`From (16) we can simply plug in` $c_{2k} = 0$, $c_{2k-1} = 1$, $A_{k+1|k} = a$, $B_k = 1$ `to the Kalman filtering equations in Sec. 6.7 of the notes to obtain the optimal Kalman filter. The KF does NOT reduce to a LTI as` $n \to \infty$ `since the optimal filter` $h(n,k)$ `of part (a) is not LTI. Note: it is also possible to solve this problem by defining a vector signal state` $\underline{s}_{2n} = [s_{2n}, s_{2n-1}]^T$.

6.16 Derive the minimum mean square error expression (**??**) for the Wiener filter and use it to find the MSE of the optimal predictors of Examples **??** and **??**.

6.17 In this exercise you will explore the extended Kalman filter (EKF) for non-linear state dynamics. Similarly to Section 6.7.1. of the notes we assume that the observations y_k obey a dynamical model, except that here we assume that the state can evolve *non-linearly* and that the additive noises are *Gaussian*:

$$\begin{aligned} y_k &= s_k + v_k \\ s_k &= \underline{c}^T \underline{\xi}_k \qquad\qquad , k = 0, 1, \ldots \\ \underline{\xi}_{k+1} &= \underline{g}(\underline{\xi}_k) + \mathbf{B}_k \underline{w}_k \end{aligned}$$

where $\underline{g}$ is a possibly non-linear p-dimensional function of the p-dimensional state vector $\underline{\xi}_k$, v_k and $\underline{w}_k$ are mutually independent zero mean temporally uncorrelated (white) noises which are *Gaussian* distributed with variance σ_v^2 and covariance matrix $\mathbf{R}_w$, respectively. All other assumptions on the model are identical to those made in Section 6.7.1. Let the posterior density of the state $\underline{\xi}_k$ given the observation sequence $\mathcal{Y}_l = \{y_l, y_{l-1}, \ldots, y_0\}$ up to time l be denoted $f_{\underline{\xi}_k|\mathcal{Y}_l}$.

(a) Show that in the linear case ($\underline{g}(\underline{\xi}_k) = \mathbf{A}\underline{\xi}_k$) the posterior density $f_{\underline{\xi}_k|\mathcal{Y}_k}$ is Gaussian:

$$f_{\underline{\xi}_k|\mathcal{Y}_l}(\underline{\xi}_k|\mathcal{Y}_k) = \frac{1}{|\mathbf{R}_{\tilde{\xi}}(k|k)|^{\frac{1}{2}}(2\pi)^{p/2}} \exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T \mathbf{R}_{\tilde{\xi}}^{-1}(k|k)(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right)$$

where $\hat{\underline{\xi}}_{k|k}$ is the Kalman filter's state estimator and $\mathbf{R}_{\tilde{\xi}}(k|k) = E[(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T]$ is the state estimator error covariance matrix.

`Soln:`

`We know that for jointly Gaussian r.v.s` W, Z `the conditional distribution of` Z `given` W `is` $\mathcal{N}(E[Z|W], \mathrm{cov}(Z|W))$. `Therefore, all that needs to be found is the conditional mean and conditional covariance matrix. However, by the same argument as in Sec. 5.4.3 we know that in this case` $E[Z|W]$ `is identical to the linear minimum mean squared error estimator of` Z `given` W, `which, for` $Z = \underline{\xi}_k$ `and` $W = \mathcal{Y}_k$ `is the output of the Kalman filter. Furthermore, the conditional covariance` $\mathrm{cov}(\underline{\xi}_k - \hat{\underline{\xi}}_k|\mathcal{Y}_k) = \mathrm{cov}(\underline{\xi}_k - \hat{\underline{\xi}}_k)$ `since (projection theorem) the error` $\underline{\xi}_k - \hat{\underline{\xi}}_k$ `is uncorrelated with past observations` $\mathcal{Y}_k$ `and therefore, since the error and the past observations are jointly Gaussian, they are statistically independent. Hence the conditional covariance is equal to the unconditional error covariance` $R_{\tilde{\xi}}(k|k)$ `of the Kalman state estimator.`

