), vii) generate the samples for x(k) from the Gaussian distribution, viii) generate the noise samples v(k), ix) from this and the measurement model h, obtain the z(k), x) then determine the mean of z(k) by using the data from {ix)}, xi) compute the estimate covariances P_{xz}, and $P_{zz,}$ the ratio of P_{xz}/P_{zz} gives the gain, xii) then one can obtain the measurement update using the predicted estimate, the gain and residuals.

II.30 It is the maximum singular value/gain of the transfer function matrix, and it signifies the maximum transfer of the collective error energies to the output estimation error energy. Then main goal in the H-infinity robust model error estimation is to minimize this gain.

==========