Mathematics of systems theory
Homework 3

Ville Wassberg
Collaborator: Klara Zimmermann

December 20, 2023

1

We can always design a feedback $u = kx$ such that the closed loop poles of the system

$$\begin{cases} \dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & a \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u, \\ y = \begin{pmatrix} 1 & 1 \end{pmatrix} x \end{cases} \tag{1}$$

can be placed at $\{-1, -1\}$. To see this we could check that the system always is controllable but considering that observability is to be checked we can instead check the eigenvalues of the matrix $A + bk$ directly:

$$\det(A + bk - \lambda \mathbf{I}_2) = \begin{vmatrix} -\lambda & 1 \\ k_1 & a + k_2 - \lambda \end{vmatrix} = -\lambda(a+k_2-\lambda)-k_1 = \lambda^2-(a+k_2)\lambda-k_1 \tag{2}$$

$$\Leftrightarrow \lambda = \frac{a + k_2}{2} \pm \sqrt{\frac{(a + k_2)^2}{4} + k_1}.$$

Thus, if we set the vector k such that

$$k = \begin{pmatrix} -1 & -2 - a \end{pmatrix} \tag{3}$$

then we get a double root of $\lambda = -1$, therefore, we can always make a pole placement at $\{-1, -1\}$.

The new system is not observable since the observability matrix becomes:

$$\Omega = \begin{pmatrix} C \\ C(A + bk) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \tag{4}$$

with rank 1 which is not full.

Now, assuming that the state is not available (not observable) we choose a such that the initial system is not observable. Since the observability matrix is

$$\Omega = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1+a \end{pmatrix}, \tag{5}$$

the system is not observable if and only if $a = -1$. To create an observer such that the estimation error converges to 0 faster than e^{-t} we can construct an estimator $\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x})$, set the error estimate as $\tilde{x} = x - \hat{x}$, for which the derivative becomes the following differential equation $\dot{\tilde{x}} = (A - LC)\tilde{x}$. Since we want the error estimation to fulfill $||x(t) - \hat{x}(t)|| < e^{-t}$, which is the norm of the solution to the differential equation, hence we need the eigenvalues of the matrix $(A - LC)$ to be less than -1 for the error estimation to converge faster than e^{-t}. Thus, we choose L accordingly. Calculating the eigenvalues of the system when $a = -1$ we get:

$$\det(A - LC - \lambda\mathbf{I}) = \lambda^2 + (l_2 + l_1 + 1)\lambda + (l_1 + l_2)$$

with solution

$$\lambda = -\frac{l_1 + l_2 + 1}{2} \pm \sqrt{\frac{(l_1 + l_2 + 1)^2}{4} - (l_1 + l_2)}, \tag{6}$$

where l_1 and l_2 are the components of L. If we choose

$$L = -k^T = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{7}$$

we see that $\lambda_1 = -2$, $\lambda_2 = -1$, satisfying our needs, thus choosing this observer, L, we get a new estimation of the system with an estimation error converging faster than e^{-t}. Our new enlarged system then gets defined by

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}). \tag{8}$$

2

In order to show that the solution $P(t)$ to the system

$$\min_u \int_0^{t_1} (y^2 + u^2)\, dt \tag{9}$$

subject to

$$\dot{x} = Ax + bu, \quad y = cx \tag{10}$$

where

$$A = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad c = \begin{pmatrix} 1 & 1 \end{pmatrix}, \tag{11}$$

is positive definite it is enough to show that the system is minimal, i.e. observable and reachable since that implies the system to have a symmetric positive definite (SPD) solution $P(t)$. We also get

$$Q = C^T C = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad R = \mathbf{I}, \quad y^2 = (cx)^T cx = x^T c^T cx = x^T Q x, \tag{12}$$

thus the system is written as:

$$\min_u \int_0^{t_1} (x^T Q x + u^T R u)\, dt. \tag{13}$$

We get the observability matrix

$$\Omega = \begin{pmatrix} c \\ cA \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix} \tag{14}$$

and the conrollability matrix

$$\Gamma = \begin{pmatrix} b & Ab \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}. \tag{15}$$

Since both the observability matrix and the reachability matrix has full rank the system is a minimal realisation and the solution $P(t)$ is thus SPD.

