Appendix A
Background Material in Linear Systems Theory

A.1 Quadratic forms

In this section, a few fundamental facts about symmetric matrices and quadratic forms are reviewed.

Symmetric matrices. Let P be a $n \times n$ *symmetric* matrix of real numbers (that is, a matrix of real numbers satisfying $P = P^{\mathrm{T}}$). Then there exist an *orthogonal* matrix Q of real numbers[1] and a diagonal matrix Λ of real numbers

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

such that

$$Q^{-1}PQ = Q^{\mathrm{T}}PQ = \Lambda\,.$$

Indeed, the numbers $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of P. Thus, a symmetric matrix P of real numbers has real eigenvalues and a purely diagonal Jordan form.

Note that the previous identity can be rewritten as

$$PQ = Q\Lambda$$

from which it is seen that the i-th column q_i of Q is an eigenvector of P, associated with the i-th eigenvalue λ_i. If P is invertible, so is the matrix Λ, and

$$P^{-1}Q = Q\Lambda^{-1}\,,$$

from which it is seen that Λ^{-1} is a Jordan form of P^{-1} and the i-th column q_i of Q is also an eigenvector of P^{-1}, associated with the i-th eigenvalue λ_i^{-1}.

[1] That is, matrix Q of real numbers satisfying $QQ^{\mathrm{T}} = I$, or – what is the same – satisfying $Q^{-1} = Q^{\mathrm{T}}$.

Quadratic forms. Let P be a $n \times n$ matrix of real numbers and $x \in \mathbb{R}^n$. The expression

$$V(x) = x^{\mathrm{T}} P x = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_i x_j$$

is called a *quadratic form* in x. Without loss of generality, in the expression above we may assume that P is *symmetric*. In fact,

$$V(x) = x^{\mathrm{T}} P x = \frac{1}{2}[x^{\mathrm{T}} P x + x^{\mathrm{T}} P x] = x^{\mathrm{T}}[\frac{1}{2}(P + P^{\mathrm{T}})]x$$

and $\frac{1}{2}(P + P^{\mathrm{T}})$ is by construction symmetric.

Let P be symmetric, express it as $P = Q \Lambda Q^{\mathrm{T}}$ (see above) with eigenvalues sorted so that $\lambda_1 \geq \cdots \geq \lambda_n$. Then

$$\begin{aligned} x^{\mathrm{T}} P x &= x^{\mathrm{T}} (Q \Lambda Q^{\mathrm{T}}) x = (Q^{\mathrm{T}} x)^{\mathrm{T}} \Lambda (Q^{\mathrm{T}} x) = \textstyle\sum_{i=1}^{n} \lambda_i (Q^{\mathrm{T}} x)_i^2 \\ &\leq \lambda_1 \textstyle\sum_{i=1}^{n} (Q^{\mathrm{T}} x)_i^2 = \lambda_1 (Q^{\mathrm{T}} x)^{\mathrm{T}} Q^{\mathrm{T}} x = \lambda_1 x^{\mathrm{T}} Q Q^{\mathrm{T}} x = \lambda_1 x^{\mathrm{T}} x \\ &= \lambda_1 \|x\|^2 . \end{aligned}$$

With a similar argument we can show that $x^{\mathrm{T}} P x \geq \lambda_n \|x\|^2$. Usually, λ_n is denoted as $\lambda_{\min}(P)$ and λ_1 is denoted as $\lambda_{\max}(P)$. In summary, we conclude that

$$\lambda_{\min}(P) \|x\|^2 \leq x^{\mathrm{T}} P x \leq \lambda_{\max}(P) \|x\|^2 .$$

Note that the inequalities are tight (*hint*: pick, as x, the last and, respectively, the first column of Q).

Sign-definite symmetric matrices. Let P be symmetric. The matrix P is said to be *positive semi-definite* if

$$x^{\mathrm{T}} P x \geq 0 \qquad \text{for all } x .$$

The matrix P is said to be *positive definite* if

$$x^{\mathrm{T}} P x > 0 \qquad \text{for all } x \neq 0 .$$

We see from the above that P is positive semi-definite if and only if $\lambda_{\min} \geq 0$ and is positive definite if and only if $\lambda_{\min} > 0$ (which in turn implies the nonsingularity of P).

There is another criterion for a matrix to be positive definite, that does not require the computation of the eigenvalues of P, known as Sylvester's criterion. For a *symmetric* matrix P, the n minors

$$D_1 = \det(p_{11}), \quad D_2 = \det \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}, \quad D_3 = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}, \ldots$$

are called the *leading principal minors*. Note that $D_n = \det(P)$.

Lemma A.1. *A symmetric matrix is positive definite if and only if all leading principal minors are positive, i.e.* $D_1 > 0,\ D_2 > 0, \dots,\ D_n > 0$.

Another alternative criterion, suited a for block-partitioned matrix, is the criterion due to Schur.

