NORGES TEKNISK- NATURVITENSKAPELIGE UNIVERSITET
INSTITUTT FOR TEKNISK KYBERNETIKK

TTK 4115 Linear System Theory

Exam Fall 2011 – Solution Suggestion

Question 1 (16%)

Explain the following concepts:

(a) (5%) Similarity transformation.

Solution: The similarity transformation tranforms objects to similar objects. The similarity transformation is a linear change of coordinates where the original object's representation is expressed with respect to different bases. Consider $\mathbf{x} \in \mathbb{R}^{n\times 1}$. The representation of $\mathbf{x}$ with respect to the basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is $\bar{\mathbf{x}}$ and, with $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n]$, the similarity transform is

$$\mathbf{x} = \mathbf{Q}\bar{\mathbf{x}}\,.$$

The state equations

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}\\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}\end{aligned}$$

can be written with respect to the basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ as

$$\begin{aligned}\dot{\bar{\mathbf{x}}} &= \bar{\mathbf{A}}\bar{\mathbf{x}} + \bar{\mathbf{B}}\mathbf{u}\\ \mathbf{y} &= \bar{\mathbf{C}}\bar{\mathbf{x}} + \bar{\mathbf{D}}u\end{aligned}$$

where

$$\bar{\mathbf{A}} = \mathbf{Q}\mathbf{A}\mathbf{Q}^{-1}\,, \qquad \bar{\mathbf{B}} = \mathbf{Q}\mathbf{B}\,, \qquad \bar{\mathbf{C}} = \mathbf{C}\mathbf{Q}^{-1}\,, \qquad \bar{\mathbf{D}} = \mathbf{D}\,.$$

The two representations are similar, meaning that they have the same eigenvalues. The similarity transformation can be used to express a system's representation with respect to different bases which may result in a more feasible realization for analysis and control purposes. It is important to note that the transfer function of a system is invariant with respect to a similarity transformation.

(b) (6%) Kalman's canonical decomposition.

Solution: Every state-space equation can be transformed, by an equivalence transformation, into four subspace equations, namely a controllable and observable, a controllable and not observable, a not controllable and observable, and a neither controllable nor observable subsystem. The controllable and observable subsystem is zero-state equivalent to the original state-space equation and they have the same transfer function.

(c) (5%) Minimal realization.

Solution: A transfer function $g(s)$ is realizable if there exists a state-space equation that has $g(s)$ as its transfer function. $g(s)$ is realizable if and only if $g(s)$ is a proper rational function. Minimal realizations are the realizations with the smallest possible dimension. A realization $(\mathbf{A}, \mathbf{b}, \mathbf{c}, d)$ is a minimal realization if and only if

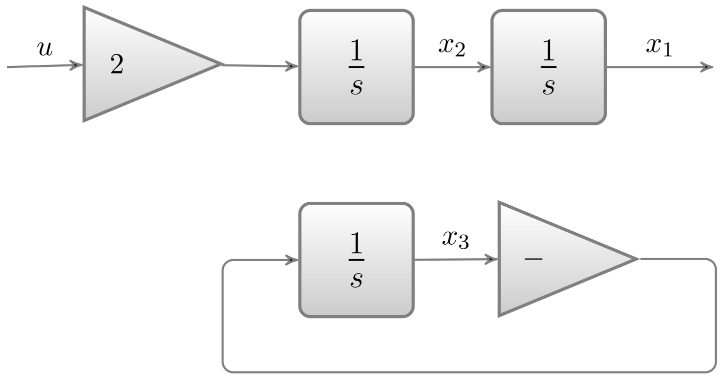
- $(\mathbf{A}, \mathbf{b})$ is controllable and $(\mathbf{A}, \mathbf{c})$ is observable, or
- $\dim \mathbf{A} = \deg g(s) = N(s)/D(s)$ where $N(s)$ and $D(s)$ are coprime factors (they have no common factors)

Question 2 (20%)
Consider the following system:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} u$$

(a) (3%) Draw a block diagram of the system.

Solution:



(b) (5%) Investigate if the system is internally stable (i.e. Lyapunov stable or marginally stable).

