ECE 486: State-space control

M.-A. Belabbas
Spring 2017, UIUC

We give a brief introduction to state-space control. In particular, we cover controllability, observability, output feedback pole placement, observer design and the separation principle.

1 Some facts from linear algebra

1.1. We start with a few facts from linear algebra. A **matrix** $A \in \mathbb{R}^{n\times n}$ is a $n \times n$ array of real numbers,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

The first index i in a_{ij} refers to the row of A where the entry is located, and the second index j to the column.

1.2. The *identity* matrix is the matrix with ones on the diagonal and zeros everywhere else:

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & & \cdots & 0 & 1 \end{bmatrix}$$

It acts just like 1 acts with the real or complex numbers: for any matrix A, we have $A = IA = AI$.

1.3. Recall that in general, the product of matrices does not commute! that is $AB \neq BA$ in general.

1.4. We can perform *block operations* on matrices: for example, if

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

is a 4 by 4 matrix partitioned into 2 by 2 blocks $A_1, \ldots, A_4$, and similarly B is partitioned as

$$B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

with 2 by 2 blocks $B_1, \ldots, B_4$, we can perform all operations on these matrices using the same rules as usual, but with the blocks as a basic unit. For example, the product AB is

$$AB = \begin{bmatrix} A_1B_1 + A_2B_3 & A_1B_2 + A_2B_4 \\ A_3B_1 + A_4B_3 & A_3B_2 + A_4B_4 \end{bmatrix}$$

Similarly, the determinant of A is

$$\det(A) = \det(A_1A_4 - A_2A_3).$$

1.5. A matrix with zero determinant is called *singular* or *non-invertible*. If A is a singular matrix, there exits a vector $v \neq 0$ such that $Av = 0$. We say that v is in the *kernel* of A.

> For example,
>
> $$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$
>
> is singular since $\det(A) = 2 - 2 = 0$. We can check that $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is such that $Av = 0$.

1.6. If A is upper triangular, or lower triangular, its determinant is the product of its diagonal entries.

> For example, if
>
> $$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 17 \\ 0 & 0 & 2 \end{bmatrix},$$
>
> then $\det(A) = 1(-1)2 = -2$.

1.7. If A is *block upper triangular* or *block lower triangular*, its determinant is the product of the determinant of the blocks.

> For example, for the 6 by 6 matrix
>
> $$A = \begin{bmatrix} A_1 & A_2 & A_3 \\ 0 & A_4 & A_5 \\ 0 & 0 & A_6 \end{bmatrix}$$

with 2 by 2 blocks (note that we write 0 here for a 2 by 2 matrix with entries all zero), we have

$$\det(A) = \det(A_1)\det(A_4)\det(A_6)$$

1.8. The eigenvalues $\lambda \in \mathbb{C}$ and eigenvectors $v \neq 0 \in \mathbb{C}^n$ of A are defined by the equation

$$Av = \lambda v.$$

1.9. Observe that we can rewrite the equation defining eigenvectors and eigenvalues as $Av - \lambda v = 0$, or

$$(A - \lambda I)v = 0.$$

This means that if λ is an eigenvalue of A, the matrix $A - \lambda I$ has the vector $v \neq 0$ in its kernel. Thus $A - \lambda I$ is a singular matrix and its determinant is zero. That is, if λ is an eigenvalue of A, $A - \lambda I$ is singular and thus

$$\det(A - \lambda I) = 0.$$

We thus have shown that the roots of the above equation are the eigenvalues of A. The polynomial $\det(A - \lambda I)$ is called the **characteristic polynomial** of A. It is a polynomial in λ of order n. We use λ or s in general as its argument.

The eigenvalues of

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix}$$

are the roots of the polynomial

$$\det(I\lambda - A) = \det\left(\begin{bmatrix} \lambda - 1 & 0 & -2 \\ -1 & \lambda - 1 & 0 \\ -2 & -2 & \lambda \end{bmatrix}\right) = \lambda^3 - 2\lambda^2 - 3\lambda = \lambda(\lambda^2 - 2\lambda - 3).$$

The eigenvalues are $3, -1$ and 0.

1.10. The transpose $A^\top$ of a matrix A is obtained by exchanging rows and columns.

For example

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 4 & 1 & 5 \end{bmatrix}^\top = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 1 & 1 \\ 3 & 2 & 5 \end{bmatrix}$$

1.11. The determinant of A and $A^\top$ are the same:

$$\det(A) = \det(A^\top).$$

We conclude from this that the characteristic polynomial of A and $A^\top$ are the same and thus A and $A^\top$ also have the same eigenvalues. This fact will be important in our derivations below!

2 Transfer functions and canonical state-space forms

2.1. Recall from the beginning of the course that after modelling a system, we can write it in the canonical form (we may need to *linearize the system*)

$$\begin{cases} \dot{x} &= Ax + Bu \\ y &= Cx \end{cases}$$

where $A \in \mathbb{R}^{n\times n}$ is a n by n matrix, B is a n-dimensional column vector (a n by 1 matrix) and C is a n-dimensional row vector (a 1 by n matrix).

2.2. The transfer function of the above system is obtained by taking the Laplace transform of the equations defining it. The first equation yields

$$sX(s) = AX(s) + BU(s)$$

where $X(s)$ is a vector containing the Laplace transforms of the entries $x_1(t), \dots, x_n(t)$ of the vector $x(t)$, and $U(s)$ is the Laplace transform of $u(t)$. The matrices A and B are unchanged. The second equation yields $Y(s) = CX(s)$. From the first equation, we get (we denote by I the n by n identity matrix)

$$\begin{aligned} sX(s) &= AX(s) + BU(s) \\ sX(s) - AX(s) &= BU(s) \\ (Is - A)X(s) &= BU(s) \\ X(s) &= (Is - A)^{-1}BU(s) \end{aligned}$$

Using the second equation, we obtain that

$$G(s) = Y(s)/U(s) = C(Is - A)^{-1}B. \tag{1}$$

2.3. Example: Let

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ and } C = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

Recall that we can easily compute the inverse of a 2 by 2 matrix as follows:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Hence, applying this formula we get

$$(Is - A)^{-1} = \begin{bmatrix} s-3 & -2 \\ -2 & s-1 \end{bmatrix}^{-1} = \frac{1}{s^2 - 4s - 1} \begin{bmatrix} 1-s & -2 \\ -2 & 3s \end{bmatrix}.$$

We then obtain the transfer function of the system described by A, B and C

$$G(s) = C(Is - A)^{-1}B = \begin{bmatrix} 0 & 1 \end{bmatrix} \frac{1}{s^2 - 3s - 1} \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{2s - 4}{s^2 - 4s - 1}$$

2.4. Using Equation (1), we can obtain the transfer function of a system described in state-space form. The other direction (that is, obtaining a triplet A, B, C from a transfer function) is called 'finding a realization of $G(s)$'. We will describe here two important realizations: the controllable canonical form (CCF) and observable canonical form (OCF). The justification of the names will become obvious once we look at controllability and observability (a bit below in these notes.)

2.5. Given $G(s)$, we first need to know what the dimensions of the matrices A, B and C are. In these notes, we only look at realizations for which the dimension of A is the *degree n of the denominator of* $G(s)$, B is a n-dimensional column vector and C is a n-dimensional row vector.