(b) Using Bayes formula show that for arbitrary linear or non-linear function $\underline{g}$ the state posterior density $\rho_k(\underline{\xi}_k) \stackrel{\text{def}}{=} f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k)$ obeys the recursion

$$\rho_{k+1}(\underline{\xi}_{k+1}) = \frac{f_{y_{k+1}|\underline{\xi}_{k+1}}(y_{k+1}|\underline{\xi}_{k+1})}{f_{y_{k+1}|\mathcal{Y}_k}(y_{k+1}|\mathcal{Y}_k)} \int_{\mathbb{R}^p} f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)\rho_k(\underline{\xi}_k)d\underline{\xi}_k. \tag{17}$$

where the density $h_1(\underline{u}) = f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{u}|\underline{\xi}_k)$ has the form of a multivariate Gaussian density with mean parameter $\underline{g}(\underline{\xi}_k)$ and covariance matrix parameter $\mathbf{B}\mathbf{R}_w\mathbf{B}^T$ and the density $h_2(z) = f_{y_{k+1}|\underline{\xi}_{k+1}}(z|\underline{\xi}_{k+1})$ has the form of a univariate Gaussian density with mean parameter $\underline{c}^T\underline{\xi}_{k+1}$ and variance parameter σ_v^2.

Soln:

Start out with Bayes formula

$$\begin{aligned} f_{\underline{\xi}_{k+1}|\mathcal{Y}_{k+1}}(\underline{\xi}_{k+1}|\mathcal{Y}_{k+1}) &= f_{\underline{\xi}_{k+1}|\mathcal{Y}_{k+1}}(\underline{\xi}_{k+1}|y_{k+1},\mathcal{Y}_k) \\ &= \frac{f_{\underline{\xi}_{k+1},y_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1},y_{k+1}|\mathcal{Y}_k)}{f_{y_{k+1}|\mathcal{Y}_k}(y_{k+1}|\mathcal{Y}_k)} \end{aligned}$$

Next we express the numerator as

$$\begin{aligned} f_{\underline{\xi}_{k+1},y_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1},y_{k+1}|\mathcal{Y}_k) &= f_{y_{k+1}|\underline{\xi}_{k+1},\mathcal{Y}_k}(y_{k+1}|\underline{\xi}_{k+1},\mathcal{Y}_k)f_{\underline{\xi}_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}|\mathcal{Y}_k) \\ &= f_{y_{k+1}|\underline{\xi}_{k+1}}(y_{k+1}|\underline{\xi}_{k+1})f_{\underline{\xi}_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}|\mathcal{Y}_k) \end{aligned}$$

where in the last step we used the fact that, given $\underline{\xi}_{k+1}$, y_{k+1} is independent of the past $\mathcal{Y}_k$ since the noise v_k is white (recall the model $y_{k+1} = \underline{c}^T\underline{\xi}_{k+1} + v_k$). As y_{k+1} is Gaussian distributed with mean $\underline{c}^T\underline{\xi}_{k+1}$ and variance σ_v^2 the first factor is the Gaussian density h_1. As for the second factor use the *law of total probablity*

$$\begin{aligned} f_{\underline{\xi}_{k+1}|\mathcal{Y}_k}(\underline{\xi}_{k+1}|\mathcal{Y}_k) &= \int_{\mathbb{R}^p} f_{\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k)d\underline{\xi}_k \\ &= \int_{\mathbb{R}^p} f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k)f_{\underline{\xi}_{k+1}|\underline{\xi}_k,\mathcal{Y}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k,\mathcal{Y}_k)d\underline{\xi}_k \end{aligned}$$

Now, as $\underline{\xi}_{k+1}$ is independent of $\mathcal{Y}_k$ given $\underline{\xi}_k$ (recall that $\underline{w}_k$ and v_k are independent Gaussian white noises), the second factor in the integrand is simply $f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)$ which is a multivariate Gaussian density as claimed. This establishes the stated recursive formula.

(c) Specialize the recursion (17) to the linear case ($\underline{g}(\underline{\xi}_k) = \mathbf{A}\underline{\xi}_k$) and derive a relation for the quadratic form $(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1})^T\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1})$ by equating the exponent on the left hand side of the equation to the exponent on the right hand side (Hint: completion of the square in the integrand).