Solution: The characteristic polynomial is $\lambda^2(\lambda+1)$. The eigenvalues are 0, 0, and -1. Since is a double eigenvalue at the origin, the system is not asymptotically stable. The system is in Jordan-form and thus the minimal polynomial can be found to be $\lambda^2(\lambda+1)$. The eigenvalue in the origin is not a simple root of the minimal polynomial and therefore, the system is neither marginally stable.

(c) (5%) Investigate if the system is controllable.

Solution: The controllability matrix is

$$\mathcal{C} = \begin{pmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{pmatrix} = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The rank of the controllability matrix is 2 which is not full rank; thus the system is not controllable.

(d) (7%) Explain why the system is stabilizable, and choose a state feedback that places the eigenvalues of the system at $\lambda_1 = -1$, $\lambda_2 = -2$, $\lambda_3 = -3$.

Solution: Every uncontrollable state equation can be transformed into a controllable and a not controllable subspace equation by Kalman's decomposition

$$\begin{pmatrix} \dot{\mathbf{x}}_c \\ \dot{\mathbf{x}}_{\bar{c}} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_c & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{\bar{c}} \end{pmatrix} \begin{pmatrix} \mathbf{x}_c \\ \mathbf{x}_{\bar{c}} \end{pmatrix} + \begin{pmatrix} \mathbf{b}_c \\ \mathbf{0} \end{pmatrix} u$$

where $(\mathbf{A}_c, \mathbf{b}_c)$ is controllable. If furthermore $\mathbf{A}_{\bar{c}}$ is stable, the system is stabilizable. Here, the system is already in the form of Kalman's decomposition with

$$\mathbf{x}_c = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ \mathbf{x}_{\bar{c}} = x_3 , \ \mathbf{A}_c = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ \mathbf{A}_{12} = \mathbf{0}_{2\times 1} , \ \mathbf{A}_{\bar{c}} = -1 , \ \mathbf{b}_c = \begin{pmatrix} 0 \\ 2 \end{pmatrix} .$$

It is easy to verify that $(\mathbf{A}_c, \mathbf{b}_c)$ is controllable and that $\mathbf{A}_c$ is stable. Thus, the system is stabilizable. Choosing the feedback control

$$u = -\begin{pmatrix} \mathbf{k}_c & \mathbf{k}_{\bar{c}} \end{pmatrix} \begin{pmatrix} \mathbf{x}_c \\ \mathbf{x}_{\bar{c}} \end{pmatrix}$$

the closed-loop system becomes

$$\begin{pmatrix} \dot{\mathbf{x}}_c \\ \dot{\mathbf{x}}_{\bar{c}} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_c - \mathbf{b}_c\mathbf{k}_c & \mathbf{A}_{12} - \mathbf{b}_c\mathbf{k}_{\bar{c}} \\ \mathbf{0} & \mathbf{A}_{\bar{c}} \end{pmatrix} \begin{pmatrix} \mathbf{x}_c \\ \mathbf{x}_{\bar{c}} \end{pmatrix}$$

The eigenvalues of the system are given by the union of the eigenvalues of $(\mathbf{A}_c - \mathbf{b}_c\mathbf{k}_c)$ and $\mathbf{A}_{\bar{c}}$ (block triangular matrix). With $\mathbf{k}_c = (k_1 \, , \, k_2)$ the characteristic polynomial is

$$\lambda^3 + (1 + 2k_2)\,\lambda^2 + (2k_1 + 2k_2)\,\lambda + 2k_1$$

Comparing the coefficients to

$$(\lambda - \lambda_1)\,(\lambda - \lambda_2)\,(\lambda - \lambda_3) = \lambda^3 + 6\lambda^2 + 11\lambda + 6$$

it follows that the feedback controller gains are

$$\mathbf{k}_c = \begin{pmatrix} 3 & 2.5 \end{pmatrix} \qquad \mathbf{k}_{\bar{c}} \text{ arbitrary} .$$