2.6. Assume given a generic proper transfer function $G(s)$ (proper means that the degree of the denominator is larger than the degree of the numerator), such that the numerator and denominator of $G(s)$ have no common factors (if you are given a $G(s)$ with common factors, e.g. a $s + 1$ term in the numerator and denominator, simplify the terms first)

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \cdots + b_{n-1} s + b_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n}.$$

2.7. We define the **CCF** to be the realization

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots & \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & & -a_2 & -a_1 \end{bmatrix}; \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} \text{ and } C = \begin{bmatrix} b_n & b_{n-1} & \cdots & b_2 & b_1 \end{bmatrix} \quad (2)$$

2.8. A direct computation of the determinant $\det(Is - A)$ for A defined as in (2) yields

$$\det(Is - A) = s^n + a_1 s^{n-1} + a_2 s^{n-2} \cdots + a_{n-1} s + a_n \quad (3)$$

This is an important fact. You can easily verify it by using the general formula to evaluate a determinant via minor expansions. We only verify it in the 2 by 2 case. We have

$$A = \begin{bmatrix} 0 & 1 \\ -a_2 & -a_1 \end{bmatrix}$$

and thus $\det(Is - A)$ is

$$\det(Is - A) = \left| \begin{bmatrix} s & -1 \\ a_2 & s + a_1 \end{bmatrix} \right| = s(s + a_1) - a_2(-1) = s^2 + a_1 s + a_2,$$

which matches (3).

2.9. We now verify that (2) is indeed a realization of $G(s)$. To do this, we need to check that $C(Is - A)^{-1}B$ yields $G(s)$. We will again only do this in the 2 by 2 case, but doing it in the general case is not any harder, except for the computation of the inverse of $Is - A$. (You don't need to know how to prove this fact in the general case).

We have

$$\begin{aligned} C(Is - A)^{-1}B = \begin{bmatrix} b_2 & b_1 \end{bmatrix} \frac{1}{s^2 + a_1 s + a_2} \begin{bmatrix} s + a_1 & 1 \\ -a_2 & s \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ = \frac{1}{s^2 + a_1 s + a_2} \begin{bmatrix} b_2 & b_1 \end{bmatrix} \begin{bmatrix} 1 \\ s \end{bmatrix} = \frac{b_1 s + b_2}{s^2 + a_1 s + a_2}, \end{aligned}$$

which is indeed $G(s)$.

2.10. If A, B, C is a realization of $G(s)$, we claim that

$$\bar{A} = A^\top, \bar{B} = C^\top, \bar{C} = B^\top$$

is also a realization of $G(s)$. We call this realization the **observable canonical form** or OCF. To verify that it is indeed a realization of $G(s)$, we need to check that

$$\bar{C}(Is - \bar{A})^{-1}\bar{B} = G(s).$$

Using the definitions of $\bar{A}, \bar{B}, \bar{C}$, the fact that $(XYZ)^\top = Z^\top Y^\top X^\top$, the fact that $(Is - A)^\top = I^\top s - A^\top = Is - A^\top$, and the fact that $(A^\top)^{-1} = (A^{-1})^\top$, we write

$$\begin{aligned} \bar{C}(Is - \bar{A})^{-1}\bar{B} &= B^\top (Is - A^\top)^{-1} C^\top \\ &= B^\top ((Is - A)^\top)^{-1} C^\top \\ &= B^\top ((Is - A)^{-1})^\top C^\top \\ &= (C(Is - A)^{-1}B)^\top \\ &= C(Is - A)^{-1}B \end{aligned}$$

where the transition to the last line is justified because $C(Is - A)^{-1}B$ is a rational function in s, i.e. a 1 by 1 matrix, and thus it is equal to its transpose.

2.11. There are many more realizations of $G(s)$ than the two we have introduced here. We can obtain them via *changes of coordinates.* Let T be a n by n invertible matrix and define

$$\bar{x} = Tx.$$

The above relation defines a **linear change of variables.** The new variables describing the system are $\bar{x}$, and because T is invertible, we can recover x from $\bar{x}$ via

$$x = T^{-1}\bar{x}.$$

2.12. If we know that our system obeys in the x coordinate the relations

$$\left\{ \begin{array}{lcl} \dot{x} & = & Ax + Bu \\ y & = & Cx \end{array} \right. ,$$

then what relations does it obey in the $\bar{x}$ coordinates? We can easily obtain them by substitution:

We have

$$\begin{aligned} \frac{d}{dt}\bar{x} &= \frac{d}{dt}(Tx) \\ &= T\frac{d}{dt}x \\ &= T(Ax + Bu) \\ &= TAx + TBu \\ &= TAT^{-1}\bar{x} + TBu \end{aligned}$$

and

$$y = Cx = CT^{-1}\bar{x}.$$

We conclude that if a system is described in the x coordinates by the matrices A, B, C, it is described in the $\bar{x} = Tx$ coordinates by the matrices TAT^{-1}, TB, CT^{-1} :

$$\begin{array}{lcl} x & \leftrightarrow & Tx \\ A & \leftrightarrow & TAT^{-1} \\ B & \leftrightarrow & TB \\ C & \leftrightarrow & CT^{-1} \end{array} \tag{4}$$

2.13. The matrices A and $\bar{A} = TAT^{-1}$ have the same characteristic polynomial. We verify by computing the characteristic polynomial of $\bar{A}$ and simplify it to the characteristic polynomial of A:

$$\begin{aligned} \det(Is - \bar{A}) &= \det(Is - TAT^{-1}) \\ &= \det(TIsT^{-1} - TAT^{-1}) \\ &= \det(T(Is - A)T^{-1}) \\ &= \det(T)\det(Is - A)\det(T^{-1}) \\ &= \det(T)\det(T^{-1})\det(Is - A) \\ &= \det(Is - A) \end{aligned}$$

We used the fact that $\det(T^{-1}) = (\det(T))^{-1}$. Note that the determinant of a matrix is a real number, and thus we can write that $\det(T^{-1})\det(Is - A) = \det(Is - A)\det(T^{-1})$. This does *not mean that* $(Is - A)T^{-1} = T^{-1}(Is - A)$. In general, the order in which we take the product of matrices is important! We conclude from this calculation that

The poles of a system do not move under changes of coordinates

2.14. We now compute the transfer function of the system in the new coordinates: let $\bar{A} = TAT^{-1}$, $\bar{B} = TB$ and $\bar{C} = CT^{-1}$ then

$$\begin{aligned}
\bar{C}(Is - \bar{A})^{-1}\bar{B} &= CT^{-1}(Is - TAT^{-1})^{-1}TB \\
&= CT^{-1}(sTT^{-1} - TAT^{-1})^{-1}TB \\
&= CT^{-1}(T(Is - A)T^{-1})^{-1}TB \\
&= CT^{-1}T(Is - A)^{-1}T^{-1}TB \\
&= C(Is - A)^{-1}B
\end{aligned}$$

we have thus shown that

The transfer function of a system does not change under changes of coordinates.

2.15. The above says that to any transfer function $G(s)$, one can associate an infinite number of state-space realizations: start for example with the CCF, then choose any invertible matrix T and change coordinates: this yields a new realization, i.e a state-space system whose transfer function is $G(s)$.

