Linear System Theory

Ph. Müllhaupt

11-10-2013

Contents

Part I

Linear Systems

Chapter 1

Introduction

1.1 References

The main reference is:

- Chi-Tsong Chen *Linear System Theory and Design*, Oxford University Press, 1999.

Noteworthy are the following references:

- H.H. Rosenbrock *State Space and Multivariable Theory*, Thomas Nelson & Sons, 1970.
- F.R. Gantmacher *The Theory of Matrices*, vol. I & II, AMS Chelsea, 1977 (1959, 1960).
- S.K. Godunov *Modern Aspects of Linear Algebra*, AMS, 1998.
- P. Lancaster *Lambda-Matrices and Vibrating Systems*, Oxford, Pregamon, 1966.

1.2 Possible Representations of Linear Systems

Among possible representations of the dynamical character of a system possessing the property of the superposition of its solutions, we will consider four of them, namely the transfer function (when the system has a single input and single output), the state space formulation (without restricting the system to have a single input and/or outputs), the matrix polynomial fractions (a generalization of the transfer function to the multi-input multi-output case), and finally polynomial system matrices, which can be seen as a hybrid between pure state space and matrix fractions.

1.3 The Superposition Principle

Consider a single-input single-output system (SISO) as a black box model in which the only way to interact with the system is to set its input u to a time function

$$u(.) : t \to u(t), \qquad t \in [0, +\infty),$$

and to collect, at its output y, the time output response

$$y(.) : t \to y(t), \qquad t \in [0, +\infty).$$

The input time function is also written $u(t)$ by abuse of notation, because one could argue that $u(t)$ represents the value of $u(t) \in \mathbb{R}$ rather than the time function $u(.) : t \to u(t)$, which we will write $u(.)$ so as to alleviate notational ambiguities whenever necessary. Sometimes we will write $u(t)$ to designate the time function for simplicity.

Definition 1 (Superposition Principle) *A single-input single-output system with input u and output y is linear when, for each possible choices of different input time functions $u_1(.)$ and $u_2(.)$ — generating the respective responses $y_1(.)$ and $y_2(.)$ — the response to the superposition of both input time functions*

$$u(.) := u_1(.) + u_2(.)$$

is the sum of the individual responses

$$y(.) = y_1(.) + y_2(.).$$

1.4 Convolution integrals

- $u(t) = \displaystyle\int_0^\infty u(\tau)\, \delta(t - \tau)\, d\tau$
- $y(t) = \displaystyle\int_0^\infty g(t - \tau)\, u(t)\, d\tau$
- $y(t) = u(t) * g(t) \Longrightarrow Y(s) = U(s) \cdot G(s)$

1.5 Transfer Functions

The transfer function approach is adequate for a linear system that has a single input u, $\dim u = 1$ and a single output y, $\dim y = 1$. For continuous-time system we will write

$$G(s) = \frac{Y(s)}{U(s)} = \frac{B(s)}{A(s)} = \frac{\sum_{i=1}^{n} b_i s^{n-i}}{s^n + \sum_{i=1}^{n} a_i s^{n-i}}$$

where $U(s)$, $Y(s)$ designate the Laplace transform of $u(t)$, $y(t)$

$$U(s) := \int_{t=0}^{+\infty} u(t) e^{-st},$$

$$Y(s) := \int_{t=0}^{+\infty} y(\tau) e^{-st},$$

and where the corresponding dynamical linear relationship is encoded by two polynomials $B(s)$, the numerator polynomial, and $A(s)$, the denominator polynomial. We will see how specific properties of the time behavior of $y(t)$ for a given $u(t)$, such as for instance stability, controllability, observability, can be read off from specific properties of the polynomials $A(s)$ and $B(s)$.

1.6 State Space

Transfer functions are clearly inadequate to handle multi-input systems. Additionally, it is not quite clear how to simulate systems arising from transfer functions, that is, to create an algorithm, a computer program, to simulate the system so as to obtain $y(t)$ knowing $u(t)$.

The main problem is in finding the right number of *initial conditions* and the *variables associated* with these initial conditions. Initial conditions can be understood as the minimal knowledge to completely specify the future time evolution of a system with the complete knowledge of the input $u(t)$, $t \in [0, +\infty)$.

We collect such initial conditions in a vector

$$x(0) = \begin{pmatrix} x_1(0) & x_2(0) & \dots & x_n(0) \end{pmatrix}^T.$$

Notice that these initial conditions are specifically indexed by a time variable set to 0. This means that these initial conditions can also be seen as evolving quantities along the system, that is, they become variables in their own right $x(t)$, and have trajectories (solutions) associated with them $x(.)$.

These variables x_1, x_2, ..., x_n are called the state space variables, and the vector x, the state vector. A very important point is that *these variables are far from unique* given *the same transfer function*. The theory of *representations* will help sort out this issue. We will need tools to check and make sure that the state variables that we have chosen are indeed associated with the transfer function (or matrix fraction description, see below) and then also, maybe, to construct a better suited set of equivalent state variables (leading to the same transfer function) which are less sensitive to numerical round off, to numerical imprecisions caused by representing these variables with an arithmetic of limited precision.

The philosophical interpretation of this, is that one can understand the state space variables, either as evolving in their own right according to the system for a given set of a priori initial conditions specified long back, namely writing $x(t)$, with $t > 0$, or as initial conditions at the current time, in which case we simply relabel the time axis with $t = 0$ so as to designate this specific time where the initial conditions are set. In the former case, $x(t)$ means that the system has evolved on its own (under for example the influence of the input $u(t)$, $t \in [0; \infty)$) from initial time $t = 0$ (where the initial condition $x(0)$ has been specified) up to the time instant t for which these same variables have acquired the values $x(t)$.

Now, the main question is: How do we describe the time evolutions of these variables ? The answer is through a set of ordinary differential equations of the first order, one ordinary differential equation per state-space variable, so that we have one initial condition per differential equation. Since all these ordinary differential equations can be coupled with each other, the simplest description is through four matrices of real numbers $A \in R^{n\times n}$, $B \in R^{n\times m}$, $C \in R^{p\times n}$ and $D \in R^{p\times m}$:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{1.1}$$

1.7 Matrix Polynomial Fractions

A matrix polynomial fraction is a generalization of the transfer function to the multi-input multi-output case. Polynomials $A(s)$ and $B(s)$ appearing in the transfer function $G(s)$ are replaced by matrices $N(s)$ and $D(s)$ having polynomials as entries, hence the name *polynomial matrix*. A polynomial matrix will be written for example as $M(s)$ with each entry in the matrix a polynomial in s. For example

$$M(s) = \begin{pmatrix} 3s^2+2s & 2s+1 & s+2 \\ s^2+s-3 & s & s^3+3 \end{pmatrix}$$

There is no need to restrict to square matrices only. Now, if one wants to mimic a transfer function, one would like to divide the matrix $M(s)$ by another polynomial matrix, say $N(s)$. Clearly, the division should be replaced by multiplying by the inverse matrix (dividing is nothing but multiplying by the inverse). Nevertheless, we stumble on two difficulties:

1. Because matrix multiplication is not a commutative operation

$$N(s)D(s)^{-1} \neq D(s)^{-1}N(s),$$

   even for square polynomial matrices $N(s)$ and $D(s)$.

2. The inverse of $D(s)$ might introduce polynomials appearing in the denominator of some, or maybe all entries of $D(s)^{-1}$.

The second aspect does not appear in transfer functions. The denominator polynomial $A(s)$ is always a polynomial. However $N(s)$ must be restricted somehow. The very interesting fact is that we need only consider a specific subset (or subclass) of polynomial matrices for the polynomial matrix $N(s)$. The matrix $M(s)$ stays a general polynomial matrix without restriction. The matrix $N(s)$, however, should belong to the set of polynomial matrices having a constant as its determinant (that is, the determinant should not a polynomial). Such matrices having a degree zero polynomial in s as determinant are called *unimodular matrices.*

Definition 2 (Polynomial Matrix) *A polynomial matrix $M(s)$ is a $n \times m$ matrix, the entries of which $m_{ij}(s)$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ are polynomials in s.*

Definition 3 (Unimodular Polynomial Matrix) *A polynomial matrix is called unimodular whenever its determinant is a constant (not a polynomial of degree higher than zero).*

1.8 Polynomial System Matrices

A question that comes naturally to mind is: Is there anyway to link the multivariable aspect of state space, which is very natural, to the polynomial matrix description ? Is there any "intermediate representation" ? The answer is surprisingly very fruitful and this new representation is what will be called polynomial system matrices.

Consider the Laplace transform of the state-space representation

$$\begin{aligned} 0 &= sX(s) - AX(s) - BU(s) \\ Y(s) &= CX(s) + DU(s) \end{aligned} \tag{1.2}$$

It is then possile to introduce a single matrix, with polynomials, that completely characterizes the system:

$$\mathfrak{S} = \begin{pmatrix} sI - A & -B \\ C & D \end{pmatrix}$$

But one could as well put any polynomial in such a matrix, that is,

$$\mathfrak{S} = \begin{pmatrix} T(s) & R(s) \\ P(s) & W(s) \end{pmatrix}$$

keeping in mind the relation

$$\begin{pmatrix} 0 \\ Y(s) \end{pmatrix} = \mathfrak{S} \begin{pmatrix} X(s) \\ U(s) \end{pmatrix}$$

for which $T(s) = sI - A$, $P(s) = -B$, $R(s) = C$, $W(s) = D$ is only a special case. This allows writing implicit systems, that is, systems for which the equations are not solved explicitly for the highest derivative. Additionally, one can also handle not only first order differential equations, as it is the case with the classical state-space equations, but also higher order differential equations.

The reasons why it can be considered as a sort of hybrid between matrix polynomials and the state-space representation is that:

(i) it uses polynomials in the entry of the system matrix $\mathfrak{S}$;

(ii) it embbodies an explcit treatment of certain variables $X(s)$ which are the Laplace transform of variables associated with state variables whenever explicit first order differential equations are considered.

These variables appear as an explicit argument to the operator associated with the system matrix $\mathfrak{S}$, which is almost never the case (unless the system has only very few states) with polynomial matrix description.

Chapter 2

Linear Algebra

In the introduction, we have presented a dynamical linear system through a combination of ordinary differential equations. All representation of such a system rely on a set of real numbers most often regrouped into matrices of scalars or polynomials. We will study in this chapter properties of the set of numbers only, which are regrouped in matrices and vectors. We will encounter also polynomials, but in a slightly different flavour than in the introduction. These new polynomials (not to be confused with those appearing in the representations of the introduction) are essentially linked to properties of the constant scalar matrices, although they will reappear also when studying matrix fraction representations and system matrices, both of which have explicitly polynomials in their definition. It is really the property of these constant matrices that interest us in this chapter irrespective of the differential equations to which they are possibly attached or not.

Linear algebra handles algebraic operations that preserve linearity, that is, preserves proportionality both mulitplicatively and additively. The most familiar concept is that of a vector space together with maps between vector spaces. These maps can be specified either in an abstract fashion or concretely through their representations by matrices. Elements of a vector space can also be represented abstractedly as subspaces of a vector space — *without choosing* any kind of coordinate systems — or as a collection of elements of the field of definition, collected in a row or column, and termed *a vector*, which explicitly admits a concrete parameterization using a well-defined choice of coordinates.

Because of the large possible ways to choose coordinates, the following natural questions come to mind: To what extent, and to what quantifiable degree, are maps and elements of a vector space identical ? For instance, given a certain matrix A, which represents a given action on a vector space described through certain coordinate systems, how can one determine whether it corresponds to another matrix B using *another* coordinate system but acting on the *same* vector space ? Otherwise stated, what are the inherent invariants of the matrix or the abstract operator that will not change with the choice of coordinates ?

2.1 Vector Space

Definition 4 *A vector space $\mathcal{V}$ over a field $\mathbb{F}$ (e.g. the field or real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$) is a set $\mathcal{V} = \{v_1, v_2, \ldots\}$, finite or infinite, together with*

- *an operation $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$*
- *an operation $\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}$*

such that $\forall v_1, v_2, v_3 \in \mathcal{V}$ and $\forall \alpha, \beta \in \mathbb{F}$ and designating by 1 the unit element of the field $\mathbb{F}$, the following identities hold:

1. *$v_1 + v_2 = v_2 + v_1$; (commutativity of $+$);*
2. *$(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$; (associativity of $+$);*
3. *There exists $0 \in \mathcal{V}$ such that $0 + v_1 = v_1 + 0 = v_1$; (neutral element);*
4. *$-v_1 \in \mathcal{V}$ such that $v_1 + (-v_1) = 0$; (opposite);*
5. *$\alpha \cdot (\beta \cdot v_1) = (\alpha\beta) \cdot v_1$; (associativity of $\cdot$);*
6. *$(\alpha + \beta) \cdot v_1 = \alpha \cdot v_1 + \beta \cdot v_1$; (distributivity of $\cdot$ over $+$ of the field $\mathbb{F}$)*
7. *$1 \cdot v_1 = v_1$; (compatibility of 1);*
8. *$\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$ (distributivity of $\cdot$ over $+$ of the vector space).*

2.2 Basis

A basis B of vector space $\mathcal{V}$ is a particular set of vectors that generate the vector space through suitable combinations. An important property so as to constitute a basis is that such a set must be minimal with respect to the number of constituting elements. The way these vectors are combined is given by *linear combinations*. From now on, we will drop explicit mention of the multiplication operator $\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}$.

Definition 5 (Linear combination) *A linear combination of elements $v_1, v_2, \ldots, v_n \in \mathcal{V}$ defined by the elements $\alpha_1, \alpha_2, \ldots, \alpha_n$ of the field $\mathbb{F}$ is, by definition, the vector $v \in \mathcal{V}$ given as*

$$v := \sum_{i=1}^{n} \alpha_i v_i.$$

Definition 6 (Linear independence) *A finite set of vectors $v_1, v_2, \ldots, v_n$ is called linearly independent whenever the solution to*

$$\sum_{i=1}^{n} \alpha_i v_i = 0$$

with $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, n$ is unique and given by

$$\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

Using this definition, it is possible to combine vectors using arbitrary linear combinations to generate a new set.

Lemma 1 (Vector space generated by a set of vectors) *Given a set of vectors $v_1, v_2, \ldots, v_p$, not necessarily linearly independent, all possible linear combination $\sum_{i=1}^{p} \alpha_i v_i$ for all possible choices of $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, p$, generate a set which is a vector space, that is, the set generated satisfies all axioms of Definition 4.*

Proof: The proof is a straightforward verification of the axioms. For example, Axiom 8 can be proved in the following way:

$$(\alpha + \beta) \sum_{i=1}^{p} \alpha_i v_i = \sum_{i=1}^{p} (\alpha + \beta) \alpha_i v_i = \sum_{i=1}^{p} (\alpha \alpha_i + \beta \beta_i) v_i =$$
$$\sum_{i=1}^{p} \alpha \alpha_i v_i + \sum_{i=1}^{p} \beta \beta_i v_i = \alpha \sum_{i=1}^{p} \alpha_i v_i + \beta \sum_{i=1}^{p} \beta_i v_i.$$

The other axioms are left to the reader as an exercise. ♠

When one imposes not only the condition of generating a vector space, but also of being linearly independent (that is, of being minimal among the spanning sets of vectors), then one has a basis. Minimality will be established through a lemma.

Definition 7 (Basis) *A basis B of a vectorspace $\mathcal{V}$ is a set of vectors $B = \{v_1, v_2, \ldots, v_n\}$ such that the vectors v_1, v_2, ..., v_n are linearly independent according to Def. 6 and generate $\mathcal{V}$ according to Lemma 1.*

Lemma 2 *The number n is independent of the generating set chosen in Def. 7.*

Proof: Suppose that there exist another generating and linearly independent set $\{w_1, w_2, \ldots, w_p\}$ with $p < n$. This means that every vector v_i, $i = 1, \ldots, n$, can be expressed through a linear combination of the w_j, $j = 1, \ldots, p$, i.e.

$$v_i = \sum_{j=1}^{p} \beta_{ij} w_j \qquad i = 1, \ldots, n. \tag{2.1}$$

Now, constitute the equation of linear independence

$$\sum_{i=1}^{n} \alpha_i v_i = 0 \tag{2.2}$$

for which the only solution is $\alpha_i = 0$, $i = 1, \ldots, n$, because the v_i are supposed to constitute a basis. We will now display a contradiction. Indeed, once the v_i, $i = 1, \ldots, n$, appearing in (2.2), are replaced by their expressions (2.1), one gets

$$\sum_{i=1}^{n} \alpha_i \left(\sum_{j=1}^{p} \beta_{ij} w_j \right) = \sum_{j=1}^{p} \left(\sum_{i=1}^{n} \alpha_i \beta_{ij} \right) w_j.$$

Then, because w_j, $j = 1, \ldots p$, constitute a basis, all coefficients in front of w_j, $j = 1, \ldots, p$, must vanish, that is,

$$\sum_{i=1}^{n} \alpha_i \beta_{ij} = 0 \qquad j = 1, \ldots, p. \tag{2.3}$$

But the system of equations (2.3) has more unknowns α_i, $i = 1, \ldots, n$, than the number of equations, p. Therefore, one can secure a set of scalars $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, n$ with at least one non-vanishing element $\alpha_k \neq 0$, $k \in 1 \ldots n$ such that (2.2) holds. This contradicts the fact that the v_i, $i = 1, \ldots, n$ are linearly independent. The case with $p > n$ follows a similar argument once p and n are swapped. Therefore, we have proved that all bases have the same number of vectors n. ♠

Definition 8 (Dimension) *The integer n appearing in Def. 7, which is guaranteed to be independent of choices through Lemma 2, is the dimension of the vector space $\mathcal{V}$.*

Lemma 3 (Coefficients w.r.t. a basis of an abstract linear map) *Given a basis $B = \{v_1, \ldots, v_n\}$ and a linear map $\phi : \mathcal{V} \to \mathcal{V}$, if $\alpha_1, \alpha_2 \in \mathcal{F}$, then $\phi(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 \phi(v_1) + \alpha_2 \phi(v_2)$.*

2.3 Matrices and Operators

Let $\mathcal{V}$ be an n dimensional vector space. We can select for such a space many different bases. Let us consider two of them $B_1 = \{v_1, \ldots, v_n\}$ and $B_2 = \{w_1, \ldots, w_n\}$ with $v_i, w_i \in \mathcal{V}$, $i = 1, \ldots, n$.

Lemma 4 (Uniqueness w.r.t. a given basis)

Once a basis $B = \{v_1, \ldots, v_n\}$ is chosen, any vector $v \in \mathcal{V}$ can be expressed as a unique linear combination with coefficients $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, n$ as

$$v = \sum_{i=1}^{n} \alpha_i v_i. \tag{2.4}$$

Proof: Suppose that there exists another set of coefficients $\bar{\alpha}_i \in \mathbb{F}$, $i = 1, \ldots, n$ such that $v = \sum_{i=1}^{n} \bar{\alpha}_i v_i$. Then substract this last expression to (2.4) so as to have

$$0 = \sum_{i=1}^{n} (\bar{\alpha}_i - \alpha_i) v_i.$$

This contradicts linear independence of the vectors of the basis B, because, by assumption the linear combination defined by the α's is different from the one defined by the $\bar{\alpha}$'s, there exists at least one k for which $\bar{\alpha}_k \neq \alpha_k$. ♠

Thus, once a basis is chosen, then one can represent a vector through a collection of n coefficients belonging to the base field $\mathbb{F}$.