Alternatively, note that the eigenvalue of the uncontrollable subsystem (x_3 dynamics) is at $\lambda_1 = -1$ and cannot be changed by feedback. The discussion may thus be limited to the controllable subsystem

$$\dot{\mathbf{x}}_c = \mathbf{A}_c\mathbf{x}_c + \mathbf{b}_c u$$

and the task to find a feedback control to place its eigenvalues at λ_2 and λ_3. Choosing

$$u = -\mathbf{k}_c\mathbf{x}_c$$

results in the closed-loop system

$$\dot{\mathbf{x}}_c = (\mathbf{A}_c - \mathbf{b}_c\mathbf{k}_c)\,\mathbf{x}_c .$$

With $\mathbf{k}_c = (k_1 \, , \, k_2)$ the characteristic polynomial is

$$\lambda^2 + 2k_2\lambda + 2k_1$$

Comparing the coefficients to

$$(\lambda - \lambda_2)\,(\lambda - \lambda_3) = \lambda^2 + 5\lambda + 6$$

it follows that the feedback controller gains are

$$\mathbf{k}_c = \begin{pmatrix} 3 & 2.5 \end{pmatrix} .$$

Question 3 (36%)

Consider the following system:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.4 \end{pmatrix} w \tag{1}$$

$$y = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + v \tag{2}$$

where w is an unknown disturbance and v is unknown measurement noise.

(a) (4%) Show that the system is observable.

Solution: The observability matrix is

$$\mathcal{O} = \begin{pmatrix} \mathbf{C} \\ \mathbf{CA} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} .$$

$\mathcal{O}$ has full rank and thus the system is observable.

(b) (6%) Show that the gain matrix of a state estimator (observer) with poles at $\lambda_1 = -10$ and $\lambda_2 = -12$ is given by:

$$\mathbf{L} = \begin{pmatrix} 19 \\ 99 \end{pmatrix}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are implicitly given by (1)–(2) above, and the state estimator is given by:

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{L}\left(y - \mathbf{C}\hat{\mathbf{x}}\right) + \mathbf{B}w .$$

Solution: Rewrite the observer as

$$\dot{\hat{\mathbf{x}}} = \left(\mathbf{A} - \mathbf{LC}\right)\hat{\mathbf{x}} + \mathbf{L}y .$$

With $\mathbf{L} = (l_1 \, , \, l_2)$ the characteristic polynomial of $(\mathbf{A} - \mathbf{LC})$ is

$$\det \begin{pmatrix} \lambda + 2 + l_1 & -1 \\ l_2 & \lambda + 1 \end{pmatrix} = \lambda^2 + (l_1 + 3)\,\lambda + l_1 + l_2 + 2 .$$

Comparing the coefficients to

$$(\lambda - \lambda_1)\,(\lambda - \lambda_2) = \lambda^2 + 22\lambda + 120$$

it follows that the gain matrix of the state observer is $\mathbf{L} = (19 \, , \, 99)^\top$.

(c) (9%) Show that the effect of the measurement noise v on the state estimates $\hat{x}_1$ and $\hat{x}_2$ can be described by the following transfer functions that are also given by the curves at the next page:

$$\frac{\hat{x}_1}{v}(s) = 19\frac{s + 6.21}{(s+10)\,(s+12)} = 0.98\frac{1 + 0.16s}{(1 + 0.1s)\,(1 + 0.08s)}$$

$$\frac{\hat{x}_2}{v}(s) = 99\frac{s + 2}{(s+10)\,(s+12)} = 1.65\frac{1 + 0.5s}{(1 + 0.1s)\,(1 + 0.08s)}$$

Solution: The state observer is

$$\dot{\hat{\mathbf{x}}} = (\mathbf{A} - \mathbf{LC})\,\hat{\mathbf{x}} + \mathbf{L}\,(\mathbf{Cx} + v)\ .$$

Taking the Laplace transform results in

$$s\hat{\mathbf{x}}\,(s) = (\mathbf{A} - \mathbf{LC})\,\hat{\mathbf{x}}\,(s) + \mathbf{L}\,(\mathbf{Cx}\,(s) + v\,(s))\ .$$