2.16. Example Consider the system of point 3, described by

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ and } C = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$

Assume someone models the same system, but instead of using x_1, he uses $x_1 + x_2$ and instead of using x_2, he uses $x_1 - 2x_2$. That is, he uses the variables

$$\bar{x}_1 = x_1 + x_2 \text{ and } \bar{x}_2 = x_1 - 2x_2.$$

In matrix form,

$$\bar{x} = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} x$$

where we recall that x and $\bar{x}$ are 2 dimensional column vector in this example. The determinant of T is non-zero and thus T is invertible. We have that

$$T^{-1} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}.$$

The system in $\bar{x}$ coordinates is thus

$$\bar{A} = TAT^{-1} = \begin{bmatrix} 13/3 & 2/3 \\ -2/3 & -1/3 \end{bmatrix}$$

We can also verify that the characteristic polynomials of A and $\bar{A}$ are the same:

$$\det(Is - \bar{A}) = \det(Is - A) = s^2 - 4s - 1 \tag{5}$$

Finally, we have that

$$\bar{B} = TB = \begin{bmatrix} 3 \\ -3 \end{bmatrix} \text{ and } \bar{C} = CT^{-1} = \begin{bmatrix} 1/3 & -1/3 \end{bmatrix}.$$

We can also verify that

$$C(Is - A)B = \bar{C}(Is - \bar{A})^{-1}\bar{C} = \frac{2s - 4}{s^2 - 4s - 1}.$$

3 Controllability and output feedback pole placement

3.1. We say that the system

$$\begin{cases} \dot{x} &= Ax + Bu \\ y &= Cx \end{cases}$$

is **controllable** if for any initial state x_0 and any final state x_T, there exists a control function $u(t)$, $t \in [0, T]$ that drives the system from $x(0) = x_0$ to $x(T) = x_T$. That is, if you solve the differential equation $\dot{x} = Ax + Bu$ from the initial condition x_0 and with the control $u(t)$ just specified, then after T seconds, we have $x(T) = x_T$. The justification of the name controllable should be clear: systems that have this property can be driven from any state to any other state, that is we have a complete control over the system through the term $Bu(t)$.

3.2. Without proof, we give the following condition: a system $\dot{x} = Ax + Bu$ is controllable if the matrix

$$C(A, B) = \begin{bmatrix} | & | & | & & | & | \\ B & AB & A^2B & \cdots & A^{n-2}B & A^{n-1}B \\ | & | & | & & | & | \end{bmatrix} \tag{6}$$

is *full rank*. That is if its determinant is non-zero. In Equation (6), each term B, A^2B, etc is a column vector. There are n such terms and thus $C(A, B)$ is a square matrix. Observe that controllability does not depend on C, which was to be expected from its definition.

3.3. Example: Consider the system described by the matrices

$$A = \begin{bmatrix} -15 & 8 \\ -15 & 7 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Can you drive the system from $x_0 = [1,\ 0]^\top$ to $x_1 = [2,\ 2]^\top$ in 1 second?
To answer the question, we check whether the system is controllable. We have

$$C(A, B) = \begin{bmatrix} 1 & -7 \\ 1 & -8 \end{bmatrix}$$

and $\det(C(A, B)) \neq 0$. We conclude that the system is controllable and thus the answer is yes. Notice that if the system were not controllable, we cannot conclude anything: there might still be a control law achieving the objective, the only fact we know is that we cannot drive the system from an *arbitrary* initial state to an *arbitrary* final state.

3.4. From its definition, it should be intuitively clear that controllability of a system does not depend on the coordinates used to describe it. Mathematically, that means that if the system

A, B is controllable, the system $\bar{A} = TAT^{-1}, \bar{B} = TB$ for any invertible matrix T should also be controllable.

We verify that it is the case: what we have to show is that if $C(A, B)$ is of full rank, then so is $C(\bar{A}, \bar{B})$. To do this, we evaluate $C(\bar{A}, \bar{B})$ explicitly

$$
\begin{aligned}
C(\bar{A}, \bar{B}) &= \begin{bmatrix} | & | & | & & | \\ \bar{B} & \bar{A}\bar{B} & \bar{A}^2\bar{B} & \cdots & \bar{A}^{n-1}\bar{B} \\ | & | & | & & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & | & & | \\ TB & TAT^{-1}TB & TAT^{-1}TAT^{-1}TB & \cdots & TAT^{-1}\cdots TAT^{-1}TB \\ | & | & | & & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & | & & | \\ TB & TAB & TAAB & \cdots & TA\cdots AB \\ | & | & | & & | \end{bmatrix} \\
&= \begin{bmatrix} | & | & | & & | \\ TB & TAB & TA^2B & \cdots & TA^{n-1}B \\ | & | & | & & | \end{bmatrix} \\
&= T\begin{bmatrix} | & | & | & & | \\ B & AB & A^2B & \cdots & A^{n-1}B \\ | & | & | & & | \end{bmatrix} \\
&= TC(A, B)
\end{aligned}
$$

We conclude that

$$\det(C(\bar{A}, \bar{B})) = \det(TC(A, B)) = \det(T)\det(C(A, B)).$$

From this last equation, we see that $\det(C(\bar{A}, \bar{B})) \neq 0$ if $\det(C(A, B))$ is controllable, since $\det(T) \neq 0$ by definition of T.

The formulas

$$C(\bar{A}, \bar{B}) = TC(A, B) \text{ and } T = C(\bar{A}, \bar{B})C(A, B)^{-1} \tag{7}$$

will be useful below.

3.5. We now consider feedback control of a system described in state-space form. We will look first at *state feedback control*, in which it is assumed that the controller has access to the *complete state of the system*, that is the vector x. We then look at *output feedback control*, is which it is assumed that we only have access to $y = Cx$, and because C is a row vector, $y \in \mathbb{R}$.

3.6. A **state feedback controller** is a controller of the form

$$u = -Kx$$

where

$$K = \begin{bmatrix} k_1 & k_2 & \cdots & k_n \end{bmatrix}.$$

We can also write it as

$$u = -k_1x_1 - k_2x_2 - \cdots - k_nx_n$$

3.7. We now show how to use state-feedback to place the poles of a *controllable system at an arbitrary desired location.*

3.8. If a system is controllable, how to find the change of coordinates that puts it in CCF? That is, if the pair A, B is controllable, how can we find T such that

$$\bar{A} = TAT^{-1} \text{ and } \bar{B} = TB$$

are such that $\bar{A}, \bar{B}$ is in CCF?

We present a *three steps procedure* to do so

1. **Evaluate the characteristic polynomial of A, and deduce $\bar{A}$ from it:** we can obtain $\bar{A}$ by putting the coefficients of the characteristic polynomial of A in the last row of $\bar{A}$ with the appropriate sign change and reordering:

   Justification: Recall that if $\bar{A} = TAT^{-1}$, for *any invertible* T, then the characteristic polynomial of $\bar{A}$ and A are the same. We *want* $\bar{A}$ to be in CCF, that is $\bar{A}$ is of the form (2). But we know from (3) that if $\bar{A}$ is in that form, the coefficients of its characteristic polynomial are in its last row. We conclude that we obtain $\bar{A}$ by putting the coefficients of the characteristic polynomial of A (which is the same as the one of $\bar{A}$ in the last row of $\bar{A}$.

Example:

Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The characteristic polynomial of A is

$$\det(Is - A) = (s-1)^2 - 2 = s^2 - 2s - 1.$$

Thus, $a_2 = -1$ and $a_1 = -2$ and

$$\bar{A} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}.$$

2. **Compute $C(\bar{A}, \bar{B})$ and $C(A, B)$.**

   Justification: Even though we don't know T, because we require $\bar{A}, \bar{B}$ to be in CCF, we know that $\bar{B} = [0,\ 1]^\top$. We can thus compute $C(\bar{A}, \bar{B})$.

Example: For the example above, we have

$$C(A, B) = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \text{ and } C(\bar{A}, \bar{B}) = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}.$$

3. **Compute $T = C(\bar{A}, \bar{B})C(A,B)^{-1}$.**

   Justification: We have shown in 4 how $C(A,B)$ and $C(\bar{A},\bar{B})$ were related. From there, the relation above is straightforward to derive.

> **Example:** From the same example as above, we find
>
> $$T = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

We can now verify our computations by checking that indeed, $\bar{A} = TAT^{-1}$ and $\bar{B} = TB$.