Soln:

Completion of the square in the integrand of (17) gives the expression

$$\begin{aligned}
&f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)\rho_k(\underline{\xi}_k) \qquad (18)\\
&= c\exp\left(-\frac{1}{2}(\underline{\xi}_{k+1}-\mathbf{A}\hat{\underline{\xi}}_{k|k}-\mathbf{A}(\underline{\xi}_k-\hat{\underline{\xi}}_{k|k}))^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1}-\mathbf{A}\hat{\underline{\xi}}_{k|k}-\mathbf{A}(\underline{\xi}_k-\hat{\underline{\xi}}_{k|k}))\right)\\
&\exp\left(-\frac{1}{2}(\underline{\xi}_k-\hat{\underline{\xi}}_{k|k})^T\mathbf{R}_{\tilde{\underline{\xi}}}^{-1}(k|k)(\underline{\xi}_k-\hat{\underline{\xi}}_{k|k})\right)\\
&= c\exp\left(-\frac{1}{2}(\underline{\xi}_k-\hat{\underline{\xi}}_{k|k}-\mathbf{Q}^{-1}\underline{u}_k)^T\mathbf{Q}(\underline{\xi}_k-\hat{\underline{\xi}}_{k|k}-\mathbf{Q}^{-1}\underline{u}_k)\right)\exp\left(-\frac{1}{2}(q_1-q_2)\right)
\end{aligned}$$

where c is an unimportant constant,

$$\begin{aligned}
\underline{u} &= \mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1}-\mathbf{A}\hat{\underline{\xi}}_{k|k})\\
\mathbf{Q} &= \mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A}+\mathbf{R}_{\tilde{\underline{\xi}}}^{-1}(k|k)\\
q_1(\underline{\xi}_{k+1}) &= (\underline{\xi}_{k+1}-\mathbf{A}\hat{\underline{\xi}}_{k|k})^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1}-\mathbf{A}\hat{\underline{\xi}}_{k|k})\\
q_2(\underline{\xi}_{k+1}) &= \underline{u}^T\mathbf{Q}^{-1}\underline{u}
\end{aligned}$$

The result of integration of (18) over $\underline{\xi}_k \in \mathbb{R}^p$ (recall that the Gaussian density integrates to 1) obtains the following expression for the integral in (17)

$$c_1\exp\left(-\frac{1}{2}(q_1-q_2)\right)$$

for some constant c_1. Now the exponent on the RHS of (17) can be easily found to be

$$-\frac{1}{2}\left((y_{k+1}-\underline{c}^T\underline{\xi}_{k+1})^2/\sigma_v^2-(y_{k+1}-\underline{c}^T\hat{\underline{\xi}}_{k+1|k+1})^2/\sigma^2+q_1(\underline{\xi}_{k+1})-q_2(\underline{\xi}_{k+1})\right) \qquad (19)$$

where $\sigma^2=\mathrm{var}(y_{k+1}-\underline{c}^T\hat{\underline{\xi}}_{k+1|k+1})=\underline{c}^T\mathbf{R}_{\tilde{\xi}}(k+1|k+1)\underline{c}+\sigma_v^2$. Thus we have the relation

$$\begin{aligned}
&(\underline{\xi}_{k+1}-\hat{\underline{\xi}}_{k+1|k+1})^T\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)(\underline{\xi}_{k+1}-\hat{\underline{\xi}}_{k+1|k+1}) = \qquad (20)\\
&\quad (y_{k+1}-\underline{c}^T\underline{\xi}_{k+1})^2/\sigma_v^2-(y_{k+1}-\underline{c}^T\hat{\underline{\xi}}_{k+1|k+1})^2/\sigma^2+q_1(\underline{\xi}_{k+1})-q_2(\underline{\xi}_{k+1})
\end{aligned}$$