The transfer matrix from the measurement noise to the state is

$$\begin{aligned}
\frac{\hat{\mathbf{x}}}{v}\,(s) &= (s\mathbf{I} - \mathbf{A} + \mathbf{LC})^{-1}\,\mathbf{L} \\
&= \begin{pmatrix} s+21 & -1 \\ 99 & s+1 \end{pmatrix}^{-1} \begin{pmatrix} 19 \\ 99 \end{pmatrix} \\
&= \frac{1}{(s+21)\,(s+1)+99} \begin{pmatrix} s+1 & 1 \\ -99 & s+21 \end{pmatrix} \begin{pmatrix} 19 \\ 99 \end{pmatrix} \\
&= \frac{1}{(s+10)\,(s+12)} \begin{pmatrix} 19\,(s+6.21) \\ 99\,(s+2) \end{pmatrix}\ .
\end{aligned}$$

And consequently:

$$\begin{aligned}
\frac{\hat{x}_1}{v}\,(s) &= 19\frac{s+6.21}{(s+10)\,(s+12)} \\
&= 19\frac{6.21\left(\frac{1}{6.21}s+1\right)}{10\left(\frac{1}{10}s+1\right)12\left(\frac{1}{12}s+1\right)} = 0.98\frac{1+0.16s}{(1+0.1s)\,(1+0.08s)} \\
\frac{\hat{x}_2}{v}\,(s) &= 99\frac{s+2}{(s+10)\,(s+12)} \\
&= 99\frac{2\left(\frac{1}{2}s+1\right)}{10\left(\frac{1}{10}s+1\right)12\left(\frac{1}{12}s+1\right)} = 1.65\frac{1+0.5s}{(1+0.1s)\,(1+0.08s)}\ .
\end{aligned}$$

(d) (4%) Assume that the measurement noise $v\,(t)$ in some cases contains a dominant component $v\,(t) = 0.2\sin(200t)$. Estimate or calculate (based on the formulas or figures in part c)) the factor this noise component is reduced by in the estimate of x_1 and x_2.

Solution: With $\omega = 200$ the gains of the transfer functions can be computed to

$$\begin{aligned}
\left|\frac{\hat{x}_1}{v}\,(\jmath\omega)\right| &= \sqrt{\frac{(19\jmath\omega+118)\,(-19\jmath\omega+118)}{(\jmath\omega+10)\,(-\jmath\omega+10)\,(\jmath\omega+12)\,(-\jmath\omega+12)}} \\
&= \sqrt{\frac{19^2\omega^2+118^2}{(\omega^2+10^2)\,(\omega^2+12^2)}} \approx 0.09476 = -20.47\,\mathrm{dB} \\
\left|\frac{\hat{x}_2}{v}\,(\jmath\omega)\right| &= \sqrt{\frac{(99\jmath\omega+198)\,(-99\jmath\omega+198)}{(\jmath\omega+10)\,(-\jmath\omega+10)\,(\jmath\omega+12)\,(-\jmath\omega+12)}} \\
&= \sqrt{\frac{99^2\omega^2+198^2}{(\omega^2+10^2)\,(\omega^2+12^2)}} \approx 0.4935 = -6.134\,\mathrm{dB}\,.
\end{aligned}$$

Alternatively, the gains can be read directly from the Bode plots.

(e) (5%) Assume the system has a constant unknown disturbance w. Discuss (calculations not needed) the effect of this disturbance on the accuracy of the estimates $\hat{x}_1$ and $\hat{x}_2$.

Solution: Define the estimation error $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$. Then the error dynamics is

$$\begin{aligned}\dot{\mathbf{e}} &= \dot{\mathbf{x}} - \dot{\hat{\mathbf{x}}} = \mathbf{A}\mathbf{x} + \mathbf{B}w - \mathbf{A}\hat{\mathbf{x}} - \mathbf{L}\left(y - \mathbf{C}\hat{\mathbf{x}}\right) \\ &= \left(\mathbf{A} - \mathbf{L}\mathbf{C}\right)\mathbf{e} + \mathbf{B}w - \mathbf{L}v\,.\end{aligned}$$