Output feedback pole placement

We now give a procedure to perform output-feedback pole placement. The type of questions we want to answer is the following: Given the system

$$\dot{x} = Ax + Bu,$$

find a feedback control law

$$u = -Kx$$

such that the poles of the *closed-loop* system are are prescribed positions. Recall that the closed-loop poles are the roots of the closed-loop characteristic polynomial. Because the closed-loop system is

$$\dot{x} = Ax - BKx = (A - BK)x, \tag{8}$$

the closed-loop characteristic polynomial is $\det(Is - A + BK)$.

We start with an example before giving a general procedure.

3.9. Consider the system from 3: it is described by the matrices

$$A = \begin{bmatrix} -15 & 8 \\ -15 & 7 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Assume that we want the closed loop system to have poles at $-10 \pm j$.

Step 1: Find the desired closed-loop characteristic polynomial.

> It is in this case given by
>
> $$(s + 10 + j)(s + 10 - j) = s^2 + 20s + 101.$$

In case you are given less desired poles than the order of the system, you can always set the non-specified poles at arbitrary locations *to the left* of the desired poles, so that the dominant poles are the desired ones.

Step 2: Find the change of variables that puts the system in CCF To do so, we use the procedure of 8. In this case, we get:

(a) Characteristic polynomial of A: $s^2 + 8s + 15$.

(b) CCF of A: $\bar{A} = \begin{bmatrix} 0 & 1 \\ -15 & -8 \end{bmatrix}$ and $\bar{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

(c) Controllability matrices in the original and CCF form to evaluate T:

$$C(A, B) = \begin{bmatrix} 1 & -7 \\ 1 & -8 \end{bmatrix}, C(\bar{A}, \bar{B}) = \begin{bmatrix} 0 & 1 \\ 1 & -8 \end{bmatrix}.$$

Thus we get $T = C(\bar{A}, \bar{B})C(A, B)^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

Step 3: Find $\bar{K}$ so that the poles of the closed-loop system in CCF are the desired poles.

This is easy to do once we notice that in CCF, because

$$\bar{B}\bar{K} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} \bar{k}_1 & \bar{k}_2 & \cdots & \bar{k}_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ \bar{k}_1 & \bar{k}_2 & \cdots & \bar{k}_{n-1} & \bar{k}_n \end{bmatrix}.$$

Thus in CCF

$$\bar{A} - \bar{B}\bar{K} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots & \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & & -a_2 & -a_1 \end{bmatrix} - \bar{B}\bar{K}$$
$$= \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots & \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & & 0 & 1 \\ -a_n - \bar{k}_1 & -a_{n-1} - \bar{k}_2 & \cdots & & -a_2 - \bar{k}_{n-1} & -a_1 - \bar{k}_n \end{bmatrix}.$$

From (3), we conclude that the characteristic polynomial of $\bar{A} - \bar{B}\bar{K}$, in CCF, is

$$\det(Is - \bar{A} - \bar{B}\bar{K}) = s^n + (a_1 + k_n)s^{n-1} + (a_2 + k_{n-2})s^{n-1} + \cdots + (+a_n + k_1).$$

It suffices now to choose k_i so that the coefficients of the above polynomial and the coefficients of the desired characteristic polynomial are the same.

> In the current example, we have
>
> $$\bar{A} - \bar{B}\bar{K} = \begin{bmatrix} 0 & 1 \\ -15 - \bar{k}_1 & -8 - \bar{k}_2 \end{bmatrix}.$$
>
> We find the relations for $\bar{K}$:
>
> $$-15 - \bar{k}_1 = -101 \text{ and } -8 - \bar{k}_2 = -20.$$
>
> We conclude that
>
> $$\bar{K} = \begin{bmatrix} 86 & 12 \end{bmatrix}.$$

Step 4 Return to the original coordinates to obtain K.

In $\bar{x}$ coordinates, the closed-loop system

$$\frac{d}{dt}\bar{x} = A\bar{x} - \bar{B}\bar{K}\bar{x}$$

has poles at the desired locations. We know that changing coordinates does *not change the locations of the poles*. This should be intuitively clear (the behavior of the system, which is described in parts by its poles, should not depend on the coordinates chosen to describe it) and is justified by point 13.

To return to the original coordinates, we recall that $x = T^{-1}\bar{x}$. Thus

$$\begin{aligned} \dot{x} &= T^{-1}\frac{d}{dt}\bar{x} \\ &= T^{-1}(\bar{A}\bar{x} - \bar{B}\bar{K}\bar{x}) \\ &= T^{-1}(TAT^{-1}Tx - TB\bar{K}Tx) \\ &= Ax - B\bar{K}Tx \end{aligned}$$

We see that if we define

$$K = \bar{K}T, \tag{9}$$

then the closed-loop system

$$\dot{x} = Ax - BKx$$

has poles at the desired positions.

For the current example, we get

$$K = \begin{bmatrix} 86 & 12 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 86 & -74 \end{bmatrix}$$

We can check that the closed loop system has the desired poles. We have

$$\det(Is - A + BK) = s^2 + 20s + 101$$

as desired.

3.10. We recapitulate the procedure:

Step 1: Evaluate the *desired* characteristic polynomial and the characteristic polynomial of A.

Step 2: From the characteristic polynomial of A, deduce $\bar{A}$. Compute $C(A, B)$ and $C(\bar{A}, \bar{B})$. Set $T = C(\bar{A}, \bar{B})C(A, B)^{-1}$.

Step 3: Find $\bar{K}$ so that the characteristic polynomial of the closed loop system in CCF is the desired characteristic polynomial.

Step 4 Set $K = \bar{K}T$: this is the desired feedback law

4 Observability and observer design

We now consider the more general case in which we do not have access to the full state of the system, x, but only to the observation variable $y = Cx$. We only look at the case in which C is a row matrix, hence y is the inner product of C and x and y is thus a real number:

$$\left\{ \begin{array}{rcl} \dot{x} &=& Ax + Bu \\ y &=& Cx \end{array} \right. \tag{10}$$

The problem we will solve is thus the following: how to do feedback pole-placement when we do not have access to the whole state x, but only to y. A controller of the type $u = -Kx$ is thus not an option. One might think about trying $u = -Ky$ and finding an appropriate K; this in general does not work. What we will do is *estimate* the state x based on y. We call that estimate $\hat{x}$. We call the machinery that estimates x given the observation y an observer. Designing observers is the objective of this section.

4.1 Observer design

4.1. We deal with the following problem: given a system as in (10), and given that we have access to the signal $y(t)$, how to estimate $x(t)$? We **denote by $\hat{x}$ the estimate of x**. Thus $x(t)$ and $\hat{x}$ are vectors of the same dimension

4.2. The main idea is to *create an auxiliary dynamical system to estimate x*. This dynamical system will take advantage of our knowledge of the matrices A and C.

4.3. We call an **observer** for the system (10) a system of the type

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}). \tag{11}$$

In the equation above L has the same dimensions as x and $\hat{x}$ and $y = Cx$ is real.

The first two terms are as in (10), with x replaced by $\hat{x}$, and observe that if $\hat{x} = x$ (that is, if the estimation of x is correct), then the last term is zero (because $y = Cx$) and thus the auxiliary system (11) follows the original system (10).

4.4. But of course, we created the auxiliary system because we do not know x. We have additional degrees of freedom however in the choice of L: *can we choose L so that $\hat{x}$ will converge to x as t increases?*

4.5. While we have not formally studied questions about convergence of a variable to another variable for dynamical systems, we know how to make sure that a linear dynamical system converges to zero: set its poles in the left-hand-side of the plane. We introduce the **error variable**

$$e = x - \hat{x}. \tag{12}$$

Observe that *convergence of e to zero is the same as convergence of $\hat{x}$ to x.*

We will thus work with the error variable instead of with $\hat{x}$ directly. We can easily pass from one to the other using (12).