Definition 9 (Vector in coordinates) *A vector — in coordinates, with respect to a basis $B = \{v_1, \ldots, v_n\}$ — is a collection of n coefficients $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, n$ written as a column (or row depending on convention): $v = \begin{pmatrix} \alpha_1 & \ldots & \alpha_n \end{pmatrix}^T$. It will be understood that this notation means that $v = \sum_{i=1}^{n} \alpha_i v_i$.*

Here a^T designates the operation of transforming a row vector a into a column vector a^T (and vice-versa).

Definition 10 (LINEAR MAP — LINEAR OPERATOR)) *A map $\Phi : \mathcal{V} \to \mathcal{V}$ will be called a linear map (or linear operator), whenever $\forall \alpha, \beta \in \mathbb{F}$, $\forall v_1, v_2 \in \mathcal{V}$, the following equality holds:*

$$\phi(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 \phi(v_1) + \alpha_2 \phi(v_2).$$

Lemma 5 (LINEAR MAP DEFINED THROUGH A BASIS) *Suppose that a basis is given as $B = \{v_1, \ldots v_n\}$, then any choice of coefficients $\alpha_{ij} \in \mathbb{F}$, $i, j \in 1, \ldots, n$ defines a linear map ϕ after setting $\phi(v) = w$ in which $v = \sum_{i=1}^n \alpha_i v_i$, and $w = \sum_j \beta_j v_j$ with the β_j's defined as*

$$\beta_i := \sum_{j=1}^n \alpha_{ij} \alpha_j \qquad i = 1, \ldots, n. \tag{2.5}$$

Proof: Let $v = \sum_{i=1}^n \alpha_i v_i$ and $w = \sum_{i=1}^n \gamma_i v_i$ then

$$\begin{aligned} \phi(\alpha v + \beta w) &= \phi\left(\sum_{i=1}^n (\alpha \alpha_i + \beta \gamma_i) v_i\right) = \sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} (\alpha \alpha_j + \beta \gamma_j)\right) v_i \\ &= \alpha \sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} \alpha_j\right) v_i + \beta \sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} \gamma_j\right) v_i \\ &= \alpha \phi(v) + \beta \phi(w) \end{aligned}$$

so that the defining property of linearity is satsfied, proving that Φ is a linear map. ♠

Lemma 6 (COEFFICIENTS OF A LINEAR MAP W.R.T. A BASIS) *Once a basis $B = \{v_1, \ldots, v_n\}$ is chosen, then each abstract linear map $\phi : \mathcal{V} \to \mathcal{V}$ leads to a unique choice of coefficients α_{ij}, $i, j \in 1, \ldots, n$ appearing in (2.5).*

Proof: Since $v_i \in \mathcal{V}$, $i = 1, \ldots, n$, one can apply the map ϕ to them so as to obtain a collection of vectors $\phi(v_i)$, $i = 1, \ldots, n$. Now, because B is a basis, each $\phi(v_i)$ admits a unique decomposition $\phi(v_i) = \sum_{j=1}^n \alpha_{ij} v_j$. Hence, all coefficicients α_{ij}, $i, j = 1, \ldots, n$ appearing in (2.5) are unique and well defined by ϕ. ♠

Definition 11 (Matrix of a linear map) *The coefficients α_{ij}, $i, j = 1, \ldots, n$ appearing in Lemma 6 can be arranged as a $n \times n$ array or matrix A that can be used to represent the linear map ϕ,*

$$A = \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & \ldots & \vdots \\ \alpha_{n1} & \ldots & \alpha_{n2} \end{pmatrix}$$

It operates on the column vector representation of a vector $v = \sum_{i=1}^{n} \alpha_i v_i$ in the following way to signify that $w = \sum_{i=1}^{n} \beta_i v_i$ is the image of v through the linear map ϕ, that is, $w = \phi(v)$.

2.4 Minimal Polynomials

We follow closely F. R. Gantmacher *Matrix Theory*, Vol. 1, chap. III. In this section, we will consider an n dimensional vector space $\mathcal{V}$ together with a vector $v \in \mathcal{V}$.

2.4.1 Polynomials define linear maps

Suppose a polynomial in λ is given as

$$\phi(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n 1$$

with all coefficients a_1, a_2, …, a_n belonging to the field $\mathbb{F}$.

Depending on the nature of the variable λ, the polynomial defines various maps. For instance, if we restrict λ to values belonging to $\mathbb{F}$ we get a map $\phi(.) : \mathbb{F} \to \mathbb{F}$. If we restrict λ to matrices we get the following:

Definition 12 (Linear map defined by a polynomial and a matrix) *Suppose that a matrix A is given together with a polynomial $\phi(\lambda)$, whose coefficients belong to $\mathbb{F}$. Then, upon substitution $\lambda \to A$, $1 \to I$, where I stands for the identity matrix, one gets a new matrix*

$$\phi(A) = A^n + \sum_{i=1}^{n} \alpha_i A^{n-i} \tag{2.6}$$

which is the linear map (in matrix form) corresponding to A and $\phi(\lambda)$.

Consider a linear map ϕ_1, leading to a matrix A_1 and another one ϕ_2, leading to A_2. Since it is linear, it's allowed to do: $\phi_1 + \phi_2$, leading to $A_1 + A_2$.

Suppose that $dim A_1 = dim A_2 = n \times n$, so we can multiply $A_1 \cdot A_2$ and $A_1 \cdot A_1$.

The question of interest is whether some the properties of this polynomial will be invariant when the basis is changed.

$$z = P \cdot x \qquad P^{-1}z = x$$
$$A : \mathcal{V} \to \mathcal{V} \qquad Ax = z \qquad PAP^{-1} = \bar{A}$$
$$\Psi(PAP^{-1}) = \Psi(A)$$

2.4.2 Annihilating polynomials

A linear map is given through a matrix A. Consider an arbitrary vector v given in column form so that A can operate on it by matrix multiplication, that is, Av is a new vector of $\mathcal{V}$ in column form. Thus, the matrix A can operate on it again so as to obtain $A(Av) = A^2 v$. Continuing this process — and since the vector space $\mathcal{V}$ is of finite dimension — there exists a p for which $A^p v$ becomes linearly dependent on the family of linearly independent vectors

$$B_v = \{v, Av, \ldots, A^{p-1}v\}.$$

This means that there exist $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, p$, such that

$$A^p = -\alpha_1 A^{p-1} - \alpha_2 A^{p-2} - \ldots - \alpha_p I. \tag{2.7}$$

Thus, the polynomial

$$\phi(\lambda) := \lambda^p + \alpha_1 \lambda^{p-1} + \ldots + \alpha_p$$

becomes an annihilating polynomial of v since (2.7) is nothing but

$$\phi(A)v = 0.$$

Definition 13 (Annihilating polynomial of a vector) *A polynomial $\phi(\lambda)$ is an annihilating polynomial of a vector $v \in \mathcal{V}$, w.r.t. a matrix A (linear operator), if*

$$\phi(A)v = 0.$$

Notice that an annihilating polynomial is specific to a linear map and might change with the chosen vector $v \in \mathcal{V}$. Thus, a natural question is to construct annihilating polynomials for a set of vectors and, therefore, for the corresponding vector subspace spanned by this set of vectors.

Definition 14 (ANNIHILATING POLYNOMIAL OF A VECTOR SPACE) *Let $\mathcal{W}$ be a vector subspace (i.e. $\mathcal{W} \subseteq \mathcal{V}$). A polynomial $\phi(\lambda)$ is an annihilating polynomial of $\mathcal{W}$ if, no matter the vector $w \in \mathcal{W}$, then*

$$\phi(A)w = 0.$$

Among all annihilating polynomials, there are some that are minimal with respect to degree. This is true not only for annihiliating polynomials of single vectors but also for annihilating polynomials of subspaces.

Definition 15 (MINIMAL ANNIHILATING POLYNOMIAL) *Among all annihilating polynomials $\phi_i(\lambda)$, $i = 1, 2, \ldots$, there are some for which their degree are all equal and of smallest possible value. They differ from each other by a factor of the field $\mathbb{F}$. Pick one of them, say $\phi_j(\lambda) = \sum_{i=0}^{p} \alpha_{p-i}\lambda^i$, and multiply this polynomial by the inverse of α_0 to get a unique polynomial $\phi(\lambda) = \alpha_0^{-1}\phi_j(\lambda)$ which will be called the minimal annihilating polynomial.*

Remark 1 *We will use at times the abreviation MAP for the (M)inimal (A)nnihilating (P)olynomial.*

2.5 Eigenvectors and Singular Values

Definition 16 (EIGENVECTOR) *Vector $\mathcal{V}$, such that there exists $\lambda_v \in \mathbb{F}$ for which $A\mathcal{V} = \lambda_v \mathcal{V}$.*

2.6 Invariant Subspaces

Definition 17 (INVARIANT SUBSPACE) *An invariant vector subspace for a linear map Φ is a vector subspace $\mathcal{I}$ included in the original vector subspace $\mathcal{V}$ (i.e. $\mathcal{I} \subseteq \mathcal{V}$) such that for any vector $v \in \mathcal{I}$ one has $\Phi(v) \in \mathcal{I}$.*

2.7 Minimal and Characteristic Polynomials

Definition 18 (CHARACTERISTIC POLYNOMIAL) *The characteristic polynomial of matrix A is $\Delta(\lambda)$ given by*

$$\Delta(\lambda) = | A - \lambda I | = 0. \tag{2.8}$$

Lemma 7 () *$\Delta(\lambda)$ is invariant with respect to similarity transform.*

Proof:

$$\begin{aligned}
\bar{A} &= PAP^{-1} \\
|\bar{A} - \lambda I| &= |PAP^{-1} - \lambda I| = |PAP^{-1} - \lambda PP^{-1}| \\
&= |P(A - \lambda I)P^{-1}| = |P||P^{-1}||\lambda I - A| \\
&= |\lambda I - A|
\end{aligned}$$

♠

2.8 Generalized Bezout and Cayley Hamilton

2.8.1 Matrix Polynomials and Multiplication

In general, the matrix product of any two square matrices M_1 and M_2 does not commute, i.e.

$$M_1 M_2 \neq M_2 M_1.$$

However, when the matrices result from polynomials $\phi_1(\lambda)$ and $\phi_2(\lambda)$, namely for a matrix A, say $M_1 = \phi_1(A)$ and $M_2 = \phi_2(A)$, things are different:

Lemma 8 (COMMUTATIVITY OF POLYNOMIAL MATRIX FUNCTIONS) *Given a $n \times n$ matrix A and two polynomials $\phi_1(\lambda)$ and $\phi_2(\lambda)$, the associated polynomial matrix functions commute, i.e.*

$$\phi_1(A)\phi_2(A) = \phi_2(A)\phi_1(A).$$

Proof: Let

$$\begin{aligned}
\phi_1(A) &= \alpha_0 A^n + \alpha_1 A^{n-1} + \ldots + \alpha_n I \\
\phi_2(A) &= \beta_0 A^m + \beta_1 A^{m-1} + \ldots + \beta_m I
\end{aligned}$$

The idea is to isolate the first term of one of the polynomials (for example $\beta_0 A^m$ of the second polynomial), and to factorize that term on the appropriate side of the other polynomial (that is, factorizing $\beta_0 A^m$ on the left in $\phi_1(A)$). This is possible because it amounts to simply regrouping repeated factors $AA \ldots A$ so that A^m appears on the other side (e.g. $A^p A^m = A^m A^p$), which gives in full detail:

$$\begin{aligned}
\phi_1(A)\phi_2(A) &= \left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right)\left(\sum_{j=0}^{m} \beta_j A^{m-i}\right) \\
&= \left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right)\left(\beta_0 A^m + \sum_{j=1}^{m} \beta_j A^{m-j}\right) \\
&= \left(\sum_{i=0}^{n} \alpha_i A^{n-i} \beta_0 A^m\right) + \left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right)\left(\sum_{j=1}^{m} \beta_j A^{m-j}\right) \\
&= \beta_0 A^m \left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right) + \left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right)\left(\sum_{j=1}^{m} \beta_j A^{m-j}\right)
\end{aligned}$$

Pursuing this factorization one step further through isolating $\beta_1 A^{m-1}$ in $\sum_{j=1}^{m} A^{m-j}$, leads to

$$\left(\beta_0 A^m + \beta_1 A^{m-1}\right)\left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right) + \left(\sum_{i=0}^{n} \alpha_i A^{n-i}\right)\left(\sum_{j=2}^{m} \beta_j A^{m-j}\right)$$

which upon iteration gives

$$\left(\sum_{j=1}^{m} \beta_j A^{m-j}\right)\left(\sum_{i=1}^{n} \alpha_i A^{n-i}\right)$$

from which the desired commutativity property results,

$$\phi_1(A)\phi_2(A) = \phi_2(A)\phi_1(A).$$

♠

2.8.2 Matrix Polynomials and Division

We consider in this section not only a polynomial of a single matrix (so that when the variable λ is replaced by a matrix, a matrix is obtained, as in the previouse paragraph) but a matrix whose entries are polynomials in the variable λ.

Definition 19 (Matrix polynomial) *A matrix polynomial $D(\lambda)$ is a $n \times m$ matrix with polynomial entries:*

$$D(\lambda) = (d_{i,j}) \qquad d_{i,j} = \sum_{k=1}^{r} d_{i,j,k}\lambda^{r-k}$$

All coeffeicients belong to the base field, that is,

$$d_{i,j,k} \in \mathbb{F} \quad i = 1, \ldots, n, \; j = 1, \ldots, m, \; k = 1, \ldots, r.$$

Of course, the case of the previous paragraph is just a special case, since by expanding and summing the polynomial matrix function for a given matrix yields a polynomial matrix with a specific structure.

Now, general matrix polynomials can, not only be multiplied by each other (in a non-commutative way when arbitrary polynomial matrices are considered), but also be divided by each other. Because matrix multiplication is not commutative, we need to consider division on the right and division on the left.

Definition 20 (Matrix polynomial right division) *Let $E(\lambda)$ and $F(\lambda)$ be two $n \times n$ matrix poylnomials. Define the right quotient $Q(\lambda)$ and the right remainder $R(\lambda)$ as the two matrix polynomials that satisfy*

$$E(\lambda) = F(\lambda)Q(\lambda) + R(\lambda). \tag{2.9}$$

Definition 21 (Matrix polynomial left division) *Let $E(\lambda)$ and $F(\lambda)$ be two $n \times n$ matrix polynomial. Define the right quotient $Q(\lambda)$ and the right remainder $R(\lambda)$ as the two matrix polynomials that satisfy*

$$E(\lambda) = \hat{Q}(\lambda)F(\lambda) + \hat{R}(\lambda). \tag{2.10}$$

Theorem 1 (Generalized Bezout) *The right remainder $R(\lambda)$ of a $n \times n$ polynomial matrix $E(\lambda)$ divided by the particular polynomial matrix $\lambda I - A$ is given by the matrix $E(A)$. This means that there exits a $Q(\lambda)$ such that*

$$E(\lambda) = Q(\lambda)(\lambda I - A) + E(A).$$

- $B(\lambda)$ is the co-adjoint matrix
  $\Delta(\lambda) = |\lambda I - A| \qquad (\lambda I - A)B(\lambda) = \Delta(\lambda)I$
  $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$
  $R(A) = 0$ zero remainder.

- $\mathcal{I}$ is an invariant subspace.
  $Rv \in \mathcal{I}$
  $\Delta(\lambda)$ cancels A
  $Rv = V_2$ $\quad R \cdot R \cdot V_2 = V_3$ $\quad V_3$ is linearly dependent on V_1 and V_2
  $R(V_1 + V_2) = RV_1 + RV_2$ it is a linear map

2.9 Orthogonalization

2.9.1 Scalar Product

Definition 22 (Scalar product) *A scalar product is a map*

$$\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$$

that satisfies, $\forall\ v_1,\ v_2,\ v_3 \in \mathcal{V}$, $\forall\ \alpha, \beta \in \mathbb{F}$:

1. $\langle v, v \rangle > 0 \qquad v \neq 0$
2. $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$
3. $\langle \alpha v_1 + \beta v_2, v_3 \rangle = \alpha \langle v_1, v_3 \rangle + \beta \langle v_2, v_3 \rangle$

Definition 23 (Norm) *We suppose that* $\mathbb{F}$ *is a valuation field with valuation* $|\cdot| : \mathbb{F} \to \mathbb{R}$. *This means that the map* $|\cdot| : \mathbb{F} \to \mathbb{R}$ *satisfies the follwoing three axioms* $\forall \alpha, \beta \in \mathbb{F}$

(i) $|\alpha| \geq 0$

(ii) $|\alpha\beta| = |\alpha|\ |\beta|$

(iii) $|\alpha + \beta| \leq |\alpha| + |\beta|$

A map $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ *is a norm when the following identities are satisfied, for all* v, v_1, v_2, α_1, α_2.

1. *Homogeneity:* $\|\alpha v\| = |\alpha| \|v\|$.
2. *Positivity:* $\|v\| > 0 \quad \forall v \neq 0$, *and,* $\|v\| = 0 \Leftrightarrow v = 0$.
3. *Triangle inequality:* $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$.

Lemma 9 *The Euclidean norm*

$$\|v\| := \sqrt{\langle v, v\rangle}$$

obeys the definition of a norm according to Def. 23.

The proof is a simple check of the corresponding axioms.

Example 1 *Let* $v = \sum_{i=1}^{n} \alpha_i v_i$ *where a basis* $B = \{v_1, \ldots, v_n\}$ *has been chosen which need not be orthonormal.*

$$\begin{aligned} \|v\|_2 &:= \sum_{i=1}^{n} |\alpha_i|^2 \\ \|v\|_1 &:= \sum_{i=1}^{n} |\alpha_i| \\ \|v\|_\infty &:= \max_{i \in 1 \ldots n} |\alpha_i| \end{aligned}$$

Remark 2 *When the basis is orthonormal the 2-norm* $\|\cdot\|_2$ *becomes the Euclidean norm* $\|\cdot\|$.

This remark leads us to the definition of an orthogonal basis and orthonormal basis.

Definition 24 (Orthogonal vectors w.r.t. a scalar product) *Let* $\langle\cdot,\cdot\rangle : \mathcal{V} \to \mathbb{R}$ *be a scalar product. Two vectors* v_1 *and* v_2*, both belonging to* $\mathcal{V}$ *will be called orthogonal if and only if*

$$\langle v_1, v_2\rangle = 0.$$

Definition 25 (Orthonormal basis)

Definition 26 (Standard scalar product) *Suppose a vector is given through its column definition as in Def. 9. The standard scalar product is defined as* $\langle v_1, v_2\rangle = v_1^T v_2$ *where on the left-hand-side,* v_1 *and* v_2 *are elements of* $\mathcal{V}$ *written in the standard way, whereas on the left-hand-side,* v_1 *and* v_2 *represent their respective representation with respect to an orthogonal basis* $B = \{e_1, e_2, \ldots, e_n\}$.