If we now let $t_1 \to \infty$ and attempt to show the existence of optimal control by solving the algebraic Riccati equation (ARE) and also verifying that P is SPD. Plugging all values into the ARE we get:

$$A^T + PA - PbR^{-1}b^T P + c^T c = 0$$

$$\Leftrightarrow$$

$$\begin{cases} p_{21}(p_{11} - p_{12}) - p_{11}(p_{11} - p_{12}) - 4p_{11} + 1 = 0, \\ p_{22}(p_{11} - p_{12}) - p_{12}(p_{11} - p_{12}) - 2p_{12} + 1 = 0 \\ p_{21}(p_{21} - p_{22}) - p_{11}(p_{21} - p_{22}) - 2p_{21} + 1 = 0, \\ p_{22}(p_{21} - p_{22}) - p_{12}(p_{21} - p_{22}) + 1 = 0. \end{cases} \tag{16}$$

Solving this system we get three alternatives for P, namely:

$$P_1 = \begin{pmatrix} 3\sqrt{2}-4 & \sqrt{2}-1 \\ \sqrt{2}-1 & \sqrt{2} \end{pmatrix}, P_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, P_3 = \begin{pmatrix} -3\sqrt{2}-4 & -\sqrt{2}-1 \\ -\sqrt{2}-1 & -\sqrt{2} \end{pmatrix}.$$

Since a two-by-two matrix is positive definite if both the trace and the determinant is strictly positive we see that P_1 is the only one satisfying this. Thus, P_1 is the optimal solution to the ARE.

3

To show that the block matrix

$$H = \begin{pmatrix} A & -BR^{-1}B^T \\ C^TC & -A^T \end{pmatrix} \tag{17}$$

has $-\lambda$ as it's eigenvalue if λ is an eigenvalue of H, we can show that H is similar to $-H^T$ which would then have the same eigenvalues. Since $-H^T$ admits the negative eigenvalues of H we then can conclude that $-\lambda$ is an eigenvalue of H if λ is. Now consider the matrix

$$J = \begin{pmatrix} 0 & \mathbf{I}_n \\ -\mathbf{I}_n & 0 \end{pmatrix}. \tag{18}$$

Then we have

$$J^{-1} = \begin{pmatrix} 0 & -\mathbf{I}_n \\ \mathbf{I}_n & 0 \end{pmatrix}, \tag{19}$$

where n is the size of A. Checking the matrix multiplication:

$$JHJ^{-1} = \begin{pmatrix} -A^T & C^TC \\ BR^{-1}B^T & A \end{pmatrix} = -H^T. \tag{20}$$

Thus, H is similar to $-H^T$ and by the reasoning above, $-\lambda$ is an eigenvalue of H if λ is an eigenvalue of H.

Assuming now that (C, A) is observable and (A, B) is reachable we can show that $\lambda = 0$ is not an eigenvalue of H. Suppose for a contradiction $\lambda = 0$ then there is an eigenvector v of H corresponding to λ such that $v \in \ker H$. Let $v = (v_1^T, v_2^T)^T$. Then we have:

$$\begin{cases} Av_1 - BR^{-1}B^Tv_2 = 0, \\ -C^TCv_1 - A^Tv_2 = 0. \end{cases} \tag{21}$$

Then we have

$$v_2 = (BR^{-1}B^T)^{-1}Av_1 \tag{22}$$

hence

$$-C^TCv_1 - A^T(BR^{-1}B^T)^{-1}Av_1 = 0. \tag{23}$$

Since C^TC is positive definite (not semi due to observability) and $A^T(BR^{-1}B^T)^{-1}A$ is positive definite due to R being positive definite and B and A non-zero due to reachability, which means it is invertible, too. However, this is a contradiction since (23) then has to be strictly negative. Thus, $\lambda = 0$ cannot be an eigenvalue of H.

Now assuming that

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \tag{24}$$

consists of n eigenvectors associated with the negative eigenvalues of H, hence

$$H \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} Z, \tag{25}$$

where Z is a stable matrix. To show that $P = X_2 X_1^{-1}$ is a solution to the ARE if X_1 is invertible we can plug it in to the equation and check. We have

$$H \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} -A^T & C^T C \\ BR^{-1}B^T & A \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} Z \tag{26}$$

and thus by inserting $-A^T X_2 = C^T C X_1 + X_2 Z$ and $BR^{-1}B^T X_2 = -AX_1 + X_1 Z$ we get

$$A^T X_2 X_1^{-1} + X_2 X_1^{-1} A - X_2 X_1^{-1} BR^{-1} B^T X_2 X_1^{-1} + C^T C = \tag{27}$$

$$-(C^T C X_1 + X_2 Z) X_1^{-1} + X_2 X_1^{-1} A + X_2 X_1^{-1}(-AX_1 + X_1 Z) X_1^{-1} + C^T C = 0$$

since all terms cancel.