(d) Using the relation found in part (c) derive Kalman filter equations specifying state estimator updates $\hat{\underline{\xi}}_{k|k}\to\hat{\underline{\xi}}_{k+1|k+1}$ and inverse covariance updates $\mathbf{R}_{\tilde{\underline{\xi}}}^{-1}(k|k)\to\mathbf{R}_{\tilde{\underline{\xi}}}^{-1}(k+1|k+1)$ in the following manner. To derive state update equation take derivative of relation (c) with respect to $\underline{\xi}_{k+1}$ and evaluate the resulting equation at $\underline{\xi}_{k+1}=\mathbf{A}\hat{\underline{\xi}}_{k|k}$. To derive covariance update equation take second derivative with respect to $\underline{\xi}_{k+1}$. You may find the Woodbury-Sherman-Morissey identity (Eq. (1) of Ch. 2) useful here. The right hand side of your inverse covariance update equation should be in a simple form involving only two simple additive terms.

Soln:

Differentiation of the LHS of (20) twice in the argument $\underline{\xi}_{k+1}$ yields $2\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)$. Likewise twice differentiating the RHS of (20) and equating to $2\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)$ gives

$$\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1) = [\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1} - [\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A}[\mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}\mathbf{A} + \mathbf{R}_{\tilde{\xi}}^{-1}(k|k)]^{-1}\mathbf{A}^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1} + \frac{\underline{c}\underline{c}^T}{\sigma_v^2}$$

Application of the WSM identity to the first two terms on the RHS gives the compact recursion for the inverse covariance

$$\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1) = [\mathbf{B}\mathbf{R}_w\mathbf{B}^T + \mathbf{A}\mathbf{R}_{\tilde{\xi}}(k|k)\mathbf{A}^T]^{-1} + \frac{\underline{c}\underline{c}^T}{\sigma_v^2} \tag{21}$$

Next we differentiate the LHS and RHS of (20) once wrt $\underline{\xi}_{k+1}$ and evaluate at $\underline{\xi}_{k+1} = \mathbf{A}\hat{\underline{\xi}}_{k|k}$ to obtain

$$\mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)(\mathbf{A}\hat{\underline{\xi}}_{k|k} - \hat{\underline{\xi}}_{k+1|k+1}) = -\frac{\underline{c}}{\sigma_v^2}(y_{k+1} - \underline{c}^T\mathbf{A}\hat{\underline{\xi}}_{k|k})$$

Which yields the Kalman filter recursion

$$\hat{\underline{\xi}}_{k+1|k+1} = \mathbf{A}\hat{\underline{\xi}}_{k|k} + \Gamma_{k|k}(y_{k+1} - \underline{c}^T\mathbf{A}\hat{\underline{\xi}}_{k|k})$$

where we have identified the Kalman gain

$$\Gamma_{k|k} = \mathbf{R}_{\tilde{\xi}}^{-1}(k+1|k+1)\frac{\underline{c}}{\sigma_v^2}.$$

6.18 This is a continuation of the previous exercise. Here we apply *Laplace's approximation* to the posterior distribution to obtain approximate state and covariance update recursions from (17). Laplace's approximation asserts that

$$f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k) \approx \frac{|\mathbf{F}_k|^{\frac{1}{2}}}{(2\pi)^{p/2}}\exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T\mathbf{F}_k(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right)$$

where $\hat{\underline{\xi}}_{k|k} = \text{argmax}_{\underline{\xi}_k} f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_k|\mathcal{Y}_k)$ is the MAP estimator of $\underline{\xi}_k$ given past observations $\mathcal{Y}_k$ and $\mathbf{F}_k = \mathbf{F}(\hat{\underline{\xi}}_{k|k})$ is the $p \times p$ *observed Fisher information matrix (FIM)* where, for $\underline{u} \in \mathbb{R}^p$

$$\mathbf{F}(\underline{u}) = -\nabla_{\underline{u}}^2 \ln f_{\underline{\xi}_k|\mathcal{Y}_k}(\underline{u}|\mathcal{Y}_k).$$

Using Laplace's approximation, and the approximation $\underline{g}(\underline{\xi}_k) = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \nabla\underline{g}(\hat{\underline{\xi}}_{k|k})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})$, in the integrand of the right hand side of (17), evaluate the integral by completion of the square.