Let $B = \{v_1, v_2, \ldots, v_n\}$ be a basis of a vector space. We

Definition 27 (General linear map) *For any vector $v : v, Av, ..., A^{p-1}v$ there exist p such that $\{v, Av, ...A^{p-1}v\}$ is a basis.*

Definition 28 (Annihilating polynomial of a vector (a subspace)) *is a λ-polynomial such that (a subspace) $\Psi(\lambda)$: $\Psi(A)v = 0$.*

$$A^p v = AA^{p-1}v = -\alpha_1 Ap - 1v - \alpha_2 A^{p-2}v - ... - \alpha_p Iv$$

$$\exists\, \alpha_1, \alpha_2, ..., \alpha_{p-1} \qquad \lambda^p + \alpha_1 \lambda^{p-1} v + ... + \alpha_{p-1}$$

Definition 29 (Cyclic invariant subspace) *is defined as*

$$\mathcal{I} = span\{v, Av, ..., A^{p-1}v\}$$

Theorem 2 *Any vector space $\mathcal{V}$ can be split as a direct sum of cyclic subspaces*

$$\mathcal{V} = \mathcal{I}_1 \oplus \mathcal{I}_2 \oplus ... \oplus \mathcal{I}_z$$

Let $V \in \mathcal{I}_1 \oplus \mathcal{I}_2$ (subspace of $\mathcal{V}$) and exists two V_1 and V_2 such that $V_1 \in \mathcal{I}_1$ and $V_2 \in \mathcal{I}_2$, then $V = V_1 + V_2$

Exercise Suppose that we know an annihilating polynomial $\Psi(\lambda)$ of the whole vector space $\mathcal{V}$, where A is given.
$\forall v \in \mathcal{V} \quad \Psi(A)v = 0$.

Definition 30 (Minimal polynomial (vector, subspace, matrix)) *It is the lowest degree polynomial which is an annihilating polynomial.*

$\Delta(\lambda)$, characteristic polynomial might not be the minimal polynomial of A.

Then, it should be shown that there exists an unique vector $\mathcal{C}$ which has the same minimal annihilating polynomial.

2.9.2 Factor space

Let $\mathcal{I} = \{v, Av, ..., A^{p-1}v\}$ be an invariant set.

Definition 31 (ABSOLUTE EQUALITY IN VECTOR SPACE) *Vectors* v_1 *and* v_2 *are absolutely equal if* $v_1 = v_2$*, meaning that* $v_1 - v_2 = 0$*.*

Definition 32 (RELATIVE (MOD $\mathcal{I}$) EQUALITY IN VECTOR SPACE) *Vectors* v_1 *and* v_2 *are relatively equal (by mod* $\mathcal{I}$*) if* $v_1 \equiv v_2$ *and iff* $\Leftrightarrow \exists v \in \mathcal{I}$, $v_1 = v_2 + v$, $v_1 - v_2 = v$*.*

Definition 33 (EQUIVALENT CLASSES OF VECTORS) *Vectors* v_1 *and* v_2 *are absolutely equal if* $v_1 = v_2$*, meaning that* $v_1 - v_2 = 0$*.*

Equivalent classes of vectors $x \equiv y$ set $\{y \,|\, \exists v \in \mathcal{I}, x = y + v\}$
1. reflexivity $x \equiv y \Rightarrow y \equiv x$
2. transitivity $x \equiv y$ $y \equiv z \Rightarrow x \equiv z$
3. simetry $x \equiv x$

Exercise Check the axioms of the vector space $\{y \,|\, \exists v \in \mathcal{I}, x = y + v\}$ if $\bar{y}$ is one of its elements.

Chapter 3

Coprimeness and Annihilating Polynomials

Recap:

- Minimal annihilating polynomial (MAP): ϕ, such that for invariant space $I_1 \subset v$ and for $\forall v, A\,v \in I_1$, it holds that $\phi(A)\,v = 0$.
- Degree of the MAP must be less than or equal to the degree of the characteristic polynomial.

3.1 Coprime polynomials

Basic integer coprimeness for given two integers (e.g. 3 and 7): $\exists a, b \in \mathbb{Z}$: $3a+7b = 1$. This holds with $a = 5$ and $b = -2$, i.e. $3{\cdot}5+7{\cdot}(-2) = 15-14 = 1$.

Definition 34 (Coprimeness polynomials) *Polynomials Ψ_1 and Ψ_2 are coprime iff $a(\lambda)\Psi_1(\lambda) + b(\lambda)\Psi(\lambda) = 1$.*

Definition 35 (Annihilating Polynomial (AP) of a Vector Space) *Let $\mathbb{R}^n$ be the general state space and consider an arbitrary subspace contained in it $\mathcal{V} \subseteq \mathbb{R}^n$.*

Let a matrix A be given which is $\mathbb{R}^{n\times n}$.

Among all polynomials $\psi(\lambda)$ consider those that are such that when λ is replaced by the matrix they annihilate all vectors that belong to the subspace $\mathcal{V}$, or in other words,

$$\{v \in \mathbb{R}^n | \psi(A)v = 0\}$$

Definition 36 (Minimal Annihilating Polynomial) *Among all AP, the one of least order will be called the minimal annihilating polynomial (written MAP).*

3.2 Invariant Subspaces and Cyclic Invariant Subspaces

Definition 37 (invariant subspace)
Let $\mathcal{I}$ be a vector subspace of $\mathbb{R}^n$ and A an $n \times n$ matrix. The subspace $\mathcal{I}$ is said to be invariant if for any $v \in \mathcal{I}$, then $Av \in \mathcal{I}$.

Definition 38 (cyclic subspace)
Let a matrix A of dimension $n \times n$ be given. A cyclic invariant subspace is an invariant vector subspace $\mathcal{I}$ generated by a single vector. Let v be one of its generator. The dimension of the cyclic subspace is the maximum integer p such that the vectors v, Av, ..., A^{p-1} constitute a family of linearly independent vectors. The integer p is independent of the chosen generator.

3.3 Decomposition into Cyclic Subspaces

Lemma 10 (vector sharing same MAP as whole space — single irreducible factor case)
Let A be a given $n \times n$ matrix. Let V be a vector subspace having MAP (w.r.t. A) which contains a single irreducible factor, i.e.

$$\psi(\lambda) = (\phi(\lambda))^l$$

where $l \in \mathbb{N}$ is the multiplicity of that factor. Then there exists at least one vector $v \in \mathcal{V}$ sharing the same MAP. The MAP (w.r.t. A) of v is $(\phi(\lambda))^l$.

Proof: Consider in succession each canonical basis vector e_i, $i = 1, \ldots, n$. For each i, construct the maximal linearly independent familiy of vectors

$$\{e_i, Ae_i, \ldots, A^{l_i-1}e_i\}.$$

To such a family of vectors, there correponds the MAP

$$\psi_i(\lambda) = (\phi(\lambda))^{l_i}$$

Notice that these MAPs have to share the same irreducible factor as the MAP of the whole space (since this latter polynomial is considered of having a single irreducible factor). The multiplicity l_i might be smaller than l. Then consider the least common mulitiple (l.c.m)

$$\text{l.c.m.}_{i=1,\ldots,n}\{(\phi_i(\lambda))^{l_i}\}$$

This means that

$$l = \max{}_{i=1,\ldots,n}\{l_i\}$$

and the the sought vector e is the basis vector with index that matches this attained maximum, i.e.

$$e := e_{\text{argmax }_{i=1,\ldots n}\{l_i\}}$$

♠

Theorem 3 (COPRIME FACTORISATION OF THE MAP OF THE WHOLE SPACE INDUCE DIRECT SUM DECOMPOSITION INTO INVARIANT SUBSPACES) *Let $\psi(\lambda)$ be the MAP of the whole vector space $\mathbb{R}^n$. Suppose that $\psi(\lambda)$ can be factorised into two coprime factors, i.e.*

$$\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda)$$

then the whole space splits into a direct sum of two invariant subspaces

$$\mathbb{R}^n = \mathcal{I}_1 \oplus \mathcal{I}_2$$

for which $\psi_1(\lambda)$ is MAP of $\mathcal{I}_1$ and ψ_2 is MAP of $\mathcal{I}_2$.

Proof: Since the polynomials $\psi_1(\lambda)$ and $\psi_2(\lambda)$ are coprime one can find two polynomials $a(\lambda)$ and $b(\lambda)$ so that the Bezout indentity

$$1 = a(\lambda)\psi_1(\lambda) + b(\lambda)\psi_2(\lambda) \tag{3.1}$$

holds which becomes after substituting the matrix A for the variable λ, and after postmultiplying by an arbitrary vector x

$$x = a(A)\psi_1(A)x + b(A)\psi_2(A)x$$

Hence, it is possible to define two new vectors

$$x_1 := a(A)\psi_1(A)x$$
$$x_2 := b(A)\psi_2(A)x$$

so that

$$x = x_1 + x_2$$

But then

$$\begin{aligned}
\psi_2(A)x_1 &= \psi_2(A)a(A)\psi_1(A)x \\
&= a(A)\psi_1(A)\psi_2(A)x \qquad (3.2)\\
&= a(A)\psi(A)x \qquad (3.3)\\
&= a(A)0 \qquad (3.4)\\
&= 0
\end{aligned}$$

where commutation of polynomial maps has been used in (3.2) so that the polynomial map $\psi(A)$ stands out in (3.3) which is the map associated with the minimal annihilating polynomial $\psi(\lambda)$ of the whole space $\mathbb{R}^n$, and hence it is also the MAP of the vector x, so that (3.4) holds and therefore leads to $\psi_2(\lambda)$ being an annihilating polynomial of x_1 (in a compact form, ψ_2 is AP of x_1). In a similar fashion, one shows that $\psi_1(\lambda)$ is AP of x_2. Hence all vectors x (being annihilated by the polynomial $\psi(\lambda)$) split into two components x_1 and x_2 according to the above procedure. One can then notice that the vector x_1 belongs to an invariant subspace (which will then be written $\mathcal{I}_1$ and defined as $\mathcal{I}_1 = \{v | \psi_1(A)v = 0\}$). Indeed, take any x_1 such that $\psi_1(A)x_1 = 0$. Multiplying by A and using the commutation of A with the polynomial map $\psi_1(A)$ gives

$$\psi_1(A)Ax_1 = A\psi_1(A)x_1 = A0 = 0$$

and hence $\psi_1(\lambda)$ is AP of Ax_1 and hence $Ax_1 \in \mathcal{I}_1$ so that $\mathcal{I}_1$ is an invariant subspace (w.r.t. A). Similarly, one shows that $\mathcal{I}_2$ is an invariant subspace (w.r.t. A). Otherwise stated, one has

$$\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2$$

Let us now show that

$$\mathcal{I} = \mathcal{I}_1 \oplus \mathcal{I}_2$$

by displaying that the only vector in common with $\mathcal{I}_1$ and $\mathcal{I}_2$ is the zero vector 0. We will exhibit a contradiction: Suppose then that there exists a vector $v \neq 0$ such that v belongs to both $\mathcal{I}_1$ and $\mathcal{I}_2$, that is,

$$\psi_1(A)v = \psi_2(A)v = 0$$

Now, postmultiplying the Bezout identity (3.1) by v (after substituting λ for A) gives

$$v = a(A)\psi_1(A)v + b(A)\psi_2(A)v = 0 + 0 = 0 \tag{3.5}$$

This exhibits the sought contradiction since $v \neq 0$ was initially assumed.

What remains to be proven is the fact that not only are $\psi_1(\lambda)$ and $\psi_2(\lambda)$ the APs of the respective supspaces, but they are in fact the minimal ones (i.e. they are MAPs). To this effect, consider an arbitrary AP $\tilde{\psi}_1$ of $\mathcal{I}_1$. Using the Bezout identity,

$$\tilde{\psi}_1(A)\psi_2(A)x = \psi_2(A)\tilde{\psi}_1(A)x_1 + \tilde{\psi}_1\psi_2(A)x_2 = 0$$

x being an arbitrary vector, $\tilde{\psi}_1(\lambda)\psi_2(\lambda)$ is an annihilating polynomial for $\mathbb{R}$ and is therefore divisible by $\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda)$ which means that $\tilde{\psi}_1(\lambda)$ is divisible by $\psi_1(\lambda)$ and thus $\psi(\lambda)$ is the MAP of $\mathcal{I}_1$. ♠

Lemma 11 (THE PRODUCT OF COPRIME POLYNOMIALS IS THE MAP OF THE SUM OF THE VECTORS ASSOCIATED WITH THE POLYNOMIALS)

- *Let A be a given $n \times n$ matrix.*
- *Let v_1 and v_2 be arbitrary vectors in $\mathbb{R}^n$.*
- *Let $\psi_1(\lambda)$ be a MAP (w.r.t. A) of the vector v_1.*
- *Let $\psi_2(\lambda)$ be a MAP (w.r.t. A) of the vector v_2.*
- *Suppose that $\psi_1(\lambda)$ and $\psi_2(\lambda)$ are coprime*

$$\Rightarrow$$

$$e := e_1 + e_2$$

has

$$\psi(\lambda) := \psi_1(\lambda)\psi_2(\lambda)$$

as MAP (w.r.t A).

Proof: Let $\chi(\lambda)$ be an AP of $e = e_1 + e_2$.

$$\begin{aligned}
\psi_2(A)\chi(A)e_1 &= \\
&= \psi_2(A)\chi(A)(e - e_2) \\
&= \psi_2(A)\chi(A)e - \chi(A)\psi_2(A)e_2 \\
&= 0 - 0 = 0
\end{aligned}$$

This means that $\psi_2(\lambda)\chi(\lambda)$ is AP of e_1 Since the AP is divisible by the MAP, the product $\psi_2(\lambda)\chi(\lambda)$ is divisible by $\psi_1(\lambda)$. In a similar fashion we establish that $\psi_1(\lambda)\chi(\lambda)$ is AP of e_2. Since the AP is divisible by the MAP, this means that the product $\psi_1(\lambda)\chi(\lambda)$ is divisible by $\psi_2(\lambda)$.

We thus come to the conclusion that because $\psi_1(\lambda)$ and $\psi_2(\lambda)$ are coprime, $\chi(\lambda)$ has to be divisible by $\psi_1(\lambda)$ and $\chi(\lambda)$ has to be divisible by $\psi_2(\lambda)$. But this also signifies that $\chi(\lambda)$ is divisible by the product $\psi_1(\lambda)\psi_2(\lambda)$. Because $\chi(\lambda)$ was supposed to be an arbitrary AP of e (not necessarily the MAP of e), we come to the conclusion that every AP of e is divisible by $\psi_(\lambda)\psi_2(\lambda)$. Therefore, $\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda)$ has to be the MAP of $e = e_1 + e_2$. ♠

Theorem 4 (THERE ALWAYS EXIST A VECTOR SHARING THE SAME MAP AS THE WHOLE SPACE)
Let A designate a matrix $\mathbb{R}^n \times \mathbb{R}^n$ and let f_A be the linear map $\mathbb{R}^n \to \mathbb{R}^n$ defined through A. Given a vector space $\mathcal{V}$ with MAP $\psi(\lambda)$ and a linear map f_A defined through a matrix A, then there always exist at least one vector $v \in \mathcal{V}$ sharing the same MAP as the whole space, that is, $\exists v \in \mathcal{V}$, such that $\psi(\lambda)$ is also the MAP of v.

Proof: Let $\psi(\lambda)$ be the MAP (w.r.t. A) of the whole space $\mathbb{R}^n$. Factorise $\psi(\lambda)$ into irreducible factors

$$\psi(\lambda) = (\phi_1(\lambda))^{l_1}(\phi_2(\lambda))^{l_2} \cdots (\phi_p(\lambda))^{l_p}$$

and then proceed as in the proof of Lemma 10 by displaying for each irreducible factor $(\phi(\lambda))^{l_i}$ the corresponding annihilated vector v_i. Then applying recursively Lemma 11 ♠

Theorem 5 (DECOMPOSITION INTO CYCLIC INVARIANT SUBSPACES WITH MAPS THAT DIVIDE EACH PREDECESSOR)
Let A be a given $n \times n$ matrix. The n dimensional vector space $\mathbb{R}^n$ splits into a direct sum of invariant cyclic subspaces

$$\mathbb{R}^n = \mathcal{I}_1 \oplus \mathcal{I}_2 \oplus \ldots \oplus \mathcal{I}_r$$

with each invariant subspace $\mathcal{I}_i$, $(i = 1, \ldots, r)$, having an associated MAP $\psi_i(\lambda)$ with the property that $\psi_{i+1}(\lambda)$ divides $\psi_i(\lambda)$.