4

To design a Kalman filter for the estimation of x, and express the covariance matrix $p(t) = E(x - \hat{x}(t))^2$ in terms of t, σ, p_0 we can use the known formulas for a discrete Kalman filter and the covariance matrix. The formulas are:

$$\begin{cases} x(t+1) = A(t)x(t) + B(t)v(t) \\ y(t) = C(t)x(t) + D(t)w(t), \end{cases}$$
$$\hat{x}(t+1) = [A - AK(t)C]\hat{x}(t) + AK(t)y(t)$$
$$K(t) = P(t)C^T[CP(t)C^T + DRD^T]^{-1}$$
$$P(t+1) = AP(t)A^T - AP(t)C^T[CP(t)C^T + DRD^T]^{-1}CP(t)A^T + BQB^T, \tag{28}$$

where $K(t)$ is the Kalman gain and $P(t)$ is the covariance matrix. Since it is a constant to be observed we have $A = 1$, $B = 0$. Since x is to be measured directly, $C = 1$ and the white noise variance is given as σ^2 thus $D = \sigma$ is sufficient, hence our Kalman gain becomes:

$$K(t) = \frac{p(t)}{p(t) + \sigma^2}. \tag{29}$$

Now expressing the Kalman filter we get

$$\hat{x}(t+1) = (1 - \frac{p(t)}{p(t) + \sigma^2})\hat{x}(t) + \frac{p(t)}{p(t) + \sigma^2}y(t). \tag{30}$$

The covariance matrix is

$$p(t+1) = p(t) - p(t))p(t) + \sigma^2)^{-1}p(t) = \frac{p(t)\sigma^2}{p(t) + \sigma^2} = \frac{1}{\sigma^{-2} + p(t)^{-1}}. \tag{31}$$

Applying induction we get:

$$p(1) = \frac{1}{\sigma^{-2} + p_0^{-1}},$$
$$p(2) = \frac{1}{\sigma^{-2} + p(1)^{-1}} = \frac{1}{\sigma^{-2} + \sigma^{-2} + p_0^{-1}} = \frac{1}{2\sigma^{-2} + p_0^{-1}}, \tag{32}$$

now suppose

$$p(k) = \frac{1}{k\sigma^{-2} + p_0^{-1}}$$

then

$$p(k+1) = \frac{1}{\sigma^{-2} + p(k)^{-1}} = \frac{1}{\sigma^{-2} + k\sigma^{-2} + p_0^{-1}} = \frac{1}{(k+1)\sigma^{-2} + p_0^{-1}}, \tag{33}$$

Thus we have an equation for the covariance matrix:

$$p(t) = \frac{1}{t\sigma^{-2} + p_0^{-1}}. \tag{34}$$

To show that $\hat{x}(t+1) = \hat{x}(t)$ as $t \to \infty$ we can inspect what happens to $p(t)$ which is quite direct, and one can see that $\lim_{t\to\infty} p(t) = 0$. Then by inspecting the equation we get

$$\begin{aligned}\lim_{t\to\infty} \hat{x}(t+1) &= \lim_{t\to\infty} (1 - \frac{p(t)}{p(t)+\sigma^2})\hat{x}(t) + \frac{p(t)}{p(t)+\sigma^2} y(t) \\ &= \lim_{t\to\infty} (1 - \frac{0}{0+\sigma^2})\hat{x}(t) + \frac{0}{0+\sigma^2} y(t) = \hat{x}(t).\end{aligned} \tag{35}$$

If we would let the variance go to infinity we can see that

$$\hat{x}(t+1) = \lim_{\sigma\to\infty} \hat{x}(t+1) = \lim_{t\to\infty} (1 - \frac{p(t)}{p(t)+\sigma^2})\hat{x}(t) + \frac{p(t)}{p(t)+\sigma^2} y(t) = \hat{x}(t),$$

too. This can be interpreted as no matter what time our measure is taken we will not get a better estimate than our initial guess.