Lemma A.2. *The symmetric matrix*

$$\begin{pmatrix} Q & S \\ S^{\mathrm{T}} & R \end{pmatrix} \tag{A.1}$$

is positive definite if and only if [2]

$$R > 0 \qquad \text{and} \qquad Q - SR^{-1}S^{\mathrm{T}} > 0\,. \tag{A.2}$$

Proof. Observe that a necessary condition for (A.1) to be positive definite is $R > 0$. Hence R is nonsingular and (A.1) can be transformed, by congruence, as

$$\begin{pmatrix} I & 0 \\ -R^{-1}S^{\mathrm{T}} & I \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} Q & S \\ S^{\mathrm{T}} & R \end{pmatrix} \begin{pmatrix} I & 0 \\ -R^{-1}S^{\mathrm{T}} & I \end{pmatrix} = \begin{pmatrix} Q - SR^{-1}S^{\mathrm{T}} & 0 \\ 0 & R \end{pmatrix}.$$

from which the condition (A.2) follows. $\triangleleft$

A symmetric matrix P is said to be *negative semi-definite* (respectively, *negative definite*) if $-P$ is *positive semi-definite* (respectively, *positive definite*). Usually, to express the property (of a matrix P) of being positive definite (respectively, positive semi-definite) the notation $P > 0$ (respectively, $P \geq 0$) is used. [3] Likewise, the notation $P < 0$ (respectively $P \leq 0$) is used to express the property that P is negative definite (respectively, negative semi-definite). If P and R are symmetric matrices, the notations

$$P \geq R \qquad \text{and} \qquad P > R$$

stand for "*the matrix* $P - R$ *is positive semi-definite*" and, respectively, for "*the matrix* $P - R$ *is positive definite*".

Any matrix P that can be written in the form $P = M^{\mathrm{T}}M$, in which M is a possibly non-square matrix, is positive semi-definite. In fact $x^{\mathrm{T}}Px = x^{\mathrm{T}}M^{\mathrm{T}}Mx = \|Mx\|^2 \geq 0$.

Conversely, any matrix P which is positive semi-definite can always be expressed as $P = M^{\mathrm{T}}M$. In fact, if P is positive semi-definite all its eigenvalues are non-negative. Let r denote the number of nonzero eigenvalues and let the eigenvalues be sorted so that $\lambda_{r+1} = \cdots = \lambda_n = 0$. Then

---

[2] The matrix $Q - SR^{-1}S^{\mathrm{T}}$ is called the *Schur's complement* of R in (A.1).

[3] Note that this *is not* the same as $p_{ij} > 0$ for all i, j. For example, the matrix

$$P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

in which the off-diagonal elements are negative, is positive definite (use Sylvester's criterion above).

$$P = Q\Lambda Q^{\mathrm{T}} = Q \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix} Q^{\mathrm{T}}$$

in which Λ_1 is an $r \times r$ diagonal matrix, whose diagonal elements are all positive. For $i = 1, \dots, r$, let σ_i denote the positive square root of λ_i, let Q_1 be the $n \times r$ matrix whose columns coincide with the first r columns of Q and set

$$M = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} Q_1^{\mathrm{T}}$$

Then, the previous identity yields

$$P = M^{\mathrm{T}} M$$

in which M is a $n \times r$ matrix of rank r.

Finally, recall that, if P is symmetric and invertible, the eigenvalues of P^{-1} are the inverse of the eigenvalues of P. Thus, in particular, if P is positive definite, so is also P^{-1}.

A.2 Linear matrix equations

In this section, we discuss the existence of solutions of two relevant linear matrix equations that arise in the analysis of linear systems. One of such equation is the so-called *Sylvester's equation*

$$AX - XS = R \tag{A.3}$$

in which $A \in \mathbb{R}^{n \times n}$ and $S \in \mathbb{R}^{d \times d}$, in the unknown $X \in \mathbb{R}^{n \times d}$. An equation of this kind arises, for instance, when it is desired to transform a given block-triangular matrix into a (purely) block-diagonal one, by means of a similarity transformation, as in

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} A & R \\ 0 & S \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix}.$$

Another instance in which an equation of this kind arises is the analysis of the stability of a linear system, where this equation assumes the special form $AX + XA^{\mathrm{T}} = Q$, known as *Lyapunov's equation*.

Another relevant linear matrix equation is the so-called *regulator* or *Francis's equation*

$$\begin{aligned} \Pi S &= A\Pi + B\Psi + P \\ 0 &= C\Pi + Q \end{aligned} \tag{A.4}$$

in which $A \in \mathbb{R}^{n \times n}$, $S \in \mathbb{R}^{d \times d}$, $B \in \mathbb{R}^{n \times m}$ and $C \in \mathbb{R}^{p \times n}$, in the unknowns $\Pi \in \mathbb{R}^{n \times d}$ and $\Psi \in \mathbb{R}^{m \times d}$. This equation arises in the study of the problem of output regulation of linear systems.

The two equations considered above are special cases of an equation of the form

$$A_1 X q_1(S) + \cdots + A_k X q_k(S) = R \tag{A.5}$$

in which, for $i = 1, \ldots, k$, $A_i \in \mathbb{R}^{\bar{n}\times\bar{m}}$ and $q_i(\lambda)$ is a polynomial in the indeterminate λ, $S \in \mathbb{R}^{\bar{d}\times\bar{d}}$, $R \in \mathbb{R}^{\bar{n}\times\bar{d}}$, in the *unknown* $X \in \mathbb{R}^{\bar{m}\times\bar{d}}$. In fact, the Sylvester's equation corresponds to the case in which $k = 2$ and

$$A_1 = A, \quad q_1(\lambda) = 1, \quad A_2 = I, \quad q_2(\lambda) = -\lambda,$$

while the Francis' equation corresponds to the case in which $k = 2$ and

$$A_1 = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix}, \quad q_1(\lambda) = 1, \quad A_2 = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad q_2(\lambda) = -\lambda, \quad R = \begin{pmatrix} -P \\ -Q \end{pmatrix}.$$

Equations of the form (A.5) are also known as *Hautus' equations*.[4] Noting that the left-hand side of (A.5) can be seen as a linear map

$$\begin{aligned} \mathscr{H} &: \mathbb{R}^{\bar{m}\times\bar{d}} \to \mathbb{R}^{\bar{n}\times\bar{d}} \\ &: \quad X \quad \mapsto \mathscr{H}(X) := A_1 X q_1(S) + \cdots + A_k X q_k(S), \end{aligned}$$

to say that (A.5) has a solution is to say that $R \in \mathrm{Im}(\mathscr{H})$.