Soln:

Similarly to (18) completion of the square gives the expression

$$
\begin{aligned}
& f_{\underline{\xi}_{k+1}|\underline{\xi}_k}(\underline{\xi}_{k+1}|\underline{\xi}_k)\rho_k(\underline{\xi}_k) \qquad (22)\\
&= c\exp\left(-\frac{1}{2}(\underline{\xi}_{k+1} - \underline{g}(\underline{\xi}_k))^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1} - \underline{g}(\underline{\xi}_k))\right)\\
& \exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T\mathbf{F}_k(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right)\\
&= c\exp\left(-\frac{1}{2}(\underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}) - \nabla\underline{g}_k(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k}))^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}) - \nabla\underline{g}_k(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k}))\right)\\
& \qquad \exp\left(-\frac{1}{2}(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})^T\mathbf{F}_k(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})\right) \qquad (23)
\end{aligned}
$$

where we have introduced the linear approximation $\underline{g}(\underline{\xi}_k) = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \nabla\underline{g}(\hat{\underline{\xi}}_{k|k})\underline{\xi}_k$ (we have defined $\nabla\underline{g}_k \overset{\text{def}}{=} \nabla\underline{g}(\hat{\underline{\xi}}_{k|k})$).
At this point, comparing the RHS of the second equality line of the above to the first line of (18) it is evident that we can make the following substitutions in (18)

$$
\begin{aligned}
\underline{\xi}_{k+1} - \mathbf{A}\hat{\underline{\xi}}_{k|k} &\rightarrow \underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}) \qquad (24)\\
\mathbf{A} &\rightarrow \nabla\underline{g}_k\\
\mathbf{R}_{\tilde{\xi}}^{-1}(k|k) &\rightarrow \mathbf{F}_k
\end{aligned}
$$

and reuse all of our previous work from (c) to obtain a analogous relation to (20):

$$
\begin{aligned}
& (\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1})^T\mathbf{F}_{k+1}(\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1}) \qquad (25)\\
&= (y_{k+1} - \underline{c}^T\underline{\xi}_{k+1})^2/\sigma_v^2 - (y_{k+1} - \underline{c}^T\hat{\underline{\xi}}_{k+1|k+1})^2/\sigma^2\\
&+ (\underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}))^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}))\\
&- (\underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}))^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}[\nabla\underline{g}_k][[\nabla\underline{g}_k]^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}[\nabla\underline{g}_k] + \mathbf{F}_k]^{-1}\\
&\cdot[\nabla\underline{g}_k]^T[\mathbf{B}\mathbf{R}_w\mathbf{B}^T]^{-1}(\underline{\xi}_{k+1} - \underline{g}(\hat{\underline{\xi}}_{k|k}))
\end{aligned}
$$

1. Using the results of part (a), and an analogous differentiation method to the one you used in part (d) of Ex 6.17, generate a recursion $\hat{\underline{\xi}}_{k|k} \rightarrow \hat{\underline{\xi}}_{k+1|k+1}$ for the MAP state estimator and a recursion $\mathbf{F}_k \rightarrow \mathbf{F}_{k+1}$ for the observed FIM (given your previous work on part (d) you should be able to quickly identify the solutions without much algebra). These equations represent the EKF filter. Represent your state estimator recursion in a form reminiscent of the Kalman filter, i.e.
$$\hat{\underline{\xi}}_{k+1|k+1} = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \Gamma_k\eta_k$$
where η_k is an analog to the Kalman innovation sequence and Γ_k is an analog to the Kalman Gain matrix (but which depends on $\hat{\underline{\xi}}_{k|k}$).