Proof: The whole space $\mathbb{R}^n$ has a MAP. It will be labeled $\psi(\lambda)$. Theorem 4 guarantees the existence of a vector $e \in \mathbb{R}^n$ having the same $\psi(\lambda)$ as MAP. This implies the existence of a cyclic invariant subspace

$$\mathcal{I}_1 := \{e, Ae, \ldots, A^{p-1}e\}$$

for a given p. Let us suppose that $p < n$, for otherwise the whole space $\mathbb{R}^n$ is a single cyclic invariant subspace, and the theorem holds. In the case that $p < n$ we can consider the relative minimal annihilating map modulo $\mathcal{I}_1$ of the whole space $\mathbb{R}^n$, which we will denote as $\psi_2(\lambda)$ and use the acronym RMAP (mod $\mathcal{I}_1$) of the whole space $\mathbb{R}^n$. It is plain to show that the RMAP has to divide the MAP of a given space. This means that $\psi_2(\lambda)$ divides $\psi_1(\lambda) = \psi(\lambda)$, i.e. $\exists \chi(\lambda)$ such that $\psi_1(\lambda) = \psi_2(\lambda)\chi(\lambda)$. Because ψ_2 is the RMAP (mod $\mathcal{I}_1$) of the whole space, there exists a vector $g^* \in \mathbb{R}^n$ sharing the same RMAP (mod $\mathcal{I}_1$), (just adapt Theorem 4 to the relative case), i.e. for which

$$\psi_2(A)g^* \equiv 0 \qquad (\text{mod } \mathcal{I}_1)$$

Translating this equivalence into an equality, there exists $\gamma(\lambda)$ such that

$$\psi_2(A)g^* = \gamma(A)e$$

Let us apply $\chi(A)$ on both sides of the last equation

$$\psi_1(A)g^* = \chi(A)\psi_2(A)g^* = \chi(A)\gamma(A)e$$

Now since $\psi_1(\lambda)$ is the MAP of the whole space, it is an AP of g^* and therefore

$$\chi(A)\psi_2(A)g^* = 0$$

which means that $\chi(\lambda)\psi_2(\lambda)$ is AP of e. It is therefore divisible by $\psi_1(\lambda)$ (since $\psi_1(\lambda)$ is MAP of e). This then means that because $\psi_1(\lambda) = \chi(\lambda)\psi_2(\lambda)$, the polynomial $\gamma(\lambda)$ must be divisible by $\psi_2(\lambda)$, i.e.

$$\gamma(\lambda) = \beta(\lambda)\psi_2(\lambda).$$

Hence,

$$\psi_2(A)g^* = \gamma(A)e = \beta(A)\psi_2(A)e$$

which leads to

$$\psi_2(A)(g^* - \beta(A)e) = 0.$$

Therefore $\psi_2(\lambda)$ is MAP of the newly defined element

$$g := g^* - \beta(A)e.$$

But $\psi_2(\lambda)$ is also the RMAP of g (mod $\mathcal{I}_1$) (since g is equal to g^* modulo an element of $\mathcal{I}_1$, namely $\beta(A)e$). Since g has MAP $\psi_2(\lambda)$, the space

$$\{g, Ag, \ldots, A^{m-1}g\} \tag{3.6}$$

is cyclic. Since $\psi_2(\lambda)$ is RMAP (mod $\mathcal{I}_1$) of g, it follows that the family (3.6) has to be linearly independent from

$$\{e, Ae, \ldots, A^{p-1}e\}.$$

If $n = m + p$ then

$$\mathbb{R}^n = \mathcal{I}_1 \oplus \mathcal{I}_2.$$

If not, we continue the process by considering the RMAP of the whole space $\mathbb{R}^n$ (mod $\mathcal{I}_1 \oplus \mathcal{I}_2$) and we finally secure a sequence of polynomials

$$\psi_1(\lambda), \psi_2(\lambda), \ldots, \psi_r(\lambda)$$

with $\psi_{i+1}(\lambda)$ dividing $\psi_i(\lambda)$ $i = 1, \ldots, r-1$. These polynomials are the MAPs of associated invariant subspaces that constitute a direct sum decomposition of the whole space:

$$\mathbb{R}^n = \mathcal{I}_1 \oplus \mathcal{I}_2 \oplus \cdots \oplus \mathcal{I}_r$$

♠

Chapter 4

Smith Form, Unimodalar Matrices and Invariant Polynomials

• Assume we have a vector space and can split it into a direc sum of invariant subspaces: $\mathcal{V} = \mathcal{I}_1 \oplus \mathcal{I}_2 \oplus ... \oplus \mathcal{I}_r$
Up to a certain index, the annihilating polynomials will divide each others: Ψ_{k+1} divides Ψ_k, where $k = 1, 2, ..., r-1$.
The question of interest is how to easily compute these polynomials Ψ_k of $\mathcal{I}_k$.

• Consider a matrix A and its determinant, $D_n(\lambda)$ of order n.
• Method to compute the minimal polynomial of $\mathcal{I}_k$:
1° choose all minors of order $n-1$ of matrix A
2° compute the greatest common divisor of all minors of order $n-1$ ($D_{n-1}(\lambda)$)
3° choose all minors of order $n-2$
4° ...
5° repeat the steps until the end.

Definition 39 (Invariant polynomials of A)

$$i_1 = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)} \qquad i_2 = \frac{D_{n-2}(\lambda)}{D_{n-3}(\lambda)} \qquad ... \qquad i_{n-1} = D_1$$

• They are invariant to the similarity transform: PAP^{-1} has the same i_k for $k = 1, 2, ...$
• We can also show consecutive division:

i_{k+1} divides i_k
$\div$ can be show by the Cauchy-Binet formula: $|AB| = |A||B|$

4.1 Smith Form

• It is possible get invariant polynomials from the Smith Form of the matrix, which can be expressed via unimodular matrices:

$$U_1(sI - A)U_2 = \begin{vmatrix} d_1(s) & 0 & \dots & 0 \\ 0 & d_2(s) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n(s) \end{vmatrix}.$$

• The idea is to find these matrices U_1 and U_2, diagonalize matrix and re-get invariant polynomial i_k.
$|U_1| \in \mathcal{F}$ and $|U_2| \in \mathcal{F}$, where $\mathcal{F}$ is a field.
• Needed building blocks:

$$T_1 = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} \qquad T_2 = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

$$T_3 = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \alpha(s) & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

$\div$ T_1 is identity matrix multiplied with a constant $c \in \mathcal{F}$.
$\div$ T_2 is identity matrix with swap of i-th row or j-th column.
$\div$ T_3 is identity matrix with an inserted polynomial.

• Construction:
$U_1 = T_{k_1} \cdot T_{k_2} \cdot ... \cdot T_{k_m}$, where $k_j \in \{1,2,3\}$ and $j = 1,2,...,m$
$U_2 = S_{k_1} \cdot S_{k_2} \cdot ... \cdot S_{k_m}$, where $S_1 = T_1$, $S_2 = T_2$, $S_3 = T_3^T$
• Another way to generate the Smith Form:

$$|sI - A| = \begin{vmatrix} P_{11}(s) & P_{12}(s) & P_{13}(s) & \dots & P_{1n}(s) \\ P_{21}(s) & P_{22}(s) & P_{23}(s) & \dots & P_{2n}(s) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{j1}(s) & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{n1}(s) & P_{n2}(s) & P_{n3}(s) & \dots & P_{nn}(s) \end{vmatrix}$$

$\div$ P_{ji} is a polynomial of maximal degree 1.

1° Choose the lowest degree element in s (e.g. in the 1st column and j-th row, P_{j1}).
2° Swap the rows to switch P_{j1} and P_{11}.
3° Divide P_{k1} by P_{j1} (and we can do it because it has the lowest degree)
4° $P_{k1}(s) = \alpha_k(s)P_{j1}(s) + \beta_k(s)$.
$\div$ α_k is from T_3 and β_k is a reminder
5° Replace P_{n1} terms by β terms.
6° ...
5° Repeat the whole process until matrix is diagonal and do the same for the columns.

• The choice will not disturb the fact that at the end we obtain an unique solution.

4.1.1 Elementary divisor

• Consider invariant polynomials $i_1(s), i_2(s), ..., i_r(s)$, where $r \leq n$.
Each of them can be factorized in the (algebraic) field $\mathcal{F}$:
$i_k(s) = (s - \lambda_{k1})^{p_{k1}} \cdot (s - \lambda_{k2})^{p_{k2}} \cdot ... \cdot (s - \lambda_{kl})^{p_{kl}}$, where $\lambda_l \in \mathbb{C}$

Definition 40 (Elementary divisor) *Each element $(s - \lambda)^p$ is an elementary divisor of A.*

4.1.2 Dynamical Systems

• $\dot{x} = Ax + Bu$ $\qquad$ $y = Cx + Du$
• If we change coordinates by $z = Px$, we have $\dot{z} = P\dot{x}$, where P is not

singular.
$x = P^{-1}z \qquad \dot{z} = P\dot{x} = PAx + PBu = PAP^{-1}z + PBu$
$y = CP^{-1}z + Du$

- SIMILARITY TRANSFORM FOR STATE-SPACE SYSTEM

$\bar{A} = PAP^{-1} \qquad \bar{B} = PB \qquad \bar{C} = CP^{-1} \qquad \bar{D} = D$

- Standard solution of a system:

$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$
$\frac{d}{dt}(e^{-At}x) = -Ae^{-At}x + e^{-At}\dot{x}$

- Exponential of the matrix:

$x \in \mathbb{R} \qquad \dot{x} = ax \qquad a \in \mathbb{R} \qquad x_0e^{at} = x$

$\frac{d}{dt}(e^{at}x_0) = ae^{at}x_0 - ax - \dot{x}$
$\div\ e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + ... \qquad x = at$
$\div\ e^{at} = 1 + at + \frac{1}{2}(at)^2 + \frac{1}{3!}(at)^3 + ...$

$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \mathcal{O}(A^4)$

Theorem 6 (CAYLEY-HAMILTON THEOREM) *Matrix A satisfies its characteristic polynomial $|\lambda I - A| = \Delta(\lambda)$ in λ.*

- $deg(\Delta(\lambda)) \leq n$

$\Delta A = 0$, we can express $A^n = \sum_{j=0}^{n-1} \alpha_i A^j$, $\alpha_i \in \mathcal{F}$
This means that in a finite sum of e^A we can take all the high degrees and express them with lower ones.
$e^A = \sum_{j=0}^{n-1} \beta_j A^j$
Though, if we have some other function, maybe it won't be easy to express it in the series.

- Suppose we have polynomials $f(\lambda) \neq g(\lambda)$

The question is if this can be true: $f(A) = g(A) \qquad f(A) - g(A) = 0$

$f(\lambda) - g(\lambda) = \Psi(\lambda)$ is an annihilating polynomial of A.
$\Psi(\lambda) = m(\lambda)\Phi(\lambda)$, where $m(\lambda)$ is minimal annihilating polynomial of A

$g(\lambda) = e^{\lambda} \qquad f(\lambda) = \sum_{j=0}^{n-1} \beta_i A^j$

• $\lambda_1, \lambda_2, ..., \lambda_r$ are roots of $m(\lambda)$
$f(\lambda_j) = g(\lambda_k)$ totally independent of matrix

• $\sum_{j=0}^{n-1} \beta_j \lambda_k^j = e^{\lambda_k}$, r equalities

• $m(\lambda) = (\lambda - \lambda_1)^{p_1}(\lambda - \lambda_2)^{p_2}$
$\frac{d}{d\lambda} m(\lambda) = p_1(\lambda - \lambda_1)^{p_1-1}(\lambda - \lambda_2)^{p_2} + p_2(\lambda - \lambda_2)^{p_2-1}(\lambda - \lambda_1)^{p_1}$

$m'(\lambda_k) = 0$

• $\dot{x} = Ax + Bu$
$-Ae^{-At}x + e^{-At}(Ax + Bu) = \frac{d}{dt}(e^{-At}x) = e^{-At}Bu$
$\int_0^{t_1} e^{-A\tau} Bu(\tau) d\tau = e^{-At_1}x(t_1) - e^{-A\cdot 0}x(0)$
$x(t_1) = e^{At_1}x_0 + e^{At_1}\int_0^{t_1} e^{-A\tau}Bu(\tau)d\tau$

Chapter 5

Stability

- Open-loop stability
- Closed-loop stability
- Asymptotic stability
- Bounded-input-bounded-output (BIBO)
- Lyapunov stability

5.1 BIBO in SISO systems

Definition 41 (Bounded-input-bounded-output)
$\forall u, \exists M \in \mathbb{R}^+ : ||u|| < M$, *where M mesures thr size of a system.*

- $||u(\cdot)|| = \int_{-\infty}^{\infty} |u(\tau)| d\tau = || \, ||_1$, norm-1, L_1
- $||u(\cdot)|| = \int_{-\infty}^{\infty} (u(\tau))^2 d\tau = || \, ||_2$, norm-2, L_2

Theorem 7 $\forall u, \exists m : ||y|| < m \Leftrightarrow$ *the system is BIBO stable.*

Proof: Consider an impulse response $g(t)$.
$y(t) = \int_{-\infty}^{\infty} u(\tau) g(t-\tau) d\tau = \int_{-\infty}^{\infty} g(\tau) u(t-\tau) d\tau \qquad y = g * u$
This means that $g(t)$ is independent of $u(t)$, which is not true for nonlinear systems.

$\exists C \in \mathbb{R}^+ : ||g(\cdot)|| < C$, because we want "bounded" $g(t)$

$$\int_{-\infty}^{\infty} ||g(\tau)|| d\tau < C$$

Proof for $\Rightarrow$:

$y = \int_{-\infty}^{\infty} u(t-\tau)g(\tau)d\tau$, and we choose M, such that $||u|| < M$

$|y| = |\int u(\tau)g(t-\tau)d\tau|$

$\diamond$ $|a \bullet b| \leq |a| \bullet |b|$ $\qquad |\int a(\cdot) \bullet b(\cdot)| < \int |a(\cdot)| \bullet |b(\cdot)|$

$|y| \leq ||u|| \cdot ||g||$ $\qquad |y| \leq M \cdot C$

Proof for $\Leftarrow$:

Assume contraposition: $g(\cdot)$ is not bounded.

$\forall N > 0, \exists t$ such that $\int_0^t |g(\tau)|d\tau > N$

$y(t) = \int_0^t g(\tau)u(t-\tau)d\tau$

Particular bounded input:

$$u(t-\tau) = \begin{cases} +1 & g(\tau) > 0; \\ -1 & g(\tau) < 0; \end{cases}$$

$\bar{y}(t) = \int_0^t |g(\tau)|d\tau > N$, which is a contradiction. ♠

- Consider the following system:

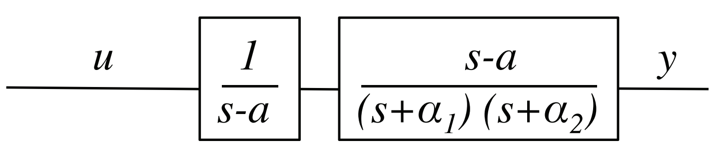


Figure 5.1: $a, \alpha_1, \alpha_2 > 0$

Transfer function of the system: $\dfrac{1}{s-a} \cdot \dfrac{s-a}{(s+\alpha_1)(s+\alpha_2)}$

If we perform canceling of the terms $(s-a)$:
$\dfrac{1}{\cancel{s-a}} \cdot \dfrac{\cancel{s-a}}{(s+\alpha_1)(s+\alpha_2)} = \dfrac{1}{(s+\alpha_1)(s+\alpha_2)}$

A possible scenario is that there can be some initial conditions that will cause the "explosion" of the system after the first block and than the zero in a might or might not be able to "calm" it down.

In other words, one should take care of *iternal states*.

5.2 Lyapunov stability

Lypunov stability is essential state stability, where the initial conditions act as inputs.

- $\dot{x} = Ax + Bu \qquad y = Cx + Du$

If we want to check stability for the states, than there is no need to know the output.

- Linear systems case:

Check stability for states with $u = 0$.
$\dot{x} = Ax \qquad \bar{x} = 0$ is equilibrium
Initial conditions $x(0) = x_0$

$\forall R < 0, \exists r : ||x_0|| < r$
Solution that depends on initial conditions: $\dfrac{d}{dt}\chi(x_0, t) = A\chi(x_0, t)$
$||\chi(x_0, t)|| < R$

5.3 Asymptotic Lyapunov stability

- $\lim\limits_{t\to\infty} \chi(x_0, t) = 0$

- $V(x) = x^T X = x^T I x$

generalize with matrix P: $x^T P x > 0$, for $\forall x \neq 0$

Definition 42 () *A positive definite matrix P is a matrix such that*

$$x^T Px > 0, \forall x \neq 0.$$

- $V = x^T Px$ is Lyapunov function.

- $\dot{V} = \dot{x}^T Px + x^T P\dot{x} = x^T A^T Px + x^T PAx \qquad \dot{x} = Ax$.

$x^T(A^T P + PA)x < 0 \qquad A^T P + PA < 0$

$\diamond$ $A^T P + PA = -Q \qquad Q > 0 \qquad x^T Qx > 0$

Definition 43 (Lyapunov stability) $\forall Q > 0, \exists P > 0$ *such that*

$$A^T P + PA = -Q.$$

Remark: Use MATLAB© function `lyap(`A^T, Q`) = `P.

- This can be also linked to the real parts of all eigenvalues of A being negative: $Re\{Eig(\lambda)\} < 0$.

$P = \int_0^\infty e^{A^T\tau} Q e^{A\tau} d\tau$

$$\begin{aligned}
A^T P + PA &= A^T \int_0^\infty e^{A^T\tau} Q e^{A\tau} d\tau + \int_0^\infty e^{A^T\tau} Q e^{A\tau} A d\tau \\
&= \int_0^\infty (A^T e^{A^T\tau} Q e^{A\tau} + e^{A^T\tau} Q e^{A\tau} A) d\tau \\
&= \int_0^\infty \frac{d}{d\tau}(e^{A^T\tau} Q e^{A\tau} d\tau) \\
&= e^{A^T\tau} Q e^{A\tau}|_0^\infty = 0 - Q = -Q \\
A^T P + PA &= -Q
\end{aligned}$$

- $A^T P + PA = -Q \Rightarrow Re\{Eig(\lambda)\} < 0$

v is an eigenvector of A: $Av = \lambda v$

$v^T = [v_1 \ v_2 \ ... \ v_n]$

$v^H = [v_1^* \ v_2^* \ ... \ v_n^*]$, complex conjugate of v ($v_1 = a + jb, v_1^* = a - jb$)

$v^H A^H = \lambda^H v^H \qquad v^T A^T = \lambda^T v^T$
$-v^H Q v = -v^T Q v = -v^T (A^T P + PA) v = \lambda^H v^T P v + v^T P v \lambda =$
$= (\lambda^H + \lambda) v^H P v$
$\quad \curvearrowright \;\; 2Re\{\lambda\} < 0 v^T P v < 0$
$2Re\{\lambda\} v^T P v < 0$

Since Q is positive definite, P is positive definite and it implicates $Re\{\lambda\} < 0$.

Chapter 6

Controllability and Observability

6.1 Controllability

- Consider the system:

$\dot{x} = Ax + Bu \qquad y = Cx + Du, \qquad \text{where } B(n,m), A(n,n), m \leq n$

Controllability is an open loop issue. It represents steering, or ability to bring an initial vector of states to the zero vector with finite time.

Observability is the ability to construct the values of unmeasured states for all times.

Definition 44 (Controllability) *System is controllable if:*
1° $rank[B\ \ AB\ \ A^2B\ \ ...\ \ A^{n-1}B] = n$
2° $rank[A - \lambda_j IB] = n \qquad \forall \lambda_j, \exists v \neq 0, j = 1, 2, ..., n$
3° $rank[\int_0^t e^{A\tau} BB^T e^{A^T\tau} d\tau] = n \qquad \forall t > 0$

Definition 45 (Controllability) *More explicit definition:*
$\forall x_0, \forall x(T), \forall T > 0, \exists u(\cdot)$ *such that*

$$x(T) = e^{AT}x_0 + \int_0^T e^{A(t-\tau)}Bu(\tau)d\tau$$

Define the controllability grammian as the following integral

$$W_c(t) = \int_0^t e^{A\tau} BB^T e^{A^T\tau} d\tau \tag{6.1}$$

The asymptotic value $W_c(\infty)$ satisfies the Lyapunov equation

$$A^T W_c(\infty) + W_c(\infty) A = BB^T$$

$3^\circ \Rightarrow$ **controllability**

We must compute a specific input that steers the system from an arbitrary initial condition x_0 to an arbitrary desired final state $x(T)$ at time T which means finding $u(.)$ in the formula

$$x(T) = s^{AT} x_0 + \int_0^T e^{A(T-\tau)} Bu(\tau) d\tau \tag{6.2}$$

On this purpose, notice that after a simple change of variables $\tau = T - \bar{\tau}$ we have the following expression for the formula (6.1)

$$W_c(T) = \int_0^T e^{A(T-\bar{\tau})} BB^T e^{A^T(T-\bar{\tau})} d\bar{\tau} \tag{6.3}$$

Let the integral in the controllability definition (6.2) be isolated on the right-hand-side

$$x(T) - e^{AT} x_0 = \int_0^T e^{A(t-\tau)} Bu(\tau) d\tau$$

then we see that the integral resembles 'half' of the controllability grammian expression, leaving the guess as to the right value of $u(.)$ to choose. Indeed setting

$$u(\tau) = B^T e^{A^T(t-\tau)} W_c(T)^{-1} (x(T) - e^{A^T} x_0) \tag{6.4}$$

we have after pulling the terms that do note depend on τ out of the integral the following trivial equality

$$
\begin{aligned}
x(T) - e^{AT}x_0 &= \int_0^T e^{A(t-\tau)}Bu(\tau)d\tau \\
&= \int_0^T e^{A(T-\tau)}B \,\left(B^T e^{A^T(T-\tau)} W_c(T)^{-1}(x(T) - e^{AT}x_0)\right)\, d\tau \\
&= \left(\int_0^T e^{A(T-\tau)}BB^T e^{A^T(T-\tau)} d\tau\right)\, W_c(T)^{-1}(x(T) - e^{AT}x_0) \\
&= W_c(T)W_c(T)^{-1}(x(T) - e^{AT}x_0) \\
&= x(T) - e^{AT}x_0 \qquad (6.5)
\end{aligned}
$$

which proves indeed that the input given in (6.4) transfers the state from x_0 to $x(T)$.