Taking the Laplace transform and disregarding the measurement noise v yields

$$\mathbf{e}\left(s\right) = \left(s\mathbf{I} - \mathbf{A} + \mathbf{L}\mathbf{C}\right)^{-1}\mathbf{B}\frac{\bar{w}}{s}\,.$$

where $w\left(s\right) = \frac{\bar{w}}{s}$ since w is a constant disturbance. Applying the final value theorem results in

$$\lim_{t\to\infty} e\left(t\right) = \lim_{s\to 0} se\left(s\right) = \left(-\mathbf{A} + \mathbf{L}\mathbf{C}\right)^{-1}\mathbf{B}\bar{w}\,.$$

Consequently, the constant disturbance w results in a constant bias of the estimates $\hat{x}_1$ and $\hat{x}_2$.
Note that calculations are not needed for this problem as long as the reasoning is sufficient and leads to the correct conclusion.

(f) (3%) Suggest (and justify) how the above estimator can be initialized.

Solution: The error dynamics are globally exponentially stable (linear system, all eigenvalues have negative real parts). So the initialization will not affect the overall stability of the observer. However, from a performance perspective, an initialization close to the true state values will result in a better performance.

(g) (5%) Discuss (without calculations) the effect if the observer poles are chosen further to the left in the left half-plane. The discussion should include changes in performance and implementation due to discretization, measurement noise (assume $v\left(t\right)$ may have any power spectral density function) and constant disturbance w.

Solution: Choosing the observer poles further to the left in the left half-plane will result in a higher bandwidth of the observer. This lead to faster decaying error dynamics and thus better performance. However, a higher bandwidth also increases the sensitivity with respect to noise. As can be seen from the sub-problem (e), the constant bias of the estimates due to the constant disturbance is inversely proportional to the observer gains. Increasing the observer gains and consequently choosing the observer poles further to the left in the left half-plane will reduce the constant bias.

Question 4 (28%)

(a) (5%) Assume the following measurement series of the discrete-time measurement $y\left[k\right]$ is recorded under conditions when the state is constant, with sampling interval $T_s =$

0.04. Explain why it is reasonable to assume $v[k]$ is a (discrete) white noise sequence, and suggest an estimate for the autocorrelation function and variance of $v[k]$.

Solution: A white noise sequence is defined as sequence of zero-mean, uncorrelated random variables. Since the measurement sequence is recorded when the state is constant, it is reasonable to conclude that the mean of the noise sequence is zero and the variance equal to the measurement series' variance.

The mean and the standard deviation of the measurement series can be roughly estimated as 0.7 and 0.1, respectively. Consequently, the variance of the noise sequence is $\mathrm{Var}(V) = \sigma_V^2 = 0.01$. The autocorrelation function for a white noise sequence is

$$R_V[m] = A\delta[m]$$

where $\delta[m]$ is the Kronecker delta. The following holds

$$R_V[0] = A = E\left(V^2\right) .$$

The last equality is due to $v[k]$ being a stationary process. The mean-square value of the process $v[k]$ can be computed to

$$\mathrm{Var}(V) = E\left(V^2\right) - \left(E(V)\right)^2 .$$

With $\mathrm{Var}(V) = 0.01$ and $E(V) = 0$ it follows that $A = E\left(V^2\right) = 0.01$.

(b) (3%) Could you conclude whether the above measurement series is the output of a stationary random process or not?

Solution: A stationary random process is characterized by the fact that the statistical properties of the process do not depend on time. As a consequence, for a stationary random process, the parameters like mean value and variance do not change with time. From the measurement series shown in the figure it is reasonable to assume that the mean and the variance do not change in time, at least for the time window shown. That leads to the conclusion that the process is most probably stationary. Note however that generally, a infinite time series of the measurement would be needed to determine conclusively whether a random process is stationary or not.