In what follows, we are interested in the case in which (A.5) has solutions for all R, i.e. in the case in which the map $\mathscr{H}$ is *surjective*. [5]

Theorem A.1. *The map $\mathscr{H}$ is surjective if and only if the $\bar{n}$ rows of the matrix*

$$A(\lambda) = A_1 q_1(\lambda) + \cdots + A_k q_k(\lambda)$$

are linearly independent for each λ which is an eigenvalue of S. If this is the case and $\bar{n} = \bar{m}$, the solution X of (A.5) is unique.

From this, it is immediate to deduce the following Corollaries.

Corollary A.1. *The Sylvester's equation (A.3) has a solution for each R if and only if $\sigma(A) \cap \sigma(S) = \emptyset$. If this is the case, the solution X is unique.*

Corollary A.2. *The Francis' equation (A.4) has a solution for each pair (P, Q) if and only if the rows of the matrix*

$$\begin{pmatrix} A - \lambda I & B \\ C & 0 \end{pmatrix}$$

are linearly independent for each λ which is an eigenvalue of S. If this is the case and $m = p$, the solution pair (Π, Ψ) is unique.

---

[4] See [3].

[5] Note that, if this is the case and $\bar{n} = \bar{m}$, tha map is also *injective*, i.e. it is an invertible linear map. In this case the solution X of (A.5) is *unique*.

A.3 The theorems of Lyapunov for linear systems

In this section we describe a powerful criterion useful to determine when a $n \times n$ matrix of real numbers has all eigenvalues with negative real part.[6]

Theorem A.2. [Direct Theorem] *Let $A \in \mathbb{R}^{n\times n}$ be a matrix of real numbers. Let $P \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix of real numbers and suppose that the matrix*

$$PA + A^{\mathrm{T}}P$$

is negative definite. Then, all eigenvalues of the matrix A have negative real part.

Proof. Let λ be an eigenvalue of A and x an associated eigenvector. Let x_R and x_I denote the real and, respectively, imaginary part of x, i.e. set $x = x_R + jx_I$ and let $x^* = x_R^{\mathrm{T}} - jx_I^{\mathrm{T}}$. Then

$$x^*Px = (x_R)^{\mathrm{T}}Px_R + (x_I)^{\mathrm{T}}Px_I.$$

Since P is positive definite and x_R and x_I cannot be both zero (because $x \neq 0$), we deduce that

$$x^*Px > 0. \tag{A.6}$$

With similar arguments, since $A^{\mathrm{T}}P + PA$ is negative definite, we deduce that

$$x^*(A^{\mathrm{T}}P + PA)x < 0. \tag{A.7}$$

Using the definition of x and λ (i.e. $Ax = x\lambda$, that implies $x^*A^{\mathrm{T}} = \lambda^*x^*$), obtain

$$x^*(A^{\mathrm{T}}P + PA)x = \lambda^*x^*Px + x^*Px\lambda = (\lambda + \lambda^*)x^*Px$$

from which, using (A.6) and (A.7) we conclude

$$\lambda + \lambda^* = 2\mathrm{Re}[\lambda] < 0. \qquad \triangleleft$$

Remark A.1. The criterion described in the previous Theorem provides a sufficient condition under which all the eigenvalues of a matrix A have negative real part. In the analysis of linear systems, this criterion is used as a sufficient condition to determine whether the equilibrium $x = 0$ of the autonomous system

$$\dot{x} = Ax \tag{A.8}$$

is (globally) asymptotically stable. In this context, the previous proof – which only uses algebraic arguments – can be replaced by the linear version of the proof of Theorem B.1, which can be summarized as follows. Let $V(x) = x^{\mathrm{T}}Px$ denote the positive definite quadratic function associated with P, let $x(t)$ denote a generic trajectory of system (A.8) and consider the composite function $V(x(t)) = x^{\mathrm{T}}(t)Px(t)$. Observe that

[6] For further reading, see e.g. [1].

$$\frac{\partial V}{\partial x} = 2x^{\mathrm{T}}P.$$

Therefore, using the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}t}V(x(t)) = \frac{\partial V}{\partial x}\Big|_{x=x(t)}\frac{\mathrm{d}x}{\mathrm{d}t} = 2x^{\mathrm{T}}(t)PAx(t) = x^{\mathrm{T}}(t)(PA + A^{\mathrm{T}}P)x(t).$$

If P is positive definite and $PA + A^{\mathrm{T}}P$ is negative definite, there exists positive numbers a_1, a_2, a_3 such that

$$a_1\|x\|^2 \leq V(x) \leq a_2\|x\|^2 \qquad \text{and} \qquad x^{\mathrm{T}}(PA + A^{\mathrm{T}}P)x \leq -a_3\|x\|^2.$$