Controllabilty $\Rightarrow 3^\circ$

Let us proceed by contradiction after assuming that both (i) $|W_c(t_1)| = 0$ at a given instant of time $t_1 > 0$ and (ii) the system is controllable. Because $|W_c(t_1)| = 0$,This means that $\exists v \neq 0$ such that $W_c(t_1)v = 0$. Then compute

$$
\begin{aligned}
v^T W_c(t_1)v &= \int_0^T v^T e^{A(t_1-\tau)}BB^T e^{A^T(t_1-\tau)} v d\tau \\
&= \int_0^{t_1} \|B^T e^{A^T(t_1-\tau)}\|^2 d\tau = 0
\end{aligned}
$$

which implies

$$
B^T e^{A^T(t_1-\tau)} \equiv 0 \qquad \forall \tau \in [0, t_1]
$$

which is also

$$
v^T e^{A^T(t_1-\tau)} B \equiv 0 \qquad \forall \tau \in [0, t_1]
$$

Now, the second part of our assumption is that the system is controllable so that there exists a $u(\cdot)$ transferring $x_0 = e^{-At_1}v$ to $x(t_1) = 0$

$$
0 = v + \int_0^{t_1} e^{A(t_1-\tau)}Bu(\tau)d\tau
$$

But taking the scalar product with the vector v gives the sought contradiction since

$$
\begin{aligned}
0 &= v^T v + v^T \int_0^{t_1} e^{A(t_1-\tau)}Bu(\tau)d\tau \\
&= v^T v \qquad (6.6)
\end{aligned}
$$

forces $v = 0$ contradicting the consequence of the first part of our assumption $v \neq 0$.

- Use Cayley-Hamilton to show that $3° \Rightarrow 1°$:

$$e^{At} = \sum_{i=0}^{n} \alpha_i A^i t^i$$

- By contraposition, we will show that $1° \Rightarrow 3°$.

If rank is lost then, $\exists v$ such that $W_c v = 0$

$v^T W_c v = 0$
$\int_0^t v^T e^{A\tau} B B^T e^{A^T \tau} v d\tau = 0$
sum of squares $\rightarrow$ each square must be 0
$B^T e^{A^T \tau} v = 0, \ \forall \tau$

- By contraposition, we will show that $1° \Rightarrow 2°$.

$\exists v \neq 0$ such that $[A - \lambda I B]v = 0$
$Av = \lambda v \qquad Bv = 0$
$A^2 v = AAv = A\lambda v = \lambda^2 v$
$rank[B \ \ \lambda B \ \ \lambda^2 \ \ ... \ \ \lambda^{n-1} B] = 1$

- Controllability is not lost under similarity transform:

$\bar{A} = PAP^{-1} \qquad \bar{B} = PB$
$rank[B \ \ AB \ \ A^2 B \ \ ... \ \ A^{n-1} B] = rank[\bar{B} \ \ \bar{A}\bar{B} \ \ \bar{A}^2 \bar{B} \ \ ... \ \ \bar{A}^{n-1} \bar{B}]$

- An incontrollable system has a specific change of coordinates where the incontrollability becomes obvious:

$\dot{x} = Ax + Bu$

$$\bar{A} = \begin{pmatrix} A_1 & A_{12} \\ 0 & A_2 \end{pmatrix} \qquad \bar{B} = \begin{pmatrix} \bar{B}_1 \\ 0 \end{pmatrix}$$

$$\curvearrowright z = \begin{pmatrix} z_c \\ z_{\bar{c}} \end{pmatrix} \quad \text{(controllable and uncontrollable states)}$$

$\dot{z}_{\bar{c}} = A_2 z_{\bar{c}}$ is not influenced by u

- Show that $2° \Rightarrow 1°$ by contraposition:

$\exists P$ such that the initial system may be transformed into controllable and

uncontrollable states.
Rank will be lost: $rank[B\ \ AB\ \ A^2B\ \ A^{n-1}B] < n$
The objective is to find all $v_i \neq 0$ such that
$rank[B\ \ AB\ \ A^2B\ \ A^{n-1}B]v_i = 0 \qquad i = 1, 2, ..., n$

$[v_i, P_j]$ is a basis

$$z = Px = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ v_1 \\ v_2 \\ \vdots \\ v_r \end{pmatrix} x$$

Compute the eignvector of A_2
$A_2 v_2 = \lambda_2 v_2$

$$\bar{A}\begin{pmatrix} 0 \\ v_2 \end{pmatrix} = \lambda_2 \begin{pmatrix} 0 \\ v_2 \end{pmatrix} \qquad \bar{B}\begin{pmatrix} 0 \\ v_2 \end{pmatrix} = 0$$

$$\curvearrowright (\bar{A} = \lambda I \bar{B}) \begin{pmatrix} 0 \\ v_2 \end{pmatrix} = 0 \qquad \text{which is a contradiction}$$

6.2 Observability

• For observability, we can generally state that if (A, B) is controllable, then (A^T, B^T) is observable.

Definition 46 (Observability) *The system is observable if :*

1° $y(\cdot), u(\cdot) \Rightarrow x(0), x(\cdot)$

$$2^\circ \ rank \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} = n$$

3° $rank \begin{pmatrix} A-\lambda I \\ C \end{pmatrix} = n$

4° $rank[\int_0^t e^{A^T\tau} C^T C e^{A\tau} d\tau] = n$

- $W_O(\infty)A + A^T W_O(\infty) = C^T C$
-

$$rank \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} = rank[C^T \;\; A^T C^T \;\; ... \;\; (A^T)^{n-1} C^T]$$

$$\Leftrightarrow$$

$$rank[B \;\; Ab \;\; A^w B \;\; ... \;\; A^{n-1}B] = rank \begin{pmatrix} B^T \\ B^T A^T \\ B^T (A^T)^2 \\ \vdots \\ B^T (A^T)^{n-1} \end{pmatrix}$$

- Show that $1^\circ \Rightarrow 4^\circ$ by contraposition:

$y = Cx + Du = Ce^{At}x_0 + C\int_0^t e^{A(t-\tau)} Bu(\tau) d\tau + Du$

y is measured and last two terms are known because we control them.

$\bar{y} = y - C\int_0^t e^{A(t-\tau)} Bu(\tau) d\tau - Du = Ce^{At}x_0$

$x_0 = W_O^{-1}(t)(\int_0^t e^{A^T\tau} C^T C e^{A\tau} d\tau) x_0$

$x_0 = W_O^{-1}(t) \int_0^t e^{A^T\tau} C^T \bar{y} d\tau$

Now, start with a state corresponding to a vector that cancels W_O
$x_0 = v \qquad u = 0 \qquad W_O \cdot v = 0$

$\curvearrowright$ we can not distinguish between a zero-state from nonzero-state (i.e. the system is unobservable).

6.3 Invariant subspaces

6.3.1 Maximal Controllable Subspace

6.3.2 Maximal Observable Subspace

Chapter 7

Realization

Assume that a transfer function of a system is known and we want to obtain a state-space model:

$$G(s) \rightsquigarrow \begin{array}{l} \dot{x} = Ax + Bu \\ y = Cx + Du \end{array}$$

7.0.3 Single-input single-output case

- $G(s) = \dfrac{\beta_1 s^3 + \beta_2 s^2 + \beta_3 s + \beta_4}{s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4} = \dfrac{N(s)}{D(s)}$
- Sequence of $N(s)$ and $D(s)$ is important for non-SISO cases.

- One possible realization:

$$\frac{1}{D(s)} = \frac{Y(s)}{U(s)} \cdot \frac{1}{N(s)} = G(s)$$

$$\frac{1}{s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4} = \frac{Y(s)}{U(s)}$$

$$v^{(4)} + \alpha_1 v^{(3)} + \alpha_2 v^{(2)} + \alpha_3 \dot{v} + \alpha_4 v = u$$

$$x_1 = v^{(2)} \qquad x_2 = v^{(2)} \qquad x_3 = \dot{v} \qquad x_4 = v \qquad x^T = [x_1 \;\; x_2 \;\; x_3 \;\; x_4]$$

$$\dot{x}_1 = v^{(4)} = -\alpha_1 v^{(3)} - \alpha_2 v^{(2)} - \alpha_3 \dot{v} - \alpha_4 v - u$$

$\dot{x}_2 = v^{(3)} = x_1$
$\dot{x}_3 = v^{(2)} = x_2$
$\dot{x}_4 = \dot{v} = x_3$

$$A = \begin{pmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad D = 0$$

For controllability: $\alpha_4 \neq 0$ or $rank[B \;\; AB \;\; A^2B \;\; A^3B] = 4$

$y = \beta_1 v^{(3)} + \beta_2 v^{(2)} + \beta_3 \dot{v} + \beta_4 v = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$

$C = [\beta_1 \;\; \beta_2 \;\; \beta_3 \;\; \beta_4]$

• $D(s)$ and $N(s)$ need to be coprime to avoid pole-zero cancellation and loss of observability.

Theorem 8 *If the controllable canonical realization is observable* (§)*, then* $N(s)$ *and* $D(s)$ *are coprime and vice versa* $(\star)$*.*

Proof:

(§) $\Rightarrow$ $(\star)$ by contraposition:
Suppose N and D have a common zero $\lambda_n \in \mathbb{R} : N(\lambda_n) = D(\lambda_n) = 0$
$\lambda_n^4 + \alpha_1 \lambda_n^3 + \alpha_3 \lambda_n^2 + \alpha_4 = 0$
$\beta_1 \lambda_n^3 + \beta_2 \lambda_n^2 + \beta_3 \lambda_n + \beta_4 = 0$

$$A = \begin{pmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad C = [\beta_1 \;\; \beta_2 \;\; \beta_3 \;\; \beta_4]$$

$$\chi = \begin{pmatrix} \lambda_n^3 \\ \lambda_n^2 \\ \lambda_n \\ 1 \end{pmatrix} \qquad \nu = \begin{pmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{pmatrix}$$

It should be shown that ν loses rank, i.e. $\exists x \neq 0, \nu x = 0$.

$Cx = \beta_1 \lambda_1^3 + \beta_2 \lambda_1^2 + \beta_3 \lambda_1 + \beta_4 = 0$

$$CAx = C \begin{pmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1^3 \\ \lambda_1^2 \\ \lambda_1 \\ 1 \end{pmatrix} = C \begin{pmatrix} -\alpha_1 \lambda_1^3 - \alpha_2 \lambda_1^2 - \alpha_3 \lambda_1 - \alpha_4 \\ \lambda_1^3 \\ \lambda_1^2 \\ \lambda_1 \end{pmatrix} =$$

$\lambda_1 Cx = 0$

It loses rank $\rightarrow$ the system is not observable.

$(\star) \Rightarrow (\S)$ by contraposition:

$$rank \begin{pmatrix} A - \lambda I \\ C \end{pmatrix} \qquad \chi = \begin{pmatrix} \lambda^3 \\ \lambda^2 \\ \lambda \\ 1 \end{pmatrix}$$

$(\S)$ implies that $\exists x \neq 0$ such that $\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} x = 0$

$Cx = 0 \qquad (A - \lambda I)x = 0 \qquad Cx = N(\lambda) = 0 \qquad |\lambda I - A| = D(\lambda)$

So, $N(\lambda) = D(\lambda) = 0$ (not coprime). ♠

- A good realization would mix controllability and observability.

- Define two state-space systems:

System I: (A, B, C, D)
System II: $(\bar{A}, \bar{B}, \bar{C}, \bar{D})$

Expansion:
$G(s) = C(sI - A)^{-1}B + D = \bar{C}(sI - \bar{A})^{-1}\bar{B} + \bar{D}$

$D + CBs + CABs^2 + CA^2Bs^3 + CA^3Bs^4 + ... =$
$\bar{D} + \bar{C}\bar{B}s + \bar{C}\bar{A}\bar{B}s^2 + \bar{C}\bar{A}^2\bar{B}s^3 + \bar{C}\bar{A}^3\bar{B}s^4 + ...$
$D = \bar{D} \qquad CB = \bar{C}\bar{B} \qquad CAB = \bar{C}\bar{A}\bar{B} \qquad ... \qquad CA^2B = \bar{C}\bar{A}^2\bar{B}$

$$\dot{z} = P\dot{x} \qquad \dot{z} = \bar{A}Px + \bar{B}u \qquad y = \bar{C}Px + \bar{D}u$$

$$\begin{array}{llll} A = P^{-1}\bar{A}P & B = P^{-1}\bar{B} & C = \bar{C}P & D = \bar{D} \\ \bar{A} = PAP^{-1} & \bar{B} = PB & \bar{C} = CP^{-1} & \bar{D} = D \end{array}$$

$$\begin{array}{l} CB = \bar{C}PP^{-1}\bar{B} = \bar{C}\bar{B} \\ CAB = \bar{C}P(P^{-1}AP)P^{-1}\bar{B} = \bar{C}(PP^{-1})\bar{A}(PP^{-1})\bar{B} = \bar{C}\bar{A}\bar{B} \\ ... \end{array}$$

- Observability and controllability matrices:

$$\mathcal{O} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} \qquad \mathcal{C} = [B \;\; AB \;\; A^2B \;\; ... \;\; A^{n-1}B]$$

- For the similarity transforms:

$$\bar{\mathcal{O}} = \begin{pmatrix} \bar{C} \\ \bar{C}\bar{A} \\ \bar{C}\bar{A}^2 \\ \vdots \\ \bar{C}\bar{A}^{n-1} \end{pmatrix} \qquad \bar{\mathcal{C}} = [\bar{B} \;\; \bar{A}\bar{B} \;\; \bar{A}^2\bar{B} \;\; ... \;\; \bar{A}^{n-1}\bar{B}]$$

- The analogs are not identical ($\mathcal{O} \neq \bar{\mathcal{O}}$, $\mathcal{C} \neq \bar{\mathcal{C}}$), but their products both give the same Hankel matrix:

$$\mathcal{O}\mathcal{C} = \bar{\mathcal{O}}\bar{\mathcal{C}} = \begin{pmatrix} CB & CAB & CA^2B \\ CAB & CA^2B & \vdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$\curvearrowright \quad \bar{C}\bar{B} = CP^{-1}PB = CB \qquad \bar{C}\bar{A}\bar{B} = CP^{-1}PAP^{-1}PB = CAB$$

- The two systems have the same transfer function $G(s) = C(sI - A)^{-1}B$.

Theorem 9 *Two minimal realizations (dimension equal to degree of $D(s)$, with $N(s)$ and $D(s)$ coprime) are similar and equivalent.*

• Product of observability, A, and controllability lends the same Markov parameters as the Hankel matrix, but in a different order:

$$\mathcal{O}A\mathcal{C} = \bar{\mathcal{O}}\bar{A}\bar{\mathcal{C}}$$

$$\begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} A \, [B \;\; AB \;\; A^2B \;\; ... \;\; A^{n-1}B] = \bar{\mathcal{O}}\bar{A}\bar{\mathcal{C}}$$

$$\bar{\mathcal{O}}^{-1}\mathcal{O}A\mathcal{C}\bar{\mathcal{C}}^{-1} = \bar{A} \qquad P = \bar{\mathcal{O}}^{-1}\mathcal{O} \qquad P^{-1} = \mathcal{C}\bar{\mathcal{C}}^{-1}$$

• Consider the systems:
$\dot{x} = Ax + Bu \qquad \dot{z} = \bar{A}x + \bar{B}u$
$y = Cx + Du \qquad y = \bar{C}x + Du$

The question is how to build the matrix P, having in mind a fact that these 2 systems have the same transfer function $G(s) = \dfrac{N(s)}{D(s)}$.

- Because of minimality, we have observability and controllability.

- To reformulate the question, we need to build the matrix P so as to "balance" between these two properties.

- Characterization of controllability:

$AW_C + W_C A^T = -BB^T$

If it is controllable, than W_C is the unique solution of the system.

$W_C = \int_0^\infty e^{A^T(BB^T)} e^{A^T\tau} d\tau$

It should be full rank.

- Likewise for observability:

$A^T W_O + W_O A = -C^T C$

Theorem 10 *Suppose that systems* (A, B, C) *and* $(\bar{A}, \bar{B}, \bar{C})$ *are minimal and equivalent, then* $W_C W_O$ *and* $\bar{W}_C \bar{W}_O$ *are similar.*

Proof:

$W_C W_O = P^{-1} \bar{W}_C \bar{W}_O P, \quad P$ connects the two properties

$\bar{A}\bar{W}_C + \bar{W}_C \bar{A}^T = -\bar{B}\bar{B}^T$

$PAP^{-1}\bar{W}_C + \bar{W}_C (P^{-1})^T A^T P^T = -PBB^T P^T$

$P^{-1}(PAP^{-1}\bar{W}_C + \bar{W}_C (P^{-1})^T A^T P^T)(P^T)^{-1} = -P^{-1}PBB^T P^T (P^T)^{-1}$

$AP^{-1}\bar{W}_C (P^T)^{-1} + P^{-1}\bar{W}_C (P^{-1})^T A^T = BB^T$

$W_C = P^{-1}\bar{W}_C (P^T)^{-1}$

$AW_C + W_C A^T = -BB^T$

The same for observability, and we get:

$W_O = P^T \bar{W}_O P$

$W_O W_C = P^T \bar{W}_O P\ P^{-1} \bar{W}_C (P^T)^{-1} = P^T \bar{W}_O \bar{W}_C (P^T)^{-1} = P^{-1} \bar{W}_O \bar{W}_C P$ ♠

Theorem 11 *The product* $W_C W_O$ *share the same eigenvalues as* $\bar{W}_C \bar{W}_O$ *(characteristic polynomials are the same).*

Theorem 12 *There exists* P*, such that* $\bar{W}_O = \bar{W}_C = \Sigma$ *(diagonal).*

- $W_C = Q^T D Q$ diagonalize it, because W_C is positive definite.