4.6. Let us write the differential equation that e satisfies:

$$\begin{aligned}
\dot{e} &= \frac{d}{dt}(x - \hat{x}) \\
&= \dot{x} - \dot{\hat{x}} \\
&= Ax + Bu - (A\hat{x} + Bu + L(y - C\hat{x})) \\
&= A(x - \hat{x}) - L(Cx - C\hat{x}) \\
&= Ae - LC(x - \hat{x}) \\
&= Ae - LCe \\
&= (A - LC)e
\end{aligned}$$

4.7. From the above point, we see that if we can choose L such that $A - LC$ has eigenvalues in the left-hand-size of the plane, then e will go to zero and thus $\hat{x}$ in (11) will go to x!

4.8. From now on, the analysis is similar to the one for feedback pole placement, where we had to choose K such that $A - BK$ had prescribed eigenvalues. When designing an observer, one can usually choose where to put the poles of the observer, and if one picks poles with very large negative real part, the observer will converge quickly to the real value of the state (that is, $\hat{x}$ will tend quickly to x). We recall that formally speaking $\hat{x} = x$ only at $t = \infty$.

4.2 Observability and duality with controllability

4.9. As already mentioned, the development here will be parallel to the one we did for output feedback pole placement. The equivalent of controllability in this setting is **observability**:

We say that a system (10) is **observable** if we can reconstruct the state x asymptotically from the knowledge of $y(t)$.

4.10. We state, without proof, the following fact: the system (10) is observable if the matrix

$$O(A,C) = \begin{bmatrix} — & C & — \\ — & CA & — \\ — & CA^2 & — \\ & \vdots & \\ — & CA^{n-2} & — \\ — & CA^{n-1} & — \end{bmatrix} \tag{13}$$

is of full rank, that is if $\det(O(A,C)) \neq 0$. Notice that C is a row vector, and thus CA, CA^2, etc. are also all row vectors and $O(A,C)$ is an $n \times n$ matrix.

4.11. We recall the OCF of a system. Given a system as (10), let

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \cdots + b_{n-1}s + b_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}$$

be the transfer function of the system. Recall that the denominator of $G(s)$ is also the characteristic polynomial of A.

The OCF of the system is given by

$$\tilde{A} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & 0 & \cdots & 0 & -a_{n-1} \\ 0 & 1 & 0 & \cdots & 0 & -a_{n-2} \\ \vdots & & \ddots & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & -a_2 \\ 0 & 0 & \cdots & 0 & 1 & -a_1 \end{bmatrix}; \quad \tilde{B} = \begin{bmatrix} b_n \\ b_{n-1} \\ \vdots \\ b_3 \\ b_2 \\ b_1 \end{bmatrix} \text{ and } \tilde{C} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{14}$$

As before, we also have that the characteristic polynomial of $\tilde{A}$ is the same as the characteristic polynomial of A and is

$$s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n$$

Thus we can easily construct $\tilde{A}$ from the characteristic polynomial of A.

4.12. If $\tilde{T}$ is a change of variables that put the system in OCF, we can, similarly to what we did for controllability, relate the observability matrix of the system in OCF and of the original system using $\tilde{T}$. Recall that after changing variables, we have

$$\tilde{A} = \tilde{T}A\tilde{T}^{-1}, \tilde{B} = \tilde{T}B \text{ and } \tilde{C} = C\tilde{T}^{-1}$$

Evaluating explicitly $O(\tilde{A}, \tilde{C})$, we get

$$\begin{aligned} O(\tilde{A},\tilde{C}) &= \begin{bmatrix} - & \tilde{C} & - \\ - & \tilde{C}\tilde{A} & - \\ - & \tilde{C}\tilde{A}^2 & - \\ & \vdots & \\ - & \tilde{C}\tilde{A}^{n-2} & - \\ - & \tilde{C}\tilde{A}^{n-1} & - \end{bmatrix} = \begin{bmatrix} - & C\tilde{T}^{-1} & - \\ - & C\tilde{T}^{-1}\tilde{T}A\tilde{T}^{-1} & - \\ - & C\tilde{T}^{-1}\tilde{T}A\tilde{T}^{-1}\tilde{T}A\tilde{T}^{-1} & - \\ & \vdots & \\ - & C\tilde{T}^{-1}\tilde{T}A\tilde{T}\cdots\tilde{T}A\tilde{T}^{-1} & - \\ - & C\tilde{T}^{-1}\tilde{T}A\tilde{T}\cdots\tilde{T}A\tilde{T}^{-1} & - \end{bmatrix} \\ &= \begin{bmatrix} - & C & - \\ - & CA & - \\ - & CA^2 & - \\ & \vdots & \\ - & CA^{n-2} & - \\ - & CA^{n-1} & - \end{bmatrix} \tilde{T}^{-1} \end{aligned}$$

We thus have

$$O(\tilde{A}, \tilde{C}) = O(A, C)\tilde{T}^{-1}$$

from which we deduce

$$\tilde{T} = O(\tilde{A}, \tilde{C})^{-1} O(A, C) \tag{15}$$

4.13. We now have enough to set up an observer for a system. We describe the procedure below and illustrate it on an example.

Question: Given

$$A = \begin{bmatrix} -15 & 8 \\ -15 & 7 \end{bmatrix} \text{ and } C = \begin{bmatrix} 1 & 1 \end{bmatrix},$$

design an observer for the system with poles at $-10 \pm j$.

Answer:

1. First check that the system is observable, that is that $O(A, C)$ is of full rank.

> We have that
>
> $$O(A, C) = \begin{bmatrix} 1 & 1 \\ -30 & 15 \end{bmatrix}$$
>
> which is of full rank, the system is thus observable.

2. Compute the desired characteristic polynomial for the observer and the characteristic polynomial of A.

> We have that the desired characteristic polynomial is
>
> $$(s + 10 - j)(s + 10 + j) = s^2 + 20s + 101$$
>
> The characteristic polynomial of A is
>
> $$\det(Is - A) = s^2 + 8s + 15$$

Recall that the problem is to find L such that the characteristic polynomial of $A - LC$ is the desired characteristic polynomial. The matrix comes from the auxiliary system

$$\dot{e} = Ae - LCe.$$

3. Find the change of variables $\tilde{T}$ that put the original system in OCF using (15). Using the same change of variables for $\hat{x}$, we set

$$\tilde{\hat{x}} = \tilde{T}\hat{x}.$$

We find that

$$\frac{d}{dt}\tilde{\hat{x}} = \tilde{T}\frac{d}{dt}\hat{x} = \tilde{T}A\tilde{T}^{-1}\tilde{\hat{x}} + \tilde{T}Bu - (y - C\tilde{T}^{-1}\tilde{\hat{x}}) = \tilde{A}\tilde{\hat{x}} + \tilde{B}u + (y - \tilde{C}\tilde{\hat{x}})$$

We can now set

$$\tilde{e} = \tilde{T}e$$

and we similarly get

$$\frac{d}{dt}\tilde{e} = \dot{\tilde{e}} = \tilde{T}A\tilde{T}^{-1}\tilde{e} - \tilde{T}LC\tilde{T}^{-1}\tilde{e} = \tilde{A}\tilde{e} - \tilde{L}\tilde{C}\tilde{e}$$

where we defined

$$\tilde{L} = \tilde{T}L.$$

Because changing coordinates does not change the position of the poles, we can try to find $\tilde{L}$ so that the poles of $\tilde{A} - \tilde{L}\tilde{C}$ are where desired. As before, if we choose the change of coordinates wisely, in this case such that the system is in OCF, we can do this easily.