From this, it is seen that $V(x(t))$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}V(x(t)) \leq -\lambda V(x(t))$$

with $\lambda = a_3/a_2 > 0$ and therefore

$$a_1\|x(t)\|^2 \leq V(x(t)) \leq e^{-\lambda t}V(x(0)) \leq e^{-\lambda t}a_2\|x(0)\|^2.$$

Thus, for any initial condition $x(0)$, $\lim_{t\to\infty} x(t) = 0$. This proves that all eigenvalues of A have negative real part. $\triangleleft$

Theorem A.3. [Converse Theorem] *Let $A \in \mathbb{R}^{n\times n}$ be a matrix of real numbers. Suppose all eigenvalues of A have negative real part. Then, for any choice of a symmetric positive definite matrix Q, there exists a unique symmetric positive definite matrix P such that*

$$PA + A^{\mathrm{T}}P = -Q. \tag{A.9}$$

Proof. Consider (A.9). This is a Sylvester equation, and – since the spectra of A and $-A^{\mathrm{T}}$ are disjoint – a unique solution P exists. We compute it explicitly. Define

$$M(t) = e^{A^{\mathrm{T}}t}Qe^{At}$$

and observe that

$$\frac{\mathrm{d}M}{\mathrm{d}t} = A^{\mathrm{T}}e^{A^{\mathrm{T}}t}Qe^{At} + e^{A^{\mathrm{T}}t}Qe^{At}A = A^{\mathrm{T}}M(t) + M(t)A.$$

Integrating over $[0,T]$ yields

$$M(T) - M(0) = A^{\mathrm{T}}\int_0^T M(t)dt + \int_0^T M(t)dt\, A.$$

Since the eigenvalues of A have negative real part,

$$\lim_{T\to\infty} M(T) = 0$$

and

$$P := \lim_{T\to\infty} \int_0^T M(t)dt < \infty\,.$$

We have shown in this way that P satisfies (A.9). It is the unique solution of this equation.

To complete the proof it remains to show that P is positive definite, if so is Q. By contradiction, suppose is not. Then there exists $x_0 \neq 0$ such that

$$x_0^{\mathrm{T}} P x_0 \leq 0\,,$$

which, in view of the expression found for P, yields

$$\int_0^\infty x_0^{\mathrm{T}} e^{A^{\mathrm{T}}t} Q e^{At} x_0 dt \leq 0\,.$$

Setting

$$x(t) = e^{At} x_0$$

this is equivalent to

$$\int_0^\infty x^{\mathrm{T}}(t) Q x(t) dt \leq 0\,,$$

which, using the estimate $x^{\mathrm{T}} Q x \geq \lambda_{\min}(Q)\|x\|^2$ in turn yields

$$\int_0^\infty \|x(t)\|^2 dt \leq 0$$

which then yields

$$x(t) = 0, \qquad \text{for all } t \in [0,\infty).$$

Bearing in mind the expression of $x(t)$, this implies $x_0 = 0$ and completes the proof.

A.4 Stabilizability, detectability and separation principle

In this section, a few fundamental facts about the stabilization of linear systems are reviewed.[7] Consider a linear system modeled by equations of the form

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \tag{A.10}$$

with state $x \in \mathbb{R}^n$, input $u \in \mathbb{R}^m$, output $y \in \mathbb{R}^p$ and we summarize the properties that determine the existence of a (dynamic) output feedback controller of the form

$$\begin{aligned} \dot{x}_{\mathrm{c}} &= A_{\mathrm{c}} x_{\mathrm{c}} + B_{\mathrm{c}} y \\ u &= C_{\mathrm{c}} x_{\mathrm{c}} + D_{\mathrm{c}} y, \end{aligned} \tag{A.11}$$

[7] For further reading, see e.g. [1].

with state $x_c \in \mathbb{R}^{n_c}$, that stabilizes the resulting closed-loop system

$$\begin{aligned}\dot{x} &= (A+BD_cC)x+BC_cx_c \\ \dot{x}_c &= B_cCx+A_cx_c.\end{aligned} \tag{A.12}$$

Definition A.1. The pair (A,B) is *stabilizable* if there exists a matrix F such that $(A+BF)$ has all the eigenvalues in $\mathbb{C}^-$.

Definition A.2. The pair (A,C) is *detectable* if there exists a matrix G such that $(A-GC)$ has all the eigenvalues in $\mathbb{C}^-$.

Noting that the closed-loop system (A.12) can be written as $\dot{x}_{c\ell} = A_{c\ell}x_{c\ell}$, with $x_{c\ell} = \mathrm{col}(x,x_c)$ and

$$A_{c\ell} = \begin{pmatrix} (A+BD_cC) & BC_c \\ B_cC & A_c \end{pmatrix}, \tag{A.13}$$

we have the following fundamental result.