$W_C = Q^T D^{1/2} D^{1/2} Q = R^T R$, because all elements of the diagonal D are positive

Lemma 12 *$W_C W_O$ has the same eigenvalues as $RW_O R^T$.*

- So, we want to show $R^T RW_O \neq RW_O R^T$, but $eig(R^T RW_O) = eig(RW_O R^T)$.

  $\curvearrowright$ Two general matrices, M and N, with dimension n:

  $det(sI - MN) = det(sI - NM)$, but obviously $MN \neq NM$

- $det(\sigma^2 I - W_C W_O) = det(\sigma^2 I - R^T RW_O) = det(\sigma^2 I - RW_O R^T)$

- Construction of matrix P:

  $W_C = R^T R$

  $P = \Sigma^{1/2} U^T (R^T)^{-1}$

  $W_C W_O$ share same eigenvalues as $R^T W_O R$

  $R^T W_O R = U\Sigma^2 U^T$

  $\bar{W}_C = PW_C P^T = \Sigma^{1/2} U^T (R^T)^{-1} W_C R^{-1} U \Sigma^{1/2} = \Sigma^{1/2} U^T U \Sigma^{1/2} = \Sigma$
  $\curvearrowright$ $U^T U = I = UU^T \qquad \Sigma^T = \Sigma$

  $RW_O R^T = U\Sigma^2 U^T$

  $\bar{W}_O = (P^T)^{-1} W_O P^{-1} = \Sigma^{-1/2} U^T RW_O R^T U \Sigma^{-1/2}$

  $\bar{W}_O = \Sigma^{-1/2} U^T U \Sigma^2 U^T U \Sigma^{-1/2} = \Sigma$

- What we did is:
  $A \quad B \quad \rightsquigarrow \quad \bar{A} \quad \bar{B}$ balancing W_C and W_O such that $\bar{W}_C = \bar{W}_O = \Sigma$
  $C \quad D \quad \rightsquigarrow \quad \bar{C} \quad \bar{D}$ balanced transformation.

- Finding a good A, B, C, and D from $G(s)$ is still an open issue.

- SISO case (it can be applied to MIMO very easily):

$$G(s) = \frac{\beta_1 s^{n-1} + \beta_2 s^{n-2} + ... + \beta_n}{s^n + \alpha_1 s^{n-1} = \alpha_2 s^{n-2} + ... + \alpha_n}$$

$$G(s) = h(1)s^{-1} + h(2)s^{-2} + h(3)s^{-3} + ...$$

$$G(s) = C(sI - A)^{-1}B = CBs^{-1} + CABs^{-2} + CA^2Bs^{-3} + ...$$

$s^{-1}I + s^{-2}A + s^{-3}A^2 + ... = s^{-1}(I + s^{-1}A + s^{-2}A^2 + ...) = s^{-1}(I - s^{-1}A)^{-1} = (sI - A)^{-1}$

$$C(sI - A)^{-1}B = CBs^{-1} + CABs^{-2} + CA^2Bs^{-3} + ...$$

$h(1) = CB \qquad h(2) = CAB \qquad h(3) = CA^2B$ ... Markov parameters

For the realization: collect Markov parameters $h(i)$-s and fill the Hankel matrix.

The problem that occurs is that the series never end.

$$T(\alpha,\beta) = \begin{pmatrix} h(1) & h(2) & h(3) & \dots & h(\beta) \\ h(2) & h(3) & h(4) & \dots & h(\beta+1) \\ h(3) & h(4) & h(5) & \dots & h(\beta+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h(\alpha) & h(\alpha+1) & h(\alpha+2) & \dots & h(\alpha+\beta-1) \end{pmatrix} =$$

$$= \begin{pmatrix} CB & CAB & CA^2B) & \dots \\ CAB & CA^2B & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

Although, the matrix is well defined, we don't know values of α and β.

$h(1) = \beta_1$, from division
$h(2) = -\alpha_1 h(1) + \beta_2$
$h(3) = -\alpha_1 h(2) - \alpha_2 h(1) + \beta_3$
$\vdots$

$h(n) = -\alpha_1 h(n-1) - \alpha_2 h(n-2) - ... - \alpha_{n-1} h(1) + \beta_n$
$\vdots$
$h(m) = -\alpha_1 h(m-1) - \alpha_2 h(m-2) - ... - \alpha_n h(m-n) + \beta_n$

$h(i)$ is nothing more else than the impulse response. So, we can do the realization by collecting the impulse response.

Theorem 13 *$\forall m \geq n$, $\exists n$, such that rank$(T(n,n))$ =rank$(T(m,m))$.*

- If $G(s)$ is minimal (means that the system is controllable), then $T(n,n) \rightarrow$ $\deg\{D(s)\} = n$.

- The second matrix of interest is:

$$\tilde{T}(n,n) = \begin{pmatrix} h(2) & h(3) & \dots \\ h(3) & h(4) & \dots \\ h(4) & h(5) & \dots \\ \vdots & \vdots & \ddots \\ h(n+1) & h(n+2) & \dots \end{pmatrix} = \quad = \begin{pmatrix} CAB & CA^2B) & \dots \\ CA^2B & \dots & \dots \\ \vdots & \ddots & \vdots \end{pmatrix}$$

C and B are easy to extract:

$$T = \mathcal{O}\mathcal{C} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} \cdot [B \;\; AB \;\; A^2B \;\; ... \;\; A^{n-1}B]$$

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

- If we collect T and $\tilde{T}$, we can rebuild A.

- So, if we can factorize T as $\mathcal{O}$ and $\mathcal{C}$, then B is the first column of $\mathcal{C}$ and C is the first row of $\mathcal{O}$. With this, we obtain a state-space representation.

- Factorization of T:

We'll use singular value decomposition (SVD).
Idea is to build positive definite matrix Λ and two other matrices (basis) K and L.

$$K^T \Lambda L = T = K^T \Lambda^{1/2} \Lambda^{1/2} L = \mathcal{O}\mathcal{C}$$

$$\Lambda = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix}, \qquad \sigma_i \in \mathbb{R}^+$$

$\curvearrowright$ Properties: $K^T K = I = L^T L \qquad LL^T = I = L^T L$
$\mathcal{O}^T \mathcal{O} = \Lambda^{1/2} K K^T \Lambda^{1/2} = \Lambda \qquad \mathcal{C}^T \mathcal{C} = \Lambda^{1/2} L L^T \Lambda^{1/2} = \Lambda$

- MIMO case (p outputs and m inputs):

$G(s) = H(0) + H(1)s^{-1} + H(2)s^{-2} + ...$
$H(i)-$ matrices of constant numbers

$$T(\alpha, \beta) = \begin{pmatrix} H(1) & H(2) & \dots & H(\beta) \\ H(2) & H(3) & \dots & H(\beta+1) \\ H(3) & H(4) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ H(\alpha) & H(\alpha+1) & \dots & H(\alpha+\beta) \end{pmatrix} = \begin{pmatrix} CB & CAB & CA^2B) & \dots \\ CAB & CA^2B & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

$\exists l \in \mathbb{N}$ such that $\exists T(l, l)$ for which rank($T(l+1.l+1)$)=rank($T(l.l)$)

Using SVD:

$$T = K^T \begin{pmatrix} \Lambda & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} L$$

where $\Lambda = \Lambda(n,n)$ and n is a size of state-space (i.e. $\dim\{x\} = n$)

For SVD we have the property that σ_i-s are ordered as in a decreasing sequence:

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_n^2 \end{pmatrix} \qquad \sigma_1^2 > \sigma_2^2 > ... > \sigma_n^2 > 0$$

We can simply refactorize T as: $T(l,l) = \bar{K}^T \Lambda \bar{L}$, meaning that we take only k first rows of K and L (because of the multiplication with zeros in matrices).

$\bar{K}$ is no more a square matrix; $K^T = (l \cdot p, n)$, $\Lambda = (n,n)$, $\bar{L} = (n, l \cdot m)$

$\mathcal{O} = \bar{K}^T \Lambda^{1/2} = \mathcal{O}(l \cdot p, m) \qquad \mathcal{C} = \Lambda^{1/2} \bar{L} = \mathcal{C}(n, n \cdot l)$

$T = \mathcal{O}\mathcal{C} \qquad \tilde{T} = \mathcal{O}A\mathcal{C}$

Need for a new way of extraction, since $\mathcal{O}$ is not a square matrix anymore.

Use pseudo-inverse methodology:
$(\mathcal{O}^T\mathcal{O})^{-1} = \mathcal{O}^+ \qquad (\mathcal{C}^T\mathcal{C})^{-1} = \mathcal{C}^+ \qquad \mathcal{O}^+ T \mathcal{C}^+ = A$
$\curvearrowright \; (\mathcal{O}^T\mathcal{O})^{-1}\mathcal{O}A\mathcal{C}\mathcal{C}^T(\mathcal{C}\mathcal{C}^T)^{-1}$

B and C can be easily extracted: B is the first m columns of $\mathcal{C}$ and C is the first p rows of $\mathcal{O}$.

- Getting $H(s)$ is tricky: can generate it from $G(s)$ if we have $G(s)$; generating experimentally requires growing-order derivatives and the validity is quickly lost in the noise.

- The similar routine as in SISO should be performed: triger one by one input and calculate the rest.

Chapter 8

State Feedback

This chapter deals with state feedback for eigenvalue assignment both in the single-input setting and in the multiple-input setting. If the system is controllable then any possible values of the eigenvalues can be assigned through a single state feedback using real numbers only. We will present this result through an example only in the multiple-input case and present a full proof in the single-input case.

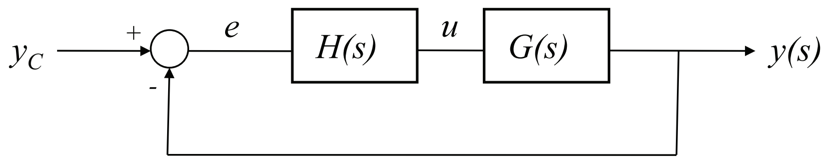
A technique relying on solving a Lyapunov equation is also introduced which can be applied toboth the single and the multiple input cases. It relies however on solving the Lyapunov equation. We omit the details as to how to solve this Lyapunov equation in an efficient and systematic way.

We will then develop a technique based on the cyclic decomposition of the state space into a direct sum of cyclic subspace with each successive summand having a MAP that divides the MAP of the previous summand. This has been explained in the chapter concerning annihilating polynomials. Although the pole placement technique presented is presented through an example only, it stresses the extra freedom of choice available in the MIMO setting. It does not rely on solving a Lyapunov equation.

Consider the system:

$$\begin{aligned} \dot{x} &= Ax + B \\ y &= x \end{aligned}$$

- The controller "does not need" any "dynamics". Its purpose is:
- to get rid of errors (in outputs, states...)
- stability - assignment of poles.



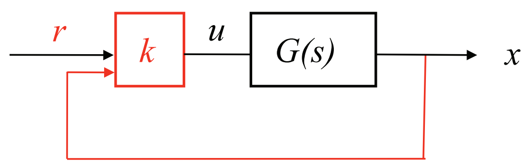
- So, we can do all this with constants.

- We need to know u, build $H(s)$, knowing the error.

- $u = -[k]x + y_C = -kx + r \qquad k = [k_1\ k_2\ \ldots\ k_n] = k(1, n), \quad k_i \in \mathbb{R}$
  $\curvearrowright$ r - reference

  $$u = -\sum_{i=1}^{n} k_i x_i + r$$

- We redraw the system:



- $\dot{x} = Ax + Bu \qquad u = -kx$

- We want to push all the states to zero, so $r = 0$

  $$\dot{x} = Ax - Bkx = (A - Bk)x$$

- If $r \neq 0$ then:

  $$\dot{x} = (A - Bk)x + Br,\ \bar{A} = A - Bk \text{ and } \bar{B} = Br$$

Theorem 14 *If (A, B) is controllable (meaning rank$[B\ AB\ \ldots\ A^{n-1}B] = n$) and we have random matrix k, then rank$[B\ (A{-}Bk)B\ (A{-}Bk)^2B\ \ldots\ (A-$*

$Bk)^{n-1}B] = n$. The controllability is never lost by changing the controller k.

Proof:
Example:

$$[B\ AB\ A^2B\ A^3B] \begin{pmatrix} 1 & -kB & -k(A-Bk)B & -k(A-Bk)^2B \\ 0 & 1 & -kB & -k(A-Bk)B \\ 0 & 0 & 1 & -kB \\ 0 & 0 & 0 & 1 \end{pmatrix} =$$

$$= [B\ (A-Bk)B]\ (A-Bk)^2B\ (A-Bk)^3B$$ ♠

- SISO case

$$\dot{x} = Ax + B \qquad u = r - kx$$

$$\mathcal{C} = [B\ \ AB\ \ A^2B\ \ A^3B]$$

$$\mathcal{C}_K = [B\ \ (A-Bk)B\ \ (A-Bk)^2B\ \ (A-Bk)^3B]$$

$$\mathcal{C}_K = \mathcal{C} \begin{pmatrix} 1 & -kB & -k(A-Bk)B & -k(A-Bk)^2B \\ 0 & 1 & -kB & -k(A-Bk)B \\ 0 & 0 & 1 & -kB \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Some eigenvalues of A might be positive (unstable), so there is a need to transform A in order to correct it.

$\dot{x} = (A - Bk)x + Br = \tilde{A}x + Br$
Need to find k such that $\text{eig}(\bar{A}) \in \mathbb{C}_-$.

Because of dimensionality issues, k takes n states and we have one input.

Use similarity transform:
$\lambda(P\tilde{A}P^{-1}) = \lambda(\tilde{A}) \qquad \lambda(PAP^{-1}) = \lambda(A)$

We want to seek for the correspondence with respect to similarity transform:

$\dot{x} = Ax + Bu \quad \rightarrow \dot{z} = \tilde{A}z + \tilde{B}u$
$\curvearrowright \quad z = Px \qquad \tilde{A} = PAP^{-1} \qquad \tilde{B} = PB$

Add the feedback:
$u = -kx + r = -kP^{-1}z + r$
$\dot{z} = \tilde{A}z + Bu = \tilde{A}z - \tilde{B}kP^{-1}z + \tilde{B}r$

Controllable canonical form is a nice system which can facilitate finding of coefficients of characteristic polynomial.

$$\tilde{A} = \begin{pmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \tilde{B} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\tilde{\tilde{A}} = (\tilde{A} - \tilde{B}kP^{-1})$ and $\tilde{k} = \tilde{B}k$, but we are still looking for k.

$|\lambda I - A| = \lambda^4 + \lambda^3\alpha_1 + \lambda^2\alpha_2 + \lambda\alpha_3 + \alpha_4$

The usefulness of transformed system:

$$\tilde{B}\tilde{k} = \begin{pmatrix} \tilde{k}_1 & \tilde{k}_2 & \tilde{k}_3 & \tilde{k}_4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \tilde{\tilde{A}} = \begin{pmatrix} -\alpha_1 - \tilde{k}_1 & -\alpha_2 - \tilde{k}_2 & -\alpha_3 - \tilde{k}_3 & -\alpha_4 - \tilde{k}_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$|\lambda I - \tilde{\tilde{A}}| = \lambda^4 + (\alpha_1 + \tilde{k}_1)\lambda^3 + (\alpha_2 + \tilde{k}_2)\lambda^2 + (\alpha_3 + \tilde{k}_3)\lambda + (\alpha_4 + \tilde{k}_4)$

To proceed:

1) Choose the 4 desired closed-loop eigenvalues: $\tilde{\lambda_1}, \tilde{\lambda_2}, \tilde{\lambda_3}, \tilde{\lambda_4}, Re(\tilde{\lambda_i}) < 0$

2) Expand the polynomial and set it equal to characteristic polynomial $(\lambda - \tilde{\lambda_1})(\lambda - \tilde{\lambda_2})(\lambda - \tilde{\lambda_3})(\lambda - \tilde{\lambda_4}) = \lambda^4 + (\alpha_1 + \tilde{k}_1)\lambda^3 + (\alpha_2 + \tilde{k}_2)\lambda^2 + (\alpha_3 + \tilde{k}_3)\lambda + (\alpha_4 + \tilde{k}_4)$

3) Express $\tilde{k}_i$ by identification of common powers of λ.

4) $k = \tilde{k}P$
This step is a bit problematic, because we don't have P yet.

- Controllable canonical form:

$$\tilde{\mathcal{C}} = \begin{pmatrix} 1 & -\alpha_1 & -\alpha_1^2-\alpha_2 & -\alpha_1^3+2\alpha_1\alpha_2+\alpha_3 \\ 0 & 1 & -\alpha_1 & \alpha_1^2-\alpha_2 \\ 0 & 0 & 1 & -\alpha_1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = [\tilde{B}\ \ \tilde{A}\tilde{B}\ \ \tilde{A}^2\tilde{B}\ \ \tilde{A}^3\tilde{B}]$$

$\mathcal{C} \leftrightarrow \tilde{\mathcal{C}}$, $\qquad$ P connects these two entities: $P = \tilde{\mathcal{C}}\mathcal{C}^{-1}$

$P^{-1} = \mathcal{C}\tilde{\mathcal{C}}^{-1}$

$\mathcal{C} = [B\ \ AB\ \ A^2B\ \ A^3B]$

$$\tilde{\mathcal{C}}^{-1} = \begin{pmatrix} 1 & \alpha_1 & \alpha_2 & \alpha_3 \\ 0 & 1 & \alpha_1 & \alpha_2 \\ 0 & 0 & 1 & \alpha_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

bullet Choosing eigenvalues can be tricky due to numerical sensitivity, so an alternative is the Lyapunov-like pole-placement method.

8.1 Lyapunov-like pole-placement method

- Ingredients:

1) Compute the eigenvalues of A (just to have a good guess where they are)

2) Choose a F matrix, n by n, with desired eigenvalues $\lambda_i(F) \neq \lambda_j(A), \forall i, j$. Should not set them on the same position as the initial one (open-loop).

3) Choose a row vector $\tilde{k}$, such that the system $(\tilde{K}, F)$ becomes observable

for example: $\text{rank} \begin{pmatrix} \tilde{K} \\ \tilde{K}F \\ \tilde{K}F^2 \\ \tilde{K}F^3 \end{pmatrix} = 4$

4) Solve $AT - TF = B\tilde{K}$ for T. This is *Lyapunov-like equation.*

5) $K = \tilde{K}T^{-1}$ $\quad$ T should be invertible.

- Assume that T is invertible:

  $AT = TF = B\tilde{K}$ $\quad K = \tilde{K}T^{-1}$ $\quad \tilde{K} = KT$
  $AT - TF = BKT$ we want to isolate closed-loop transfer function
  $(A - BK)T = TF$ $\quad T^{-1}(A - BK)T = F$
  F and $(A - BK)$ are similar and have the desired eigenvalues.