We need to compute $O(A, C)$ and $O(\tilde{A}, \tilde{C})$. We already have the first one. From the characteristic polynomial of A, we can get $\tilde{A}$ and $\tilde{C}$ is always the same in OCF:

$$\tilde{A} = \begin{bmatrix} 0 & -15 \\ 1 & -8 \end{bmatrix} \text{ and } \tilde{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

We now compute

$$O(\tilde{A}, \tilde{C}) = \begin{bmatrix} 0 & 1 \\ 1 & -8 \end{bmatrix}$$

Thus, we get

$$\tilde{T} = \begin{bmatrix} -22 & 23 \\ 1 & 1 \end{bmatrix}$$

4. In OCF, one can easily choose the value of $\tilde{L}$ so that the system

$$\dot{\tilde{e}} = (\tilde{A} - \tilde{L}\tilde{C})\tilde{e}$$

has poles where desired. Indeed, we have that

$$\tilde{A} - \tilde{L}\tilde{C} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_n - \tilde{L}_1 \\ 1 & 0 & 0 & \cdots & 0 & -a_{n-1} - \tilde{L}_2 \\ 0 & 1 & 0 & \cdots & 0 & -a_{n-2} - \tilde{L}_3 \\ \vdots & & \ddots & & & \vdots \\ 0 & \cdots & 0 & 1 & 0 & -a_2 - \tilde{L}_{n-1} \\ 0 & 0 & \cdots & 0 & 1 & -a_1 - \tilde{L}_n \end{bmatrix}$$

We thus know that the characteristic polynomial of $\tilde{A} - \tilde{L}\tilde{C}$ is

$$s^n + (a_1 + \tilde{L}_n)s^{n-1} + \cdots + (a_2 + \tilde{L}_{n-1})s + a_n + \tilde{L}_1$$

We can choose $\tilde{L}_i$ so that this polynomial is the desired characteristic polynomial.

From the desired characteristic polynomial and the characteristic polynomial of A, we get that

$$-15 - \tilde{L}_1 = -101 \text{ and } -8 - \tilde{L}_2 = -20$$

from which we obtain

$$\tilde{L}_1 = 86 \text{ and } \tilde{L}_2 = 12.$$

5. We now return to the original coordinates. We had $\tilde{L} = \tilde{T}L$ and thus

$$L = \tilde{T}^{-1}\tilde{L} \tag{16}$$

With the values of $\tilde{L}$ and $\tilde{T}$ found above, we obtain

$$L = \begin{bmatrix} 4.222 \\ 7.778 \end{bmatrix}$$

6. You can verify that $\det(Is - A + LC)$ with the value of L just calculated is indeed the desired characteristic polynomial.

With the values of A, C and L as above, we indeed get that $\det(Is - A + LC) = s^2 + 20s + 101$.

5 Duality between controllability and observability

The procedure to set the poles using state feedback and to design an observer are rather similar. This is no accident: we have seen that the OCF was simply the 'transpose' of the CCF, where transpose has to be taken with a grain of salt since we also switch B and C. One say that controllability and observability are *dual*, in the sense that a system is controllable is similar to saying that the transpose system is observable. We can use this fact to design an observer using the procedure for controllability. This is in some sense an exercise in keeping track of what matrix in OCF corresponds to what matrix in CCF. We elaborate on the duality here, provide a general procedure to design an observer using this duality (this is similar to the procedure above, you can choose whichever you prefer to solve problem) and we illustrate it on an example.

5.1. Recall the form of the controllability matrix (6). We have seen that if a system is controllable, we can find a change of coordinates T that puts it in CCF. A similar statement holds here: if a system is *observable*, then there exists a change of coordinates that puts it in OCF form.

To see this, note that the system described by matrices A, B and C

$$\begin{cases} \dot{x} &= Ax + Bu \\ y &= Cx \end{cases} \tag{17}$$

is observable if and only if the system described by $A^\top, C^\top, B^\top$

$$\begin{cases} \dot{x} &= A^\top x + C^\top u \\ y &= B^\top x \end{cases} \tag{18}$$

is controllable.

It suffices to use the definitions of controllability and observability to verify that statement: saying that the system A, B, C is observable is equivalent to saying that $\det(O(A,C)) \neq 0$ and saying that $A^\top, C^\top, B^\top$ is controllable is saying that $\det(C(A^\top, C^\top)) \neq 0$. We write $C(A^\top, C^\top)$ explicitly:

$$\begin{aligned} C(A^\top, C^\top) &= \begin{bmatrix} | & | & | & & | \\ C^\top & A^\top C^\top & (A^\top)^2 C^\top & \cdots & (A^\top)^{n-1} C^\top \\ | & | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} — & C & — \\ — & CA & — \\ — & CA^2 & — \\ & \vdots & \\ — & CA^{n-2} & — \\ — & CA^{n-1} & — \end{bmatrix}^\top = O(A,C)^\top \end{aligned}$$

Recall that that for an arbitrary matrix X, we always have $\det(X) = \det(X^\top)$. We thus conclude that

$$\det(C(A^\top, C^\top)) \neq 0 \Leftrightarrow \det(O(A,C)) \neq 0.$$

5.2. Setting up an observer with specific poles requires finding L such that the characteristic polynomial

$$\det(Is - A + LC)$$

is a specific polynomial (obtained from the poles just mentioned). When we do output feedback pole placement, we want to find K such that the characteristic polynomial of

$$\det(Is - A + BK)$$

is a desired polynomial. Recall that the characteristic polynomials of any matrix X and its transpose $X^\top$ are the same; in the current setting, this yields

$$\det(Is - A + BK) = \det(Is - A^\top - K^\top B^\top)$$

If we let $K = L^\top$, we see that an observer design problem for (18) is the same as an output feedback pole placement problem for (17).

5.3. We can now give a general procedure to set up an observer using duality:

1. Check that the system is observable, that is that $\det(O(A, C)) \neq 0$. If the system is not observable, then one cannot set-up an observer.
2. If the system is observable, consider the auxiliary system

   $$\dot{e} = A^\top e - C^\top L^\top e = (A^\top - C^\top L^\top)e.$$

   It is similar to the system (8) we considered in the feedback pole-placement problem, with

   $$A \leftrightarrow A^\top, B \leftrightarrow C^\top, K \leftrightarrow L^\top.$$

3. Choose desired poles for the observer, make sure that they have negative real part.
4. Use the procedure developed in 10 page 17 to set the poles of a closed loop system to the desired positions.
5. The observer is

   $$\dot{\hat{x}} = A\hat{x} + Bu - LCy$$

   for the L just calculated.

5.4. Example

> We illustrate the procedure above on an example.
> *Question:* Given the system described by the matrices A, B given by
>
> $$A = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

construct a dynamical system whose state will converge to the state of x given that you only have access to $x_1 + x_2 + x_3$.

Answer: The last part of the question says that we only have access to $y = Cx$ for

$$C = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

1. We first check that the system is observable. We have

$$O(A, C) = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ 1 & 4 & -4 \end{bmatrix}$$

   and $\det(O(A, C)) \neq 0$. Hence, the system is observable and an observer exists.

2. We choose, arbitrarily because no poles were specified, to put the poles of the observer at $-10 \pm j$ and -10. This gives us a desired characteristic polynomial for $A - LC$:

$$(s + 10 - j)(s + 10 + j)(s + 10) = s^3 + 30s^2 + 301s + 1010$$

3. Consider the auxiliary system $\dot{e} = A^\top e - C^\top L^\top e$:

$$\dot{e} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 1 & 1 & -1 \end{bmatrix} e + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} L_1 & L_2 & L_3 \end{bmatrix} e$$

   We want the characteristic polynomial of this system to be $s^3 + 30s^2 + 301s + 1010$. We know that if our original system is observable, this auxiliary system is controllable. We thus use the procedure outline before.