Theorem A.4. *There exists matrices A_c, B_c, C_c, D_c such that (A.13) has all the eigenvalues in $\mathbb{C}^-$ if and only if the pair (A,B) is stabilizable* and *pair (A,C) is detectable.*

Proof. [Necessity] Suppose all eigenvalues of (A.13) have negative real part. Then, by the converse Lyapunov's Theorem, there exists a unique, symmetric, and positive definite solution $P_{c\ell}$ of the matrix equation

$$P_{c\ell}A_{c\ell} + A_{c\ell}^{\mathrm{T}}P_{c\ell} = -I\,. \tag{A.14}$$

Let $P_{c\ell}$ be partitioned as in

$$P_{c\ell} = \begin{pmatrix} P & S \\ S^{\mathrm{T}} & P_c \end{pmatrix}$$

consistently with the partition of $A_{c\ell}$ (note, in this respect, that the two diagonal blocks may have different dimensions n and n_c). Note also that P and P_c are necessarily positive definite (and hence also nonsingular) because so is $P_{c\ell}$. Consider the matrix

$$T = \begin{pmatrix} I & 0 \\ -P_c^{-1}S^{\mathrm{T}} & I \end{pmatrix}.$$

Define $\tilde{P} := T^{\mathrm{T}}P_{c\ell}T$ and note that

$$\tilde{P} = \begin{pmatrix} P - SP_c^{-1}S^{\mathrm{T}} & 0 \\ 0 & P_c \end{pmatrix}. \tag{A.15}$$

Define $\tilde{A} := T^{-1}A_{c\ell}T$ and, by means of a simple computation, observe that

$$\tilde{A} = \begin{pmatrix} A + B(D_cC - C_cP_c^{-1}S^{\mathrm{T}}) & * \\ * & * \end{pmatrix}$$

in which we have denoted by an asterisk blocks whose expression is not relevant in the sequel.

From (A.14) it is seen that

$$\begin{aligned}\tilde{P}\tilde{A}+\tilde{A}^{\mathrm{T}}\tilde{P} &= (T^{\mathrm{T}}P_{\mathrm{c}\ell}T)(T^{-1}A_{\mathrm{c}\ell}T)+(T^{\mathrm{T}}A_{\mathrm{c}\ell}^{\mathrm{T}}(T^{-1})^{\mathrm{T}})(T^{\mathrm{T}}P_{\mathrm{c}\ell}T)\\ &= T^{\mathrm{T}}(P_{\mathrm{c}\ell}A_{\mathrm{c}\ell}+A_{\mathrm{c}\ell}^{\mathrm{T}}P_{\mathrm{c}\ell})T=-T^{\mathrm{T}}T.\end{aligned} \tag{A.16}$$

This shows that the matrix $\tilde{P}\tilde{A}+\tilde{A}^{\mathrm{T}}\tilde{P}$ is negative definite (because T is nonsingular) and hence so is its upper-left block. The latter, if we set

$$P_0=P-SP_{\mathrm{c}}^{-1}S^{\mathrm{T}}, \qquad F=D_{\mathrm{c}}C-C_{\mathrm{c}}P_{\mathrm{c}}^{-1}S^{\mathrm{T}}$$

can be written in the form

$$P_0(A+BF)+(A+BF)^{\mathrm{T}}P_0\,. \tag{A.17}$$

The matrix P_0 is positive definite, because it is the upper-left block of the positive definite matrix (A.15). The matrix (A.17) is negative definite, because it is the upper-left block of the negative definite matrix (A.16). Thus, by the direct criterion of Lyapunov, it follows that the eigenvalues of $A+BF$ have negative real part. This completes the proof that, if $A_{\mathrm{c}\ell}$ has all eigenvalues in $\mathbb{C}^-$, there exists a matrix F such that $A+BF$ has all eigenvalues in $\mathbb{C}^-$, i.e. the pair (A,B) is stabilizable. In a similar way it is proven that the pair (A,C) is detectable.

[Sufficiency] Assuming that (A,B) is stabilizable and that (A,C) is detectable, pick F and G so that $(A+BF)$ has all eigenvalues in $\mathbb{C}^-$ and $(A-GC)$ has all eigenvalues in $\mathbb{C}^-$. Consider the controller

$$\begin{aligned}\dot{x}_{\mathrm{c}} &= (A+BF-GC)x_{\mathrm{c}}+Gy\\ u &= Fx_{\mathrm{c}},\end{aligned} \tag{A.18}$$

i.e. set

$$A_{\mathrm{c}}=A+BF-GC, \quad B_{\mathrm{c}}=G, \quad C_{\mathrm{c}}=F, \quad D_{\mathrm{c}}=0.$$

This yields a closed-loop system

$$\begin{aligned}\dot{x} &= Ax+BFx_{\mathrm{c}}\\ \dot{x}_{\mathrm{c}} &= GCx+(A+BF-GC)x_{\mathrm{c}}\,.\end{aligned}$$

The change of variables $z=x-x_{\mathrm{c}}$ changes the latter into the system

$$\begin{pmatrix}\dot{x}\\ \dot{z}\end{pmatrix}=\begin{pmatrix}A+BF & -BF\\ 0 & A-GC\end{pmatrix}\begin{pmatrix}x\\ z\end{pmatrix}$$

This system is in block-triangular form and both diagonal blocks have all eigenvalues in $\mathbb{C}^-$. Thus the controller (A.18) guarantees that the matrix $A_{\mathrm{c}\ell}$ has all eigenvalues in $\mathbb{C}^-$. [8]

[8] Observe that the choices of F and G are *independent* of each other, i.e. F is only required to place the eigenvalues of $(A+BF)$ in $\mathbb{C}^-$ and G is only required to place the eigenvalues of $(A-GC)$

To check whether the two fundamental properties in question hold, the following tests are useful.