- To justify the assumption that T is invertible:

  $$\begin{aligned} \alpha_4(IT - TI) &= \\ \alpha_3(AT - TF) &= \dots \\ \alpha_2(A^2T - TF^2) &= \dots \qquad \text{sum these} \\ \alpha_1(A^3T - TF^3) &= \dots \\ (A^4F - TF^4) &= \dots \end{aligned}$$

  $\Delta(\lambda) = \lambda^4 + \alpha_1\lambda^3 + \alpha_2\lambda^2 + \alpha_3\lambda + \alpha_4$
  $A^4 + \alpha_1 A^3 + \alpha_2 A^2 + \alpha_3 A + \alpha_4 = 0$

  If we sum everything up, al the members will vanish and the result will be 0.

- If λ_i is an eigenvalue of F, then $\Delta(\lambda_i)$ is also an eigenvalue of $\Delta(F)$.
  $|\bar{\lambda}I - F| = |\Delta(\bar{\lambda}_i)I - \Delta(F)| = 0$, prove by Cauchy-Binet.

- $AT - TF = B\bar{K}$ $\quad AT = B\bar{K} + TF$
  $$\begin{aligned} A^2T - TF^2 &= A(AT) - TF^2 = A(B\bar{K} + TF) - TF^2 = \\ &= AB\bar{K} - ATF - TF^2 = (AT - TF)F + AB\bar{K} = AB\bar{K} + B\bar{K}F \end{aligned}$$
  and so on for higher terms..

- We thus finish to justify the assumption of T being full rank:

  $$\begin{aligned} \alpha_4(IT - TI) &= 0 \\ \alpha_3(AT - TF) &= (B\bar{K})\alpha_3 \\ \alpha_2(A^2T - TF^2) &= (AB\bar{K} + B\bar{K}F)\alpha_2 \\ \alpha_1(A^3T - TF^3) &= (A^2B\bar{K} + AB\bar{K}F + B\bar{K}F^2)\alpha_1 \\ (A^4F - TF^4) &= (A^3B\bar{K} + A^2B\bar{K}F + AB\bar{K}F^2 + B\bar{K}F^3) \cdot 1 \\ \hline &= \Delta(A)T - T\Delta(F) \end{aligned}$$

$$\Delta(A)T - T\Delta(F) = [B\ AB\ A^2B\ A^3B]\begin{pmatrix} \alpha_3 & \alpha_2 & \alpha_1 & 1 \\ \alpha_2 & \alpha_1 & 1 & 0 \\ \alpha_1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \bar{K} \\ \bar{K}F \\ \bar{K}F^2 \\ \bar{K}F^3 \end{pmatrix}$$

This implies that T is full-rank:

- first term cancels by Cayley-Hamilton
- controllability matrix is full-rank since the system is controllable
- the middle matrix is full-rank by construction
- the last matrix is full-rank by observability assumption.

8.2 Langenhop vectors and canonical form

Let a given matrix A of size $n \times n$ be given having a MAP (w.r.t. A) of the whole space $\mathbb{R}^n$ of maximum order n. This means that there exists at least one vector v of cyclicity index of n. We will introduce particular vectors linked with this cyclic vector v. The MAP (w.r.t. A) is equal to the characteristic polynomial and is of order n

$$\Delta(\lambda) = |\lambda I - A| = \lambda^n + \alpha_1\lambda^{n-1} + \alpha_2\lambda^{n-2} + \ldots + \alpha_n$$

Let v be a vector of cyclicity index of n which means that the vectors in the familiy of vectors

$$\{v, Av, A^2v, \ldots, A^{n-1}v\}$$

are linearly independent and spans the whole space $\mathbb{R}^n$.

Let us introduce in succession the following vectors (in reverse numeratotion)

$$\begin{array}{rcl} e_n &=& v \\ e_{n-1} &=& Ae_n + \alpha_1 v \\ e_{n-2} &=& Ae_{n-1} + \alpha_2 v \\ e_{n-3} &=& Ae_{n-2} + \alpha_3 v \\ \vdots && \vdots \\ e_1 &=& Ae_2 + \alpha_{n-1} v \end{array}$$

which, in-closed form, leads to the following defining formula (whose right-hand side only depends on the vector v)

$$e_k = A^{n-1}v + \sum_{j=1}^{n-k} \alpha_j A^j v \qquad k = 1, \dots n \tag{8.1}$$

One can then deduce that the matrix A acts on these vectors according to

$$Ae_k = e_{k-1} - \alpha_k e_k \tag{8.2}$$

or in the canonical basis $\{e_1, e_2, \dots, e_n\}$ the A matrix becomes after introducing the change of basis matrix

$$P = \begin{pmatrix} e_1 & e_2 & \dots & e_n \end{pmatrix} \tag{8.3}$$

the following matrix

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_n & -\alpha_{n-1} & -\alpha_{n-2} & \dots & -\alpha_1 \end{pmatrix}$$

where the effect of formula (8.2) reveals itself.

8.3 Pole assignment in single input controllable case

We will now apply the computation of the previous section to a control system with a single input

$$\dot{x} = Ax + bu$$

so as to assign the eigenvalues to a specific set of n complex numbers in complex conjugate pairs.

On this purpose, let us introduce the change of coordinates (associated with the change of canonical basis vectors given by the P matrix above)

but with the cyclic vector v being set to the input vector b. This is possible since under the assumption of controllability the b vectors generates a cyclic family of vectors of maximum order, i.e. a set of n linearly independent vectors

$$\{b, Ab, A^2b, \ldots, A^{n-1}b\}$$

$$z = P^{-1}x$$

which give the followinig expression

$$\dot{z} = P^{-1}APz + P^{-1}bu \tag{8.4}$$

which gives a transparent transcription and visible expresion of the characteristic polynomial since the transformed matrix

8.4 Eigenvalue assignment through cyclic decomposition

Consider the following square 4×4 matrix A together with the 4×2 matrix B

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$

that define the multi-input open-loop control system

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

with $x = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix}^T$ and $u = \begin{pmatrix} u_1 & u_2 \end{pmatrix}^T$.

Notice that the open-loop eigenvalues equal $2, 3, 4$, with 2 a repeated eigenvalue.

The system is not controllable using one input channel only. For instance, defining b_1 and b_2 as the first and second column of $B = \begin{pmatrix} b_1 & b_2 \end{pmatrix}$, the first input does not achieve controllability since

$$\text{rank} \begin{pmatrix} b_1 & Ab_1 & A^2b_1 & A^3b_1 \end{pmatrix} = \text{rank} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix} = 3 < 4$$

Similarly, the second input does not achieve controllability either

$$\text{rank} \begin{pmatrix} b_2 & Ab_2 & A^2b_2 & A^3b_2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 0 & 0 & 0 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{pmatrix} = 3 < 4$$

Nevertheless using both inputs does the system become controllable since

$$\text{rank} \begin{pmatrix} B & AB & A^2B & A^3B \end{pmatrix} = \text{rank} \begin{pmatrix} 0 & 1 & 0 & 2 & 0 & 4 & 0 & 8 \\ 1 & 0 & 2 & 0 & 4 & 0 & 8 & 0 \\ 1 & 1 & 3 & 3 & 9 & 9 & 27 & 27 \\ 1 & 1 & 4 & 4 & 16 & 16 & 64 & 64 \end{pmatrix} = 4$$

Problem 1 *Design a feedback mechanism so as to set the closed-loop eigenvalues at purely imaginary values $+i, -i, +2i, -2i$ so that the closed-loop system is purely oscillatory with two distinct frequencies (the frequency of the second mode being twice the frequency of the first mode).*

This problem seems quite trivial at first sight, but looking closer we encounter a difficulty. The maximum cyclicity index of the A matrix is an odd number, namely 3 which is not a multiple of 2. This means that all possible combinations of the column vectors of the B matrix will at most achieve a cyclicity index of 3.

Unfortunately, complex poles can be assigned only by conjugate pairs and therefore cannot be embedded into a third order polynomial unless a purely real eigenvalue is accounted for, a condition clearly not allowed by the design considerations (i.e. imposing only complex conjugate eigenvalues).

The origin of this problem stems from the multiple eigenvalue 2 which accounts for the MAP (w.r.t. A) of the whole space $\mathbb{R}^4$ being

$$\psi(\lambda) = (\lambda - 2)(\lambda - 3)(\lambda - 4) = \lambda^3 - 9\lambda^2 + 26\lambda - 24 \tag{8.5}$$

a third order polynomial. Indeed,

$$\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 27 & 0 \\ 0 & 0 & 0 & 64 \end{pmatrix} - 9 \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix} + 26 \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$
$$- \begin{pmatrix} 24 & 0 & 0 & 0 \\ 0 & 24 & 0 & 0 \\ 0 & 0 & 24 & 0 \\ 0 & 0 & 0 & 24 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

since we have simultaneously $8-9\times4+26\times2-24=0$, $27-9\times9+26\times3-24=0$ and $64-9\times16+26\times4-24=0$.

So as to circumvent the difficulty of the maximum index being 3, we first construct a feedback $u=Kx$ so as to assign — after feedback — a new MAP (w.r.t. $A+KB$) of order 4. This is achieved by changing one of the eigenvalues 2 to another value which we choose equal to 1 (so as to be distinct from 2, 3 and 4). This automatically implies that for the transformed matrix $A+BK$, the existence of a cyclic vector of cyclicity index 4.

Since the original system is controllable, there is a nice theorem that states that there exists a vector in the span of the columns of the B matrix sharing the exact same MAP (w.r.t. $A+BK$) as a vector having maximum cyclicity index, i.e. there exists a combination of the columns of the B matrix that attain maximum cyclicity index of 4. This combination is then used to set the eigenvalues at the desired position by imposing that the MAP of the newly transformed system (through this second feedback) is the desired closed-loop characteristic polynomial equal to

$$\psi_d(\lambda)=(\lambda-i)(\lambda+i)(\lambda-2i)(\lambda+2i)=(\lambda^2+1)(\lambda^2+2)$$

Therefore, so as to initiate the first step of the design, let us choose the first Leangenhop vector equal to b_1. Since b_1 generates a subspace of cyclicity index of 3, this vector is set to

$$e_3:=b_1=\begin{pmatrix}0&1&1&1\end{pmatrix}^T$$

Then define inductively according to the scheme (8.1)

$$e_2 \quad := \quad Ae_1-9b_1=\begin{pmatrix}0&-7&-6&-5\end{pmatrix}^T \tag{8.6}$$

$$e_1 \quad := \quad Ae_2+26b_1=\begin{pmatrix}0&12&8&6\end{pmatrix}^T \tag{8.7}$$

Notice the appearance of all the coefficients of the MAP (w.r.t. A) of the vector b_1 except the last coefficient (first coefficients of the polynomial (8.5)) in the inductive computations (8.6)-(8.7).

Then we pick a complementary vector v in the range of the B matrix so that the family of vectors $\{e_1,e_2,e_3,v\}$ becomes linearly independent. For example setting

$$v:=b_2=\begin{pmatrix}1&0&1&1\end{pmatrix}^T$$

allows the change of basis defined through the matrix

$$P = \begin{pmatrix} e_1 & e_2 & e_3 & v \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 12 & -7 & 1 & 0 \\ 8 & -6 & 1 & 1 \\ 6 & -6 & 1 & 1 \end{pmatrix}$$

to transform the initial matrix A to

$$\bar{A} = P^{-1}AP = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 24 & -26 & 9 & 7 \\ 0 & 0 & 0 & 2 \end{pmatrix} \tag{8.8}$$

the structure of which is

$$\begin{pmatrix} 0 & 1 & 0 & X \\ 0 & 0 & 1 & X \\ -\alpha_3 & -\alpha_2 & -\alpha_1 & X \\ 0 & 0 & 0 & -\beta \end{pmatrix} \tag{8.9}$$

where $\alpha_i, i = 1, \ldots, 3$ are the coefficients of the MAP of $b_1 = e_3$ (which is $\lambda^3 + \alpha_1\lambda^2 + \alpha_2\lambda + \alpha_3$) and where β is the coefficient of the polynomial $\lambda + \beta$ which the RMAP (w.r.t. A) of v (mod span $\{e_1, e_2, e_3\}$). Notice that because $\lambda + \beta$ is not the absolute AP but only the relative one, do — not necessarily zero values — appear in the upper part of the last column of the transformed matrix in (8.9), where the symbol X accounts for any number which is not necessarily zero.

Concerning the effect of P on the input matrix, because both, $P^{-1}P = I$, and P contains, in its third column position, b_1 and, in its fourth column position, b_2, the transformed B matrix is

$$\bar{B} = P^{-1}B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.10}$$

Let us construct a first feedback so as to change the last eigenvalue 2 to the value 1, distinct from the three first eigenvalues 2, 3, and 4. This is achieved after setting

$$\bar{K} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

So that after applying the feedback

$$u = -\bar{K}z + v$$

to the transformed system

$$\dot{z} = \bar{A}z + \bar{B}u$$

we have a new system

$$\dot{z} = (\bar{A} + \bar{B}\bar{K})z + \bar{B}v = \tilde{A}z + \bar{B}v$$

with distinct four eigenvalues with the matrix

$$\tilde{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 24 & -26 & 9 & 7 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{8.11}$$

Now want to find a vector in the range of the $\bar{B}$ matrix having cyclicity index of 4. This is achieved for the vector

$$v_4 = \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}^T$$

since the four following vectors

$$\{v_4, \tilde{A}v_4, \tilde{A}^2v_4, \tilde{A}^3v_4\} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 16 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 17 \\ 99 \\ 1 \end{pmatrix}, \begin{pmatrix} 17 \\ 100 \\ 504 \\ 1 \end{pmatrix} \right\}$$

are linearly independent.

Let us compute the Langenhop vectors associated with v_4 and $\tilde{A}$ characteristic polynomial

$$\begin{aligned} (\lambda-1)(\lambda-2)(\lambda-3)(\lambda-4) &= \lambda^4 - 10\lambda^3 + 35\lambda^2 - 50\lambda + 24 \\ &= \lambda^4 + \beta_1\lambda^3 + \beta_2\lambda^2 + \beta_3\lambda + \beta_4 \end{aligned}$$

which gives successively

$$\begin{aligned} \tilde{e}_4 &= v_4 = \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}^T \\ \tilde{e}_3 &= \tilde{A}e_4 - \beta_3 v_4 = \tilde{A}e_4 + 50v_4 = \begin{pmatrix} 0 & 2 & 6 & -9 \end{pmatrix}^T \\ \tilde{e}_2 &= \tilde{A}e_3 - \beta_2 v_4 = \tilde{A}e_3 - 35v_4 = \begin{pmatrix} 2 & -3 & -26 & 26 \end{pmatrix}^T \\ \tilde{e}_1 &= \tilde{A}e_2 - \beta_1 v_4 = \tilde{A}e_2 + 10v_4 = \begin{pmatrix} -3 & 0 & 24 & -24 \end{pmatrix}^T \end{aligned}$$

Therefore, after defining the new change of variables

$$\tilde{P} = \begin{pmatrix} \tilde{e}_1 & \tilde{e}_2 & \tilde{e}_3 & \tilde{e}_4 \end{pmatrix} = \begin{pmatrix} -3 & 2 & 0 & 0 \\ 0 & -3 & 2 & 0 \\ 24 & -26 & 6 & 1 \\ -24 & 26 & -9 & 1 \end{pmatrix}$$

brings the $\tilde{A}$ matrix into the canonical form

$$\hat{A} = \tilde{P}^{-1}\bar{A}\tilde{P} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -24 & 50 & -35 & 10 \end{pmatrix}$$

where the characteristic polynomial coefficients are clearly visible. Now, define the desired characteristic polynomial as

$$(\lambda^2+1)(\lambda^2+4) = \lambda^4 + 5\lambda^2 + 1 = \lambda^4 + \gamma_1\lambda^3 + \gamma_2\lambda^2 + \gamma_3\lambda + \gamma_4$$

and set the second feedback to

$$\hat{K} = \begin{pmatrix} \beta_4-\gamma_4 & \beta_3-\gamma_3 & \beta_2-\gamma_2 & \beta_1-\gamma_1 \end{pmatrix}^T = \begin{pmatrix} 20 & -50 & 30 & -10 \end{pmatrix}^T$$

so that

$$\hat{A} + \hat{B}\hat{K} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -4 & 0 & -5 & 0 \end{pmatrix}$$

with the required desired characteristic polynomial. Summing up, we have the following feedback to be applied to the original control system,

$$\begin{aligned} K_t &= \bar{K}P^{-1} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \hat{K}\tilde{P}^{-1}P^{-1} \\ &= \begin{pmatrix} -\frac{26}{3} & -20 & \frac{130}{3} & -34 \\ -\frac{29}{3} & -20 & \frac{130}{3} & -34 \end{pmatrix} \end{aligned} \tag{8.12}$$

and we check indeed that this feedback assigns the required eigenvalues since

$$\lambda(A + K_tB) = \{i, -i, 2i, -2i\}$$

Formula (8.12) is justified after considering the following sequences of systems and the change of coordinates associated with them (i.e. those defined successively by P and $\tilde{P}$ already computed).

$$\begin{aligned}
\dot{x} &= Ax + Bu \\
z &= P^{-1}x \\
\dot{z} &= P^{-1}APz + P^{-1}Bu \\
u &= \bar{K}x + \begin{pmatrix} 1 \\ 1 \end{pmatrix} v \\
w &= \tilde{P}^{-1}z \\
\dot{w} &= \tilde{P}^{-1}\tilde{A}\tilde{P}w + \tilde{P}^{-1}\tilde{B}v \\
v &= \hat{K}w + v_2 \\
\dot{w} &= \left(\tilde{P}^{-1}\tilde{A}\tilde{P} + \tilde{P}^{-1}\tilde{B}\hat{K}\right) w + \tilde{P}^{-1}\tilde{B}v
\end{aligned}$$

Considering only what affects the inputs v and u

$$\begin{aligned}
v &= \hat{K}w + v_2 \\
&= \hat{K}\tilde{P}^{-1}z + v_2 \\
&= \hat{K}\tilde{P}^{-1}P^{-1}x + v_2 \\
u &= \bar{K}z + \begin{pmatrix} 1 \\ 1 \end{pmatrix} v \\
&= \left(\bar{K}P^{-1} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \hat{K}\tilde{P}^{-1}P^{-1}\right) x + v_2 \qquad (8.13)
\end{aligned}$$

which justifies formula (8.12).

Chapter 9

Pole Assignment and Model-Matching

9.1 Pole placement - SISO

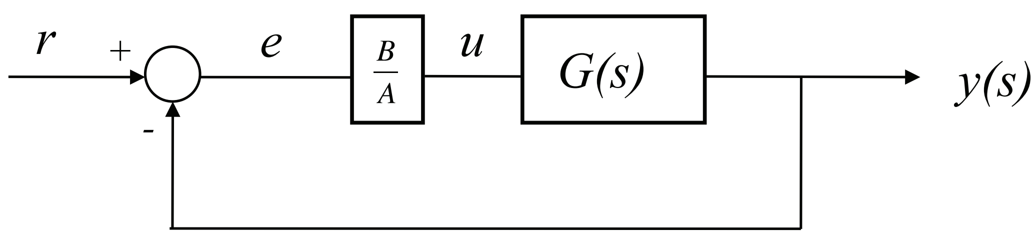


Figure 9.1:

- The purpose is to build controller B/A.