`Soln:`
`By twice differentiating (25) wrt` $\underline{\xi}_{k+1}$ `we obtain the analogous recursion to (21) which applies to the EKF`

$$\mathbf{F}_{k+1} = [\mathbf{B}\mathbf{R}_w\mathbf{B}^T + [\nabla \underline{g}_k]\mathbf{F}_k^{-1}[\nabla \underline{g}_k]^T]^{-1} + \frac{\underline{c}\underline{c}^T}{\sigma_v^2} \tag{26}$$

`To obtain a recursion for the EKF state estimator, analogously to Exercise 6.17, we again differentiate once wrt` $\underline{\xi}_{k+1}$ `but now evaluate the derivative at` $\underline{\xi}_{k+1} = \underline{g}(\hat{\underline{\xi}}_{k|k})$ `to obtain`

$$\mathbf{F}_{k+1}(\underline{g}(\hat{\underline{\xi}}_{k|k}) - \hat{\underline{\xi}}_{k+1|k+1}) = -\frac{\underline{c}}{\sigma_v^2}(y_{k+1} - \underline{c}^T\underline{g}(\hat{\underline{\xi}}_{k|k}))$$

`Which yields the EKF state recursion`

$$\hat{\underline{\xi}}_{k+1|k+1} = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \Gamma_{k|k}(y_{k+1} - \underline{c}^T\underline{g}(\hat{\underline{\xi}}_{k|k}))$$

`where we have identified the EKF gain`

$$\Gamma_{k|k} = \mathbf{F}_{k+1}^{-1}\frac{\underline{c}}{\sigma_v^2}.$$

2. Evaluate the EKF specified by the recursions found in (b) for the case of a scalar ($p = 1$) state ξ_k, scalar state noise w_k, scalar c, and the quadratic plant

$$g(\xi_k) = a\xi_k^2,$$

where $|a| < 1$. If the Fisher recursion is initialized by $F_{-1} > 0$ will the observed Fisher information F_k positive for all $k \geq 0$?

`Soln:`
`For this example things simplify considerably as` $\nabla g_k = 2a\hat{\xi}_{k|k}$. `Simply substitute this into the EKF equations to obtain the state update equation`

$$\hat{\xi}_{k+1|k+1} = a\hat{\xi}_{k|k}^2 + F_{k+1}^{-1}\frac{c}{\sigma_v^2}(y_{k+1} - ca\hat{\xi}_{k|k}^2)$$

`where the Fisher info is updated by`

$$F_{k+1} = F_k/(b^2 \ \ \sigma_w^2 F_k + (2a\hat{\xi}_{k|k})^2) + c^2/\sigma_v^2.$$

`As long as` $F_{-1} > 0$ `the right hand side of the above will remain positive for all` k `and therefore` $F_k > 0$ `for all` $k \geq 0$.

6.19 This is a continuation of the previous exercise. An approach which does not require making the supplemental approximation $\underline{g}(\underline{\xi}_k) = \underline{g}(\hat{\underline{\xi}}_{k|k}) + \nabla\underline{g}(\hat{\underline{\xi}}_{k|k})(\underline{\xi}_k - \hat{\underline{\xi}}_{k|k})$ is to apply the Laplace approximation to the posterior of the successive pair of states

$$f_{\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k}(\underline{\xi}_{k+1},\underline{\xi}_k|\mathcal{Y}_k) \approx \frac{|\mathbf{Q}_k|^{\frac{1}{2}}}{(2\pi)^p}\exp\left(-\frac{1}{2}\begin{bmatrix}\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1}\\ \underline{\xi}_k - \hat{\underline{\xi}}_{k|k}\end{bmatrix}^T \mathbf{Q}_k \begin{bmatrix}\underline{\xi}_{k+1} - \hat{\underline{\xi}}_{k+1|k+1}\\ \underline{\xi}_k - \hat{\underline{\xi}}_{k|k}\end{bmatrix}\right)$$

where $\mathbf{Q}_k = \mathbf{Q}(\hat{\underline{\xi}}_{k+1|k+1}, \hat{\underline{\xi}}_{k|k})$ is the $2p \times 2p$ observed FIM where, for $\underline{u}, \underline{v} \in \mathbb{R}^p$

$$\mathbf{Q}(\underline{u}, \underline{v}) = -\left[\begin{array}{cc} \nabla^2_{\underline{u}} & \nabla_{\underline{u}}[\nabla_{\underline{v}}]^T \\ \nabla_{\underline{v}}[\nabla_{\underline{u}}]^T & \nabla^2_{\underline{v}} \end{array}\right] \ln f_{\underline{\xi}_{k+1}, \underline{\xi}_k | \mathcal{Y}_k}(\underline{u}, \underline{v} | \mathcal{Y}_k).$$

Find a set of EKF equations by using this approximation and compare to your solution in part (b) of Exercise 6.18. (You can assume that the state is 1 dimensional if you like).