Lemma A.3. *The pair (A,B) is stabilizable if and only if*

$$\operatorname{rank}\begin{pmatrix} A-\lambda I & B \end{pmatrix} = n \tag{A.19}$$

for all $\lambda \in \sigma(A)$ having non-negative real part.

Lemma A.4. *The pair (A,C) is detectable if and only if*

$$\operatorname{rank}\begin{pmatrix} A-\lambda I \\ C \end{pmatrix} = n \tag{A.20}$$

for all $\lambda \in \sigma(A)$ having non-negative real part.

Remark A.2. For the sake of completeness, we recall how the two properties of stabilizability and detectability invoked above compare with the properties of reachability and observability. To this end, we recall that the linear system (A.25) is *reachable* if and only if

$$\operatorname{rank}\begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix} = n \tag{A.21}$$

or, what is the same, if and only if the condition (A.19) holds *for all* $\lambda \in \sigma(A)$ (and not just for all such λ's having non-negative real part). The linear system (A.25) is *observable* if and only if

$$\operatorname{rank}\begin{pmatrix} C \\ CA \\ \cdots \\ CA^{n-1} \end{pmatrix} = n \tag{A.22}$$

or, what is the same, if and only if the condition (A.20) holds *for all* $\lambda \in \sigma(A)$ (and not just for all such λ's having non-negative real part).

It is seen from this that, in general, reachability is a property stronger than stabilizability and observability is a property stronger than detectability. The two (pairs of) properties coincide when all eigenvalues of A have non-negative real part. If the rank of the matrix on the left-hand side of (A.21) is $n_1 < n$, the system is *not* reachable and there exists a nonsingular matrix T such that [9]

$$TAT^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \qquad TB = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}$$

in which $A_{11} \in \mathbb{R}^{n_1 \times n_1}$ and the pair (A_{11}, B_1) is reachable. This being the case, it is easy to check that the pair (A,B) is stabilizable if and only if all eigenvalues of A_{22} have negative real part. A similar criterion determines the relation between

in $\mathbb{C}^-$. For this reason, the controller (A.18) is said to be a controller inspired by a *separation principle*.

[9] This is the well-know *Kalman's* decomposition of a system into reachable/unreachable parts.

detectability and observability. If the rank of the matrix on the left-hand side of (A.22) is $n_1 < n$, the system is *not* observable and there exists a nonsingular matrix T such that

$$TAT^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \qquad CT^{-1} = \begin{pmatrix} 0 & C_2 \end{pmatrix}$$

in which $A_{22} \in \mathbb{R}^{n_1 \times n_1}$ and the pair (A_{22}, C_2) is observable. This being the case, it is easy to check that the pair (A, C) is detectable if and only if all eigenvalues of A_{11} have negative real part.

A.5 Steady-state response to harmonic inputs

Invariant subspaces. Let A be a fixed $n \times n$ matrix. A subspace $\mathscr{V}$ of $\mathbb{R}^n$ is *invariant* under A if

$$v \in \mathscr{V} \quad \Rightarrow \quad Av \in \mathscr{V}.$$

Let $d < n$ denote the dimension of $\mathscr{V}$ and let $\{v_1, v_2, \ldots, v_d\}$ be a basis of $\mathscr{V}$, that is a set of d linearly independent vectors $v_i \in \mathbb{R}^n$ such that

$$\mathscr{V} = \mathrm{Im}(V)$$

where V is the $n \times d$ matrix

$$V = \begin{pmatrix} v_1 & v_2 & \cdots & v_d \end{pmatrix}.$$

Then, it is an easy matter to check that $\mathscr{V}$ is invariant under A if and only if there exists a $d \times d$ matrix $A_{\mathscr{V}}$ such that

$$AV = VA_{\mathscr{V}}.$$

The map $z \mapsto A_{\mathscr{V}} z$ characterizes the *restriction* to $\mathscr{V}$ of the map $x \mapsto Ax$. This being the case, observe that if (λ_0, z_0) is a pair eigenvalue-eigenvector for $A_{\mathscr{V}}$ (i.e. a pair satisfying $A_{\mathscr{V}} z_0 = \lambda_0 z_0$), then $(\lambda_0, V z_0)$ is a pair eigenvalue-eigenvector for A.

Let the matrix A have n_s eigenvalues in $\mathbb{C}^-$, n_a eigenvalues in $\mathbb{C}^+$ and n_c eigenvalues in $\mathbb{C}^0 = \{\lambda \in \mathbb{C} : \mathrm{Re}[\lambda] = 0\}$, with (obviously) $n_\mathrm{s} + n_\mathrm{a} + n_\mathrm{c} = n$. Then (passing for instance through the Jordan form of A) it is easy to check that there exist three invariant subspaces of A, denoted $\mathscr{V}_\mathrm{s}$, $\mathscr{V}_\mathrm{a}$, $\mathscr{V}_\mathrm{c}$, of dimension $n_\mathrm{s}, n_\mathrm{a}, n_\mathrm{c}$ that are complementary in $\mathbb{R}^n$, i.e. satisfy

$$\mathscr{V}_\mathrm{s} \oplus \mathscr{V}_\mathrm{a} \oplus \mathscr{V}_\mathrm{c} = \mathbb{R}^n, \tag{A.23}$$