- $G(s) = \dfrac{N(s)}{D(s)}$ $\quad$ $\deg N = m \quad \deg D = n \quad m < n$

  $N(s) = N_0 + sN_1 + s^2N_2 + ... + s^m N_m \qquad N_i \in \mathbb{R}$
  $D(s) = D_0 + sD_1 + s^2D_2 + ... + s^n D_n \qquad D_i \in \mathbb{R}$

- Closed loop transfer function:

$$G_0 = \frac{\frac{B}{A}\cdot\frac{N}{D}}{1+\frac{B}{A}\cdot\frac{N}{D}} = \frac{Y}{r} = \frac{BN}{AD+BN}$$

we can't change the zeros, but can change poles

- In some ways, this is more useful that state-space since we dont need to measured all the states.
- We want to enforce:

$$G_0 = \frac{E}{F} = \frac{BN}{AD+BN}$$

$E = BN$ and $F = AD + BN$ have to be solved for A and B

- $F = F_0 + sF_1 + s^2F_2 + ...?$ we don't know where to stop
  $F = AD + BN$
  $A = A_0 + sA_1 + s^2A_2 + ...?$
  $B = B_0 + sB_1 + s^2B_2 + ...?$
  $F_0 = A_0D_0 + B_0N_0$
  $F_1 = A + 0D_1 + A_1D_0 + B_0N_1 + B_1N_0$
  $F_2 = A_0D_2 + A_1D_1 + A_2AD_0 + B_0N_2 + B_1N_1 + B_2N_0$
  $\vdots$

- Sylvester matrix (N, D initial data) multiplied by the unknowns (A, B) gives what we want to assign (F):

$$S\begin{pmatrix} \\ \\ \end{pmatrix} = \begin{pmatrix} \\ F \\ \\ \end{pmatrix}$$

- How to build the S and calculate A, B:

$$[A_0\ B_0\ A_1\ B_1\ A_2\ B_2\ ...]\begin{pmatrix} D_0 & D_1 & D_2 & \dots & D_n & 0 \\ N_0 & N_1 & N_2 & \dots & N_n & 0 \\ 0 & D_0 & D_1 & \dots & \dots & 0 \\ 0 & 0 & D_0 & \dots & \dots & 0 \\ \vdots & 0 & N_0 & \dots & \dots & 0 \\ \vdots & 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & \dots & 0 \end{pmatrix} = [F_0\ F_1\ F_2\ ...\ F_{n+1}]$$

- Assume that $\deg A = l$ and $\deg B = l$.
  $\dim S = 2(l+1) \times (n+l+1)$

  Distinguish between different cases:
  $2(l+1) > n+l+1 \quad AB = F * pinv(s)$
  $2(l+1) = n+l+1 \quad AB = FS^{-1}$
  $2(l+1) < n+l+1 \quad$ not useful
  We can choose l which is a degree of compensator, but can't change n which is a degree of initial transfer function.
  As we change (increase) l, the first case will become dominant.

9.2 Disturbance rejecting

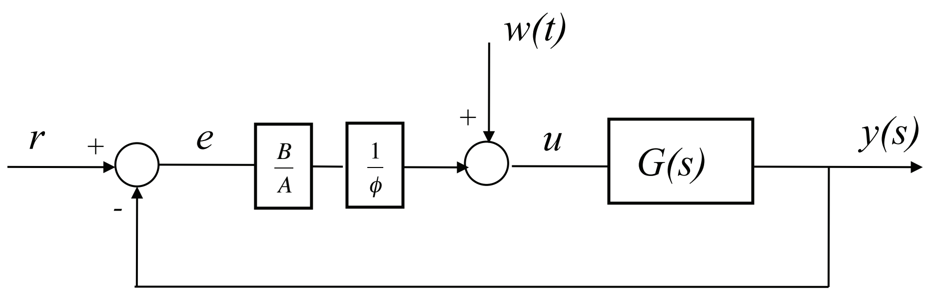


Figure 9.2:

- Perfect tracking with perfect knowledge and "no control" on zeros.

- Assume disturbance entering (the $1/\Phi$ is used to compensate for the disturbance).

- Relations to compensate for reference tracking and disturbance rejection:

$$G_{YW} = \frac{Y(s)}{W(s)} = \frac{G(s)}{1 + G(s)\frac{B}{A}\frac{1}{\Phi}} = \frac{\frac{N}{D}}{1 + \frac{N}{D}\frac{B}{A}\frac{1}{\Phi}} = \frac{NA\Phi}{DA\Phi + BN}$$

$$G_{YR} = \frac{Y(s)}{R(s)} = \frac{\frac{B}{A\Phi}\frac{N}{D}}{1 + \frac{B}{A\Phi}\frac{N}{D}} = \frac{BN}{A\Phi D + BN}$$

- Assume that we know $w(t)$ and thus $W(s)$:

$$W(s) = \frac{N_W}{D_W} \qquad Y(s) = \frac{NA\Phi}{DA\Phi + BN}\frac{N_W}{D_W}$$

poles of the perturbation can be offset with the poles of the model.

9.3 Internal model principle

- $D_W = \Phi$

$$G_{YE} = \frac{Y(s)}{Y(s) - r(s)} \qquad e = (1 - G_{YR})r = \frac{A\Phi B + BN - BN}{A\Phi D + BN}\frac{N_R}{D_R}$$

If we want to take care of both, then the composed Φ is $\Phi = \Phi_R \Phi_W$.

- Degree of l tends to dumb quickly. The trick is to put only the least possible information in Φ (the least common multiple of the unstable poles of $\frac{N_W}{D_W}$ and $\frac{N_r}{D_r}$).

- Idea is to compensate only the unstable poles (the stable part will die out naturally).

- Solving for the Sylvester:

$$A\Phi D + BN = F \qquad \tilde{D} = \Phi D \qquad A\tilde{D} + BN = F$$

solve as previously.

9.4 Model matching

- Second degree freedom controller.

- $C_1 = \dfrac{L}{A_1}$ compensator $\qquad C_2 = \dfrac{M}{A_2} \qquad A = A_1 = A_2$

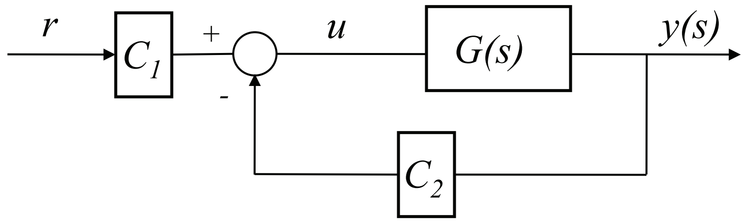


Figure 9.3:

so, three polynomials to deal with L, M, A.

- Now we want to match the whole transfer function, not just the poles and we impose an arbitrary:

$$G_0 = \frac{Y(s)}{R(s)} = C_1 \cdot \frac{G(s)}{1 + C_2 G(s)} = \frac{L}{A} \cdot \frac{\frac{N}{D}}{1 + \frac{M}{A}\frac{N}{D}} = \frac{LN}{AD + MN} = \frac{N_0}{D_0}$$

- Restrictions:

1) pole-zero excess inequality:
deg$D-$deg$N- \leq$ degD_0- degN_0 (should never decrease)
2) must retain "unstable" zeros (zeros in right-half plane of $N(s)$), as cancelling them will introduce unstable poles.

$$\frac{G_0}{N} = \frac{N_0}{ND_0} = \frac{\bar{N}_0}{\bar{D}_0} \qquad G_0 = \frac{N\bar{N}_0}{\bar{D}_0}$$

$$\frac{L}{AD + MN} = \frac{\bar{N}_0}{\bar{D}_0} \;\Leftrightarrow\; \frac{LN}{AD + MN} = \frac{\bar{N}_0 N}{\bar{D}_0}$$

$$LN = \bar{N}_0 N \qquad AD + MN = \bar{D}_0$$

We know hoe to solve it using Sylvester matrix.

Tempted to set $L = \bar{N}_0$, could possibly induce non-properness of $C_1 (C_1 = L/A)$, because deg(L) might be greater than deg(A).

l can be increased by adding a "dummy" polynomial $\hat{D}$, so that:

$$AD + MN = \bar{D}_0\hat{D} \quad \text{and} \quad \frac{LN}{AD + MN} = \frac{\bar{N}_0 N}{\bar{D}_0\hat{D}}.$$

- Theres no guarantee for the compensators being stable in the above procedure (the elements are, but the whole setup is not, due to the possible instability of A).

- Alternative I:

  The close-loop will stabilize the system even if L/A is unstable.
  Too many states (state-space for L/A and M/L).
  Might not be proper ($deg(M) > deg(L)$).

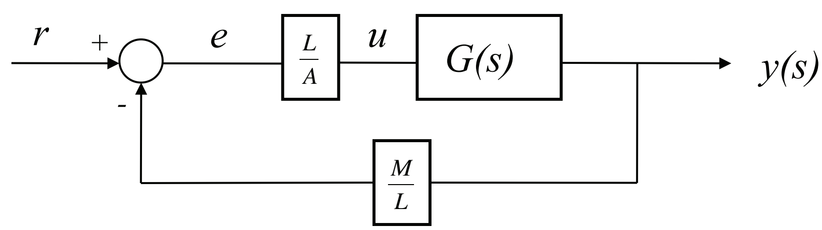


Figure 9.4:

- Alternative II:

  Assumption: M is stable.
  It is guaranteed to be proper.

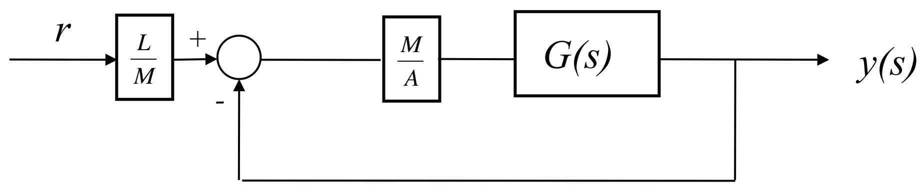


Figure 9.5:

- Alternative III:

  This is more efficient, as it needs less states.
  Noise gets amplified by M, however.

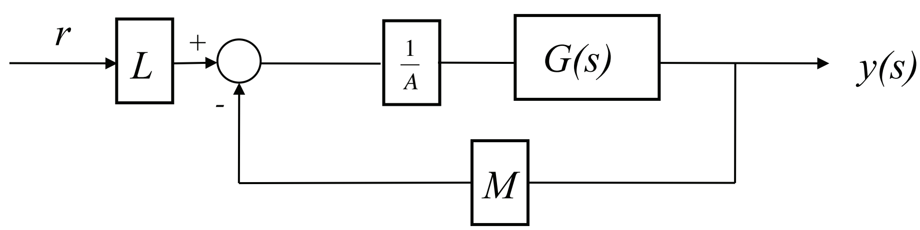


Figure 9.6:

- No leakage: all closed-loop signals, transfer functions, should be stable.

- So, all the 3 cases are better than the initial one, but then the individual ones differ by properties.

9.5 MIMO setting

- Transfer function matrix with p inputs and q outputs:

$$G(s) = \begin{pmatrix} \frac{n_{11}(s)}{d(s)} & \frac{n_{12}(s)}{d(s)} & \cdots & \frac{n_{1p}(s)}{d(s)} \\ \frac{n_{21}(s)}{d(s)} & \frac{n_{22}(s)}{d(s)} & \cdots & \frac{n_{2p}(s)}{d(s)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{n_{q1}(s)}{d(s)} & \frac{n_{q2}(s)}{d(s)} & \cdots & \frac{n_{qp}(s)}{d(s)} \end{pmatrix} = \bar{D}^{-1}(s)\bar{N}(s) = N(s)D^{-1}(s)$$

  $\bar{D}(s)$ is unimodular ($det(\bar{D}(s))$ doesn't depend on s)
  $\bar{N}(s)$ doesn't need to be unimodular.

- The question is if we have matrix description of $G(s)$, how do we get the factorization.

- Objective is to find unimodular $N(s)$ and $\bar{N}(s)$, and $D(s)$ and $\bar{D}(s)$,

- $G(s) = N(s)D^{-1}(s) = \bar{D}^{-1}(s)\bar{N}(s)$
  multiply with $\bar{D}$ on the left and D on the right

  $\bar{D}(s)N(s)D^{-1}(s)D(s) = \bar{D}(s)\bar{D}^{-1}(s)\bar{N}(s)D(s)$

  $\bar{D}(s)N(s) = \bar{N}(s)D(s)$, advantageous because $\bar{D}(s)$ and $\bar{N}(s)$ are known.

- Bezout equation:

  $\bar{D}(s)N - \bar{N}(s)D = 0$

- $\bar{N}(s) = \bar{N}_0 + s\bar{N}_1 + s^2\bar{N}_2 + ...$ $\quad \bar{N}_i \in \mathbb{R}_{q \times p}$
  $\bar{D}(s) = \bar{D}_0 + s\bar{D}_1 + s^2\bar{D}_2 + ...$ $\quad \bar{D}_i \in \mathbb{R}_{q \times p}$

- $\bar{D}(s)N(s) - \bar{N}(s)D(s) = 0$

$$
\begin{pmatrix}
\bar{D}_0 & \bar{N}_0 & 0 & 0 & 0 & 0 \\
\bar{D}_1 & \bar{N}_1 & \bar{D}_0 & \bar{N}_0 & 0 & 0 \\
\bar{D}_2 & \bar{N}_2 & \bar{D}_1 & \bar{N}_1 & \bar{D}_0 & \bar{N}_0 \\
\vdots & \vdots & \bar{D}_2 & \bar{N}_2 & \bar{D}_1 & \bar{N}_1 \\
\bar{D}_n & 0 & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}
\begin{pmatrix}
-N_0 \\ D_0 \\ -N_1 \\ D_1 \\ -N_2 \\ D_2 \\ \vdots
\end{pmatrix} = 0
$$

all elements belong to $\mathbb{R}$.

- The question is how to solve this system.

- We should built the first matrix to the dimension of n, until we get full rank.

- Get different solutions by first taking one column, then two, etc...

- We're looking for p solutions (width of second matrix).

- Need a $n \times q$ rank matrix for the D's.

- In order to get N's we should keep increasing column index until rank deficiency is reached.

Example Testing loss of ranks

Try to use qr() in MATLAB©

$A = QR$, Q is full rank, R is rank deficient.
The rank of A is the number of nonzero entries in the main diagonal of A.
We need to perform QR at each iteration to find where rank is lost.
Keep the number of columns that preserve rank.
Alternative: look for zeros that appear in the SVD.

Chapter 10

Polynomial system matrix

- If state-space (A, B, C, D) is a minimal representation of $G(s)$, then the $deg(G(s)) = dim(x) = n$

- $G(s) = C(sI - A)^{-1}B = \bar{D}(s)^{-1}N^{-1} = N(s)D(s)^{-1}$

  where $C(q, n)$ and $B(p, n)$ are constants polynomial matrices.

- $G(s) = V(s)T^{-1}(s)U(s) + W(s)$
  special case $T(s) = sI + A$

- Implicit linear system

  $T(s)\cdot\xi = \mathbb{U}(s)\cdot U(s) \qquad deg(T(s)) > 1 \qquad Y(s) = V(s)\cdot\xi + W(s)\cdot U(s)$
  $\curvearrowright$ $\mathbb{U}$ is a matrix of inputs

- Explicit linear system

  $(sI - A)X(s) = B \cdot U(s) \qquad T(s)(r, r)$

- $P(s) \equiv \begin{pmatrix} T(s) & -U(s) \\ V(s) & W(s) \end{pmatrix}$ convention for system matrix

- **Example**

Compute unimodular matrices $N(s)$ and $D(s)$ such that $G(s) = N(s)D^{-1}(s)$

$$G(s) = \begin{pmatrix} \dfrac{1}{(s+1)^2} & \dfrac{1}{(s+1)(s+2)} \\ \dfrac{1}{(s+1)(s+2)} & \dfrac{s+3}{(s+2)^2} \end{pmatrix}$$

$$P(s) = \left(\begin{array}{ccc|cc} I_2 & 0 & 0 & 0 & 0 \\ 0 & (s+1)^2(s+2) & 0 & s+2 & s+1 \\ 0 & -(s+1) & s+2 & 0 & 1 \\ \hline 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{array}\right)$$

$1^\circ\ T(s) \qquad 2^\circ\ -U(s) \qquad 3^\circ\ V(s) \qquad 4^\circ\ W(s)$

- First check that it gives $G(s)$.
  - Find a state-space.
  - Compute the Smith form of $T(s)$.
  - Compute the Smith form of $P(s)$.
  - Where are the poles and zeros? • Compare with $G(s)$.

- System matrix:

State-space: $\dot{x} = Ax + Bu \qquad y = Cx + Du$

$$\begin{pmatrix} sI - A & -B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

$P(s)$ - polynomial matrix of special kind - system matrix

$$T(s) = sI - A \qquad V(s) = -B \qquad U(s) = C \qquad W(s) = D$$

• Be aware that it needs to be 1st order, because of the state-space representation.

$$G(s) = C(sI - A)^{-1}B + D \qquad G(s) = U(s)T^{-1}(s)V(s) + W(s)$$

- Smith form

  $P(s) = M_{P1}\Lambda_p M_{P2}$ $\qquad$ M_i are two unimodular matrices, $det M_i \in \mathbb{R}$

  $T(s) = M_{T1}\Lambda_p M_{T2}$

- $\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & p_1(s) & 0 & 0 \\ 0 & 0 & p_2(s) & 0 \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$ $\qquad$ $\Lambda_T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & t_1(s) & 0 & 0 \\ 0 & 0 & t_2(s) & 0 \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$

  $\xi_1 = p_1(s) \qquad \xi_2 = p_2(s) \qquad \xi_3 = p_3(s) \qquad \dots$
  $\psi_n = t_1(s) \qquad \psi_{n-1} = t_2(s) \qquad \psi_{n-2} = t_3(s) \qquad \dots$

- $G(s) = \bar{D}^{-1}\bar{N}$

  $$\bar{D} = \begin{vmatrix} (1+s)^2(2+s)^2 & 0 \\ 0 & (1+s)^2(2+s)^2 \end{vmatrix}$$

  $$\bar{N} = \begin{vmatrix} (2+s)^2 & (1+s)(2+s) \\ (1+s)(2+s) & (1+s)^2(3+s) \end{vmatrix}$$

- $\bar{N} = M_{\bar{N}_1}\Lambda_{\bar{N}} M_{\bar{N}_2} \qquad \Lambda_{\bar{N}} = C_1 \bar{N} C_2$
  $C_1 = M_{\bar{N}_1}^{-1} \qquad C_2 = M_{\bar{N}_2}^{-1}$

- $\Lambda_{\bar{N}} = \begin{pmatrix} 1 & 0 \\ 0 & (1+s)^2(2+s)^3 \end{pmatrix}$

- $$\theta = \frac{\Lambda_{\bar{N}}}{\bar{D}} = \begin{pmatrix} \dfrac{1}{(s+1)^2(s+2)^2} & 0 \\ 0 & s+2 \end{pmatrix}$$