4. Find T that puts the auxiliary system in CCF.

   (a) We don't need to calculate $C(A^\top, C^\top)$, since we know it is $O(A, C)^\top$, which we have calculated already:

$$C(A^\top, C^\top) = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & 4 \\ 1 & -1 & -4 \end{bmatrix}.$$

   (b) We compute the characteristic polynomial of $A^\top$ (We have seen that it is the same as the characteristic polynomial of A) We have

$$p_A(s) = \det(Is - A) = (s + 1)(s + 2)(s + 1) = s^3 + 4s^2 + 5s + 2.$$

   Thus

$$\bar{A}^\top = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix}.$$

   We know that in CCF, $\bar{C}^\top = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\top$. (This is $\bar{B}^\top$) We thus have

$$C(\bar{A}^\top, \bar{C}^\top) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -4 \\ 1 & -4 & 11 \end{bmatrix}$$

Using 4 on page 11, we find that

$$T = C(\bar{A}^\top, \bar{B}^\top) C(A^\top, C^\top)^{-1}.$$

For the current example, we find

$$T = \begin{bmatrix} -3 & 2 & 1 \\ 4 & -3 & -1 \\ -5 & 5 & 1 \end{bmatrix}$$

(c) Find the value of $\bar{L}$ that sets the characteristic polynomial of $\bar{A}^\top - \bar{C}^\top \bar{L}^\top$ to the desired characteristic polynomial. This is the same as step 3 page 15 where we were choosing $\bar{K}$. We get

$$-2 - \bar{L}_1 = -1010, -5 - \bar{L}_2 = -301, -4 - \bar{L}_3 = -30.$$

We get that

$$\bar{L}_1 = 1008, \bar{L}_2 = 296, \bar{L}_3 = 26.$$

(d) Recall that we had $K = \bar{K}T$. Here, we have $\bar{L} = \bar{K}^\top$ and $L = K^\top$. Taking the transpose of the previous equation, we get $L = K^\top = T^\top \bar{K}^\top = T^\top \bar{L}$. We thus set

$$L = T^\top \bar{L}.$$

For this example, we get

$$L = \begin{bmatrix} -1970 \\ 1258 \\ 738 \end{bmatrix}$$

(e) We can verify that the procedure worked:

$$\det(Is - A + LC) = s^3 + 30s^2 + 301s + 1010$$

as desired.

6 Output feedback pole placement

This section is the crowning achievement of our coverage of state-space control. We now know how to accomplish the following two tasks for the system

$$\left\{ \begin{array}{rcl} \dot{x} & = & Ax + Bu \\ y & = & Cx \end{array} \right. \tag{19}$$

1. Choose a feedback control law $u = -Kx$ so that the closed-loop system $\dot{x} = Ax - BKx$ has poles at a desired position.
2. Choose an L so that for the observer $\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x})$, $\hat{x}$ converges to x at a desired speed (that is, $A - LC$ has the desired eigenvalues)

Of course, in most situations encountered in practice, one does not have access to the full state, but only to the observable y. Can we replace x by $\hat{x}$ in the control law $\dot{u} = Kx \simeq K\hat{x}$? Quite remarkably, we can indeed do that. The fact that this would work is a consequence of what is called the **separation principle**. We will state it precisely below.

6.1. We now want to solve the following problem: given a system as in (19), and desired pole positions for the closed-loop system (for example, these poles can be obtained from rising time, etc.), can you design an *output feedback controller* so that the closed loop system has the desired poles?

6.2. We can solve the above problem if the system is *both observable and controllable.* The solution is then straightforward:

1. Find K so that the system with output feedback (that is $u = -Kx$) has poles at the desired location.
2. Find L which yields an observer

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

   for the system, preferably with poles faster than the ones required for the feedback.

3. Use the control law

$$u = -K\hat{x}.$$

6.3. The fact that we can set-up the observer and controller independently is called the **separation principle**. We describe now why this works.

6.4. The dynamics of the overall closed-loop system is

$$\left\{ \begin{array}{rcl} \dot{x} & = & Ax + Bu \\ y & = & Cx \\ u & = & -K\hat{x} \\ \dot{\hat{x}} & = & A\hat{x} + Bu + L(y - C\hat{x}) \end{array} \right.$$

which, making substitutions, can be written as

$$\left\{\begin{array}{rcl} \dot{x} & = & Ax - BK\hat{x} \\ \dot{\hat{x}} & = & A\hat{x} - BK\hat{x} + L(Cx - C\hat{x}) \end{array}\right. .$$

We can write the system in matrix form as follows:

$$\frac{d}{dt}\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \underbrace{\begin{bmatrix} A & -BK \\ LC & A - BK - LC \end{bmatrix}}_{A_{cl}} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} \tag{20}$$

The poles of the closed-loop system are thus the poles of the matrix A_{cl} appearing in (20). Observe that if our original system was of dimension n, the closed-loop system is now of dimension $2n$. This is not unlike what we did earlier with transfer functions: e.g. using a PID controller added poles/zeros to the original system (that is the one without controller).

6.5. We now show that the eigenvalues of A_{cl} (recall that it has $2n$ eigenvalues and they are the closed-loop poles) are the eigenvalues of $A - BK$ (there are n of them) and the eigenvalues of $A - LC$ (there are also n of them). **This remarkable fact is the separation principle**. This also justifies why we want the eigenvalues of the observer to be faster than the ones from the state feedback poles placement: we want these latter eigenvalues/poles to shape the behavior of the closed-loop system, not the ones of the observer.

6.6. In order to prove the separation principle, we find a change of coordinates that makes the computations of the eigenvalues of A_{cl} easy. We take the $2n$ by $2n$ matrix

$$T = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}$$

where I is the n by n identity matrix (careful, below we use I sometimes for the $2n$ by $2n$ identity matrix and sometimes for the n by n identity matrix.) Using this change of variables, the new coordinates are

$$T\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} x \\ x - \hat{x} \end{bmatrix} = \begin{bmatrix} x \\ e \end{bmatrix}$$

where we recall that we defined $e = x - \hat{x}$. We could compute the matrix $\bar{A}_{cl} = TA_{cl}T^{-1}$ or use a direct computation. We do both but start with the latter.

1. We note that $\hat{x} = x - e$. We first find a differential equation for x in terms of x and e:

$$\begin{array}{rcl} \dot{x} & = & Ax - BK\hat{x} \\ & = & Ax - BK(x - e) \\ & = & (A - BK)x + BKe \end{array}$$

   Similarly, we get for e

$$\begin{aligned}
\dot{e} &= \dot{x} - \dot{\hat{x}} \\
&= Ax - BK\hat{x} - (LCx + (A - BK - LC)\hat{x}) \\
&= (A - LC)x - BKx + BKe - (A - BK - LC)x + (A - BK - LC)e \\
&= (A - LC)e
\end{aligned}$$

Putting the two above together, we get

$$\frac{d}{dt}\begin{bmatrix} x \\ e \end{bmatrix} = \underbrace{\begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix}}_{\bar{A}_{cl}} \begin{bmatrix} x \\ e \end{bmatrix} \tag{21}$$

2. Alternatively, we have that $\bar{A}_{cl} = TAT^{-1}$. We check that $T^{-1} = T$, that is T is its own inverse:

$$TT = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

which is the $2n$ by $2n$ identity matrix.

We evaluate $\bar{A}_{cl}$ as follows:

$$\begin{aligned}
\bar{A}_{cl} &= \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} A & -BK \\ LC & A - BK - LC \end{bmatrix} \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \\
&= \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} A - BK & BK \\ A - BK & -A + BK + LC \end{bmatrix} \\
&= \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix}
\end{aligned}$$

6.7. Recall that the determinant of an upper triangular matrix is the product of its diagonal entries. Using rules for block computations, we see that the characteristic polynomial of $\tilde{A}_{cl}$, which is the same as the characteristic polynomial of A_{cl} since they are related by a change of variables, is

$$\det(Is - A_{cl}) = \det\left(\begin{bmatrix} Is - A + BK & -BK \\ 0 & Is - A + LC \end{bmatrix}\right) = \det(Is - A + BK)\det(Is - A + LC)$$

The roots of $\det(Is - A_{cl})$ are thus the roots of $\det(Is - A + BK)$ (the roots of our state-feedback system) union the roots of $\det(Is - A + LC)$ (the roots of our observer). This proves the separation principle.