with the property that the restriction of A to $\mathscr{V}_\mathrm{s}$ is characterized by a $n_\mathrm{s} \times n_\mathrm{s}$ matrix A_s whose eigenvalues are precisely the n_s eigenvalues of A that are in $\mathbb{C}^-$, the restriction of A to $\mathscr{V}_\mathrm{a}$ is characterized by a $n_\mathrm{a} \times n_\mathrm{a}$ matrix A_a whose eigenvalues are precisely the n_a eigenvalues of A that are in $\mathbb{C}^+$ and the restriction of A to $\mathscr{V}_\mathrm{c}$ is characterized by a $n_c \times n_c$ matrix A_c whose eigenvalues are precisely the n_c eigenvalues of A that are in $\mathbb{C}^0$. These three subspaces are called the *stable eigenspace*, the *antistable eigenspace* and the *center eigenspace*.

Consider now the autonomous linear system

$$\dot{x} = Ax \tag{A.24}$$

with $x \in \mathbb{R}^n$. It is easy to check that if subspace $\mathscr{V}$ is invariant under A, then for any $x^\circ \in \mathscr{V}$, the integral curve $x(t)$ of (A.24) passing through x° at time $t = 0$ is such that $x(t) \in \mathscr{V}$ for all $t \in \mathbb{R}$. [10] Because of (A.23), any trajectory $x(t)$ of (A.24) can be uniquely decomposed as

$$x(t) = x_s(t) + x_a(t) + x_c(t)$$

with $x_s(t) \in \mathscr{V}_s$, $x_a(t) \in \mathscr{V}_a$, $x_c(t) \in \mathscr{V}_c$. Moreover, since the restriction of A to $\mathscr{V}_s$ is characterized by a matrix A_s whose eigenvalues are all in $\mathbb{C}^-$,

$$\lim_{t \to \infty} x_s(t) = 0$$

and, since the restriction of A to $\mathscr{V}_a$ is characterized by a matrix A_a whose eigenvalues are all in $\mathbb{C}^+$,

$$\lim_{t \to -\infty} x_a(t) = 0\,.$$

A geometric characterization of the steady-state response. It is well known that a stable linear system

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

subject to a (harmonic) input of the form

$$u(t) = u_0 \cos(\omega_0 t) \tag{A.25}$$

exhibits a well-defined *steady state response*, which is itself a harmonic function of time. The response in question can be easily characterized by means of a simple geometric construction. Observe that the input defined above can be viewed as generated by an autonomous system of the form

$$\begin{aligned} \dot{w} &= Sw \\ u &= Qw \end{aligned} \tag{A.26}$$

in which $w \in \mathbb{R}^2$ and

[10] Note that also the converse of such implication holds. If $\mathscr{V}$ is a subspace with the property that, for any $x^\circ \in \mathscr{V}$, the integral curve $x(t)$ of (A.24) passing through x° at time $t = 0$ satisfies $x(t) \in \mathscr{V}$ for all $t \in \mathbb{R}$, then $\mathscr{V}$ is invariant under A. The property in question is sometimes referred to as the *integral version* of the notion of invariance, while the property indicated in the text above is referred to as the *infinitesimal version* on the notion of invariance.

$$S = \begin{pmatrix} 0 & \omega_0 \\ -\omega_0 & 0 \end{pmatrix}, \qquad Q = u_0 \begin{pmatrix} 1 & 0 \end{pmatrix},$$

set in the initial state [11]

$$w(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{A.27}$$

In this way, the *forced* response of the given linear system, from any initial state $x(0)$, to the input (A.25) can be identified with the *free* response of the composite system

$$\begin{pmatrix} \dot{w} \\ \dot{x} \end{pmatrix} = \begin{pmatrix} S & 0 \\ BQ & A \end{pmatrix} \begin{pmatrix} w \\ x \end{pmatrix} \tag{A.28}$$

from the initial state $(x(0), w(0))$ with $w(0)$ given by (A.27).

Since A has all eigenvalues with negative real part and S has eigenvalues on the imaginary axis, the Sylvester equation

$$\Pi S = A\Pi + BQ \tag{A.29}$$

has a unique solution Π. The composite system (A.28) possesses two complementary invariant eigenspaces: a *stable eigenspace* and a *center eigenspace*, which can be respectively expressed as

$$\mathscr{V}^{\mathrm{s}} = \operatorname{span}\begin{pmatrix} 0 \\ I \end{pmatrix}, \qquad \mathscr{V}^{\mathrm{c}} = \operatorname{span}\begin{pmatrix} I \\ \Pi \end{pmatrix}.$$

The latter, in particular, shows that the center eigenspace is the set of all pairs (w, x) such that $x = \Pi w$.