`Soln:`

`Let` ξ_k `be scalar and define` q^k_{ij} `the entries of the matrix` $\mathbf{Q}_k$, $i, j = 0, 1$. `A similar completion of the square argument to used above yields the following`

$$\begin{aligned} (\xi_{k+1} - \hat{\xi}_{k+1|k+1})^2 |\mathbf{Q}_{k+1}| / q^{k+1}_{00} &= (\xi_{k+1} - \hat{\xi}_{k+1|k})^2 |\mathbf{Q}_k| / q^k_{11} \\ &\quad + (y_{k+1} - c\xi_{k+1})^2 / \sigma^2_v - (y_{k+1} - c\hat{\xi}_{k+1|k})^2 / \sigma^2 \end{aligned} \qquad (27)$$

`where` $|\mathbf{Q}|$ `denotes determinant of` $\mathbf{Q}$ `and` $\sigma^2 = \mathrm{var}(y_{k+1}|\mathcal{Y}_k) = c^2 |\mathbf{Q}_{k+1}| / q^{k+1}_{00} + \sigma^2_v$. `By differentiating this relation wrt` ξ_{k+1} `and evaluating at` $\xi_{k+1} = \hat{\xi}_{k+1|k}$ `we obtain the state update`

$$\hat{\xi}_{k+1|k+1} = \hat{\xi}_{k+1|k} + \frac{c q^{k+1}_{00}}{\sigma^2_v |\mathbf{Q}_{k+1}|}(y_{k+1} - c\hat{\xi}_{k+1|k})$$

`This state update can be transformed to a true recursion by invoking the large sample approximation (invariance of MAP):` $\hat{\xi}_{k+1|k} = E[g(\xi_k)|\mathcal{Y}_k] \approx g(\hat{\xi}_{k|k})$.

`By twice differentiating relation (27) wrt` ξ_{k+1} `we obtain the recursion`

$$\frac{|\mathbf{Q}_{k+1}|}{q^{k+1}_{00}} = \frac{|\mathbf{Q}_k|}{q^k_{11}} + c^2/\sigma^2_v$$

`A simple calculation establishes that`

$$\begin{aligned} q^k_{00} &= 1/(B^2 R_w) \\ q^k_{11} &= F_k + [(g'(\hat{\xi}_{k|k}))^2 - (\hat{\xi}_{k+1|k+1} - g(\hat{\xi}_{k|k}))g''(\hat{\xi}_{k|k})]/(B^2 R_w) \\ q^k_{01} &= -g'(\hat{\xi}_{k|k})/(B^2 R_w) \end{aligned}$$

`and`

$$|\mathbf{Q}| = q_{11} q_{00} - q^2_{01}.$$

`Using these relations we obtain a recursion for` F_k. `As` q^k_{11} `depends on` $\hat{\xi}_{k+1|k+1}$ `the state and FIM recursions are not explicit. There are two solutions to this. One is to search for a stationary point` $\hat{\xi}_{k+1|k+1}$ `that satisfies both recursions; this can be done by numerical root finding on the pair of equations`

$$\begin{aligned} \frac{|\mathbf{Q}_{k+1}|}{q^{k+1}_{00}} - \frac{|\mathbf{Q}_k|}{q^k_{11}} &= 0 + c^2/\sigma^2_v \\ \hat{\xi}_{k+1|k+1} - \hat{\xi}_{k+1|k} + \frac{c q^{k+1}_{00}}{\sigma^2_v |\mathbf{Q}_{k+1}|}(y_{k+1} - c\hat{\xi}_{k+1|k}) &= 0 \end{aligned}$$

`Another approach is the ‘‘one step late’’ algorithm which substitutes` $|\mathbf{Q}_k|/q^k_{00}$ `for` $|\mathbf{Q}_{k+1}|/q^{k+1}_{00}$ `in the state update equation.`