(c) (10%) Consider the following system:

$$\begin{pmatrix} x_1[k+1] \\ x_2[k+1] \end{pmatrix} = \begin{pmatrix} 0.8 & -0.2 \\ 0 & 0.7 \end{pmatrix} \begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix} + \begin{pmatrix} 0.2 \\ 0.4 \end{pmatrix} w[k]$$

$$y = \begin{pmatrix} 1 & 0.1 \end{pmatrix} \begin{pmatrix} x_1[k] \\ x_2[k] \end{pmatrix} + v[k]$$

where $w[k]$ is a disturbance and $v[k]$ is measurement noise. Describe how a discrete-time Kalman-filter algorithm can be used to estimate the system states if $w[k]$ is a slowly time-varying disturbance (Hint: How to model such a disturbance?). Include the augmented model and the main formulas.

Solution: Since $w[k]$ is a slowly time-varying disturbance, the system has to be augmented with one additional state. The disturbance can be modelled as a Gaussian random-walk process, that is

$$w[k+1] = w[k] + \bar{w}[k]$$

where $\bar{w}[k]$ is a white noise sequence. Including the disturbance $w[k]$ as an additional state $x_3[k]$, the the augmented model is

$$\begin{pmatrix} x_1[k+1] \\ x_2[k+1] \\ x_3[k+1] \end{pmatrix} = \overbrace{\begin{pmatrix} 0.8 & -0.2 & 0.2 \\ 0 & 0.7 & 0.4 \\ 0 & 0 & 1 \end{pmatrix}}^{\Phi_k} \begin{pmatrix} x_1[k] \\ x_2[k] \\ x_3[k] \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \bar{w}[k]$$

$$y = \underbrace{\begin{pmatrix} 1 & 0.1 & 0 \end{pmatrix}}_{\mathbf{H}_k} \begin{pmatrix} x_1[k] \\ x_2[k] \\ x_3[k] \end{pmatrix} + v[k] \,.$$

Note that in case it cannot be assumed that the measurement noise is a white noise sequence, the model has to be augmented with the correlated part of the measurement noise as well.

The main formulas for the Kalman-filter are:

- Compute the Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^\top \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^\top + R_k \right)^{-1}$$

- Update estimate with measurement y_k:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(y - \mathbf{H}_k \hat{\mathbf{x}}_k^- \right)$$

- Compute error covariance for updated estimate:

$$\mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right) \mathbf{P}_k^- \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right)^\top + \mathbf{K}_k R_k \mathbf{K}_k^\top$$

- Project ahead:

$$\begin{aligned} \hat{\mathbf{x}}_{k+1} &= \Phi_k \hat{\mathbf{x}}_k \\ \mathbf{P}_{k+1}^- &= \Phi_k \mathbf{P}_k \Phi_k^\top + \mathbf{Q}_k \end{aligned}$$

with

$$\begin{aligned} \mathbf{Q}_k &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & q_k \end{pmatrix} \\ E\left[\bar{w}_k \bar{w}_i^\top \right] &= \begin{cases} q_k \,, & i = k \\ 0 \,, & i \neq k \end{cases} \\ E\left[v_k v_i^\top \right] &= \begin{cases} R_k \,, & i = k \\ 0 \,, & i \neq k \end{cases} \\ E\left[\bar{w}_k v_i^\top \right] &= 0 \quad \forall k, \, i \,. \end{aligned}$$

(d) (10%) Describe the main idea behind an extended Kalman-filter for estimating the state of a nonlinear system, based on the discrete-time ordinary Kalman-filter. Suggest means for numerically robust implementation and initialization of an extended Kalman-filter to ensure convergence and stability.

Solution: The main idea behind an extended Kalman-filter (EKF) is to linearize the nonlinear system about the current state estimate and then to consequently implement a linear Kalman-filter.

Means for a numerically robust implementation and initialization of the EKF are:

- Use double precision floating point computations.
- Ensure symmetric $\mathbf{P}_k$
  - Use a symmetric update formula.
  - Make $\mathbf{P}_k$ symmetric at each sample.
  - Use symmetric factorization methods (UD-factorization or "square root factors").
- Choose $\mathbf{Q}$ positive definite rather than positive semidefinite to avoid singular or negative definite $\mathbf{P}_k$.
- Divide the prediction step in several "integration steps".
- Good initialization of $\mathbf{P}_0^-$ is important to avoid numerical round-off errors.
- Good initialization of $\mathbf{x}_0^-$ is essential to avoid large linearization errors initially.