Consider now the change of variables $\tilde{x} = x - \Pi w$ which, after a simple calculation which uses (A.29), yields

$$\begin{aligned} \dot{w} &= Sw \\ \dot{\tilde{x}} &= A\tilde{x}. \end{aligned}$$

Since the matrix A is Hurwitz, for any initial condition

$$\lim_{t\to\infty} \tilde{x}(t) = 0,$$

which shows that the (unique) projection of the trajectory along the stable eigenspace asymptotically tends to zero. In the original coordinates, this reads as

$$\lim_{t\to\infty} [x(t) - \Pi w(t)] = 0,$$

---

[11] To check that this is the case, simply bear in mind that the solution $w(t)$ of (A.26) is given by

$$w(t) = e^{St} w(0) = \begin{pmatrix} \cos(\omega_0 t) & \sin(\omega_0 t) \\ -\sin(\omega_0 t) & \cos(\omega_0 t) \end{pmatrix} \begin{pmatrix} w_1(0) \\ w_2(0) \end{pmatrix}$$

.

from which we see that the *steady-state response* of the system to any input generated by (A.26) can be expressed as

$$x_{\mathrm{ss}}(t) = \Pi w(t)\,. \tag{A.30}$$

It is worth observing that the steady-state response $x_{\mathrm{ss}}(t)$ thus defined can also be identified with an *actual* forced response of the system to the input (A.25), provided that the initial state $x(0)$ is appropriately chosen. In fact, since the center eigenspace is invariant for the composite system (A.28), if the initial condition of the latter is taken on $\mathscr{V}^{\mathrm{c}}$, i.e. if $x(0) = \Pi w(0)$, the motion of such system remains confined to $\mathscr{V}^{\mathrm{c}}$ for all t, i.e. $x(t) = \Pi w(t)$ for all t. Thus, if $x(0) = \Pi w(0)$, the actual forced response $x(t)$ of the system to the input (A.25) coincides with the steady-state response $x_{\mathrm{ss}}(t)$. Note that, in view of the definition of $w(0)$ given by (A.27), this initial state $x(0)$ is nothing else than the first column of the matrix Π.

The calculation of the solution Π of the Sylvester equation (A.29) is straightforward. Set

$$\Pi = \begin{pmatrix} \Pi_1 & \Pi_2 \end{pmatrix}$$

and observe that the equation in question reduces to

$$\Pi \begin{pmatrix} 0 & \omega_0 \\ -\omega_0 & 0 \end{pmatrix} = A\Pi + Bu_0 \begin{pmatrix} 1 & 0 \end{pmatrix}\,.$$

An elementary calculation (multiply first both sides on the right by the vector $\begin{pmatrix} 1 & j \end{pmatrix}^{\mathrm{T}}$) yields

$$\Pi_1 + j\Pi_2 = (j\omega_0 I - A)^{-1} B u_0\,,$$

i.e.

$$\Pi = \begin{pmatrix} \mathrm{Re}[(j\omega_0 I - A)^{-1}B]u_0 & \mathrm{Im}[(j\omega_0 I - A)^{-1}B]u_0 \end{pmatrix}\,.$$

As shown above, the steady state response has the form (A.30). Hence, in particular, the periodic input

$$u(t) = u_0 \cos(\omega_0 t)$$

produces the periodic state response

$$x_{\mathrm{ss}}(t) = \Pi w(t) = \Pi_1 \cos(\omega_0 t) - \Pi_2 \sin(\omega_0 t)\,, \tag{A.31}$$

and the periodic output response

$$\begin{aligned} y_{\mathrm{ss}}(t) &= Cx_{\mathrm{ss}}(t) + Du_0 \cos(\omega_0 t) \\ &= \mathrm{Re}[T(j\omega_0)]u_0 \cos(\omega_0 t) - \mathrm{Im}[T(j\omega_0)]u_0 \sin(\omega_0 t)\,, \end{aligned} \tag{A.32}$$

in which

$$T(j\omega) = C(j\omega I - A)^{-1}B + D\,.$$

A.6 Hamiltonian matrices and algebraic Riccati equations

In this section, a few fundamental facts about algebraic Riccati equations are reviewed.[12] An algebraic Riccati equation is an equation of the form

$$A^{\mathrm{T}}X + XA + Q + XRX = 0 \tag{A.33}$$

in which all matrices involved are $n \times n$ matrices and R, Q are *symmetric* matrices. Such equation can also be rewritten the equivalent form as

$$\begin{pmatrix} X & -I \end{pmatrix} \begin{pmatrix} A & R \\ -Q & -A^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} I \\ X \end{pmatrix} = 0\,.$$

From either one of these expressions, it is easy to deduce the following identity

$$\begin{pmatrix} A & R \\ -Q & -A^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} I \\ X \end{pmatrix} = \begin{pmatrix} I \\ X \end{pmatrix} (A + RX) \tag{A.34}$$

and to conclude that X is a solution of the Riccati equation (A.33) if and only if the subspace

$$\mathscr{V} = \mathrm{Im} \begin{pmatrix} I \\ X \end{pmatrix} \tag{A.35}$$

is an (n-dimensional) invariant subspace of the matrix

$$H = \begin{pmatrix} A & R \\ -Q & -A^{\mathrm{T}} \end{pmatrix}. \tag{A.36}$$

In particular, (A.34) also shows that if X is a solution of (A.33), the matrix $A + RX$ characterizes the restriction of H to its invariant subspace (A.35). A matrix of the form (A.36), with real entries and in which R and Q are symmetric matrices, is called a Hamiltonian matrix. Some relevant features of the Hamiltonian matrix (A.36) and their relationships with the Riccati equation (A.33) are reviewed in what follows.

Lemma A.5. *The spectrum of the Hamiltonian matrix (A.36) is symmetric with respect to the imaginary axis.*

Proof. Set

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

and note that

$$J^{-1}HJ = \begin{pmatrix} -A^{\mathrm{T}} & Q \\ -R & A \end{pmatrix} = -H^{\mathrm{T}}\,.$$

Hence H and -