6.8. Note that when we used root locus or Bode plot methods to design controllers, we did not have a separation principle, and thus we needed to iterate to place to poles exactly where desired.

Translated here, this would have meant that after using the control law $u = -K\hat{x}$ instead of $u = -Kx$, the poles would move, but we see it is not the case!

6.9. *Question: Consider a state-space system of type* (19) *with matrices*

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \text{ and } C = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix},$$

construct an output-feedback controller that stabilizes the system and such that the rise-time is of at most 2 *seconds and the overshoot to a unit step is of no more than 25%.*

Answer: We can solve this problem in different ways: one way would be to first compute the transfer function of the system:

$$G(s) = C(Is - A)^{-1}B = \frac{(s-1)^2}{s^3 - 6\,s^2 + 11\,s - 7}$$

Based on this transfer function, we can use the root locus method to set-up a lead-lag controller with the desired closed-loop poles.

We use a state-space approach here instead. Because we do output feedback, we have to set-up an observer as well.

1. Check whether the system is controllable and observable.

We have

$$C(A,B) = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 4 \\ 1 & 3 & 10 \end{bmatrix} \text{ and } O(A,C) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & 3 \\ 4 & 9 & 10 \end{bmatrix}.$$

Both $\det(C(A,B)) \neq 0$ and $\det(O(A,C))) \neq 0$ and this the system is controllable and observable.

2. Find the desired characteristic polynomial for the output feedback closed loop system.

It is a system of order 3, it will thus have 3 poles. The specifications give us two poles:

$$\omega_n \geq \frac{1.8}{t_r} \text{ and } M_p < 25\%.$$

We take, for example $\omega_n = 1$ and $\xi = 0.6$. We now use the canonical second order polynomial to find 2 desired poles:

$$s^2 + 2\xi\omega_n s + \omega_n^2 = s^2 + 1.2s + 1$$

This yields the two desired poles: $-0.6 \pm 0.8j$.

We need to choose one additional pole, making sure it is faster than the 2 desired ones (that is with a smaller negative real part) We take for example a third pole at -5. The desired characteristic polynomial is

$$(s+5)(s+0.6+0.8j)(s+0.6-0.8j) = s^3 + 31/5s^2 + 7s + 5.$$

3. Find the desired characteristic polynomial for the observer.

No poles have been specified for the observer. From the separation principle, we know that the poles of the system with an output feedback controller and an observer are the poles of the output feedback portion and the poles of the observer section. We want to make sure that the poles of the observer are not the dominant poles. We set them, for example, at -20. The desired characteristic polynomial for the observer is thus

$$(s+20)^3 = s^3 + 60s^2 + 1200s + 8000.$$

4. Evaluate the characteristic polynomial of A. (It is needed to find the canonical forms)

We find

$$\det(Is - A) = s^3 - 6s^2 + 11s - 7.$$

5. Find T that puts the system in CCF and $\tilde{T}$ that puts the system in OCF.

We have that

$$\bar{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 7 & -11 & 6 \end{bmatrix}; \bar{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Thus

$$C(\bar{A}, \bar{B}) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 6 \\ 1 & 6 & 25 \end{bmatrix}.$$

From formula (7), we get

$$T = C(\bar{A}, \bar{B})C(A, B)^{-1} = \begin{bmatrix} -1 & -1 & 1 \\ -2 & -1 & 2 \\ -3 & 0 & 4 \end{bmatrix}.$$

Similarly for $\tilde{T}$, we find evaluate $\tilde{A}$, which is just the transpose of $\bar{A}$:

$$\tilde{A} = \begin{bmatrix} 0 & 0 & 7 \\ 1 & 0 & -11 \\ 0 & 1 & 6 \end{bmatrix} \text{ and } \tilde{C} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

We then get

$$O(\tilde{A}, \tilde{C}) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 6 \\ 1 & 6 & 25 \end{bmatrix}.$$

Note that the above is just the transpose of $C(\bar{A}, \bar{B})$ (and here, because the matrix is symmetric, it is equal to its transpose, but it is not always the case).

We thus get from equation (15) that

$$\tilde{T} = O(\tilde{A}, \tilde{C})^{-1} O(A, C) = \begin{bmatrix} -2 & 2 & 3 \\ 1 & -3 & -3 \\ 0 & 1 & 1 \end{bmatrix}.$$

You can double check the calculation by verifying that

$$TAT^{-1} = \bar{A} \text{ and } \tilde{T}A\tilde{T}^{-1} = \tilde{A}.$$

6. Find $\bar{K}$ and $\tilde{L}$ so we get the desired characteristic polynomial for the output feedback and the observer.

From the characteristic polynomial of A and the desired characteristic polynomial for the output feedback we deduce

$$6 - \bar{K}_3 = -31/5; -11 - \bar{K}_2 = -7; 7 - \bar{K}_1 = -5.$$

This yields

$$\bar{K} = \begin{bmatrix} 12 & -4 & 12.2 \end{bmatrix}$$

We can verify that $\det(Is - \bar{A} + \bar{B}\bar{K}) = s^3 + 31/5s^2 + 7s + 5$. From the characteristic polynomial of A and the desired characteristic polynomial of the observer, we get

$$-6 - \tilde{L}_3 = -60; -11 - \tilde{L}_2 = -1200; 7 - \tilde{L}_1 = 8000.$$

This yields

$$\tilde{L} = \begin{bmatrix} 8007 \\ 1189 \\ 66 \end{bmatrix}$$

You can check that $\det(Is - \tilde{A} + \tilde{L}\tilde{C}) = s^3 + 60s^2 + 1200s + 8000$.

7. Find K and L.

Using formula (9), we have

$$K = \bar{K}T = \begin{bmatrix} -40.6 & -8 & 52.8 \end{bmatrix}$$

We can verify that $\det(Is - A + BK) = s^3 + 31/5s^2 + 7s + 5$.

Using formula (16), we have

$$L = \tilde{T}^{-1}\tilde{L} = \begin{bmatrix} 1387 \\ -10583 \\ 10649 \end{bmatrix}.$$

We can verify that $\det(Is - A + LC) = s^3 + 60s^2 + 1200s + 8000$.

8. Write down the closed-loop system.

We have

$$\begin{cases} \dot{x} &= Ax + Bu \\ y &= Cx \\ u &= -K\hat{x} \\ \dot{\hat{x}} &= A\hat{x} + Bu + L(y - C\hat{x}) \end{cases}$$

with A, B and C as given in the statement and L and K as above. We can check that the dynamics of the closed loop system have poles where desired.We know from (20) that the overall dynamics is (You need to evaluate products BK and LC to obtain this matrix)

$$\frac{d}{dt}\begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & \frac{203}{5} & 8 & -\frac{264}{5} \\ 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & \frac{203}{5} & 8 & -\frac{264}{5} \\ 0 & 1387 & 1387 & \frac{208}{5} & -1379 & -\frac{7194}{5} \\ 0 & -10583 & -10583 & 1 & 10585 & 10583 \\ 0 & 10649 & 10649 & \frac{203}{5} & -10640 & -\frac{53494}{5} \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

Using Matlab, we find that the eigenvalues are $-20, -20, -20, -5, -0.6 \pm 0.8j$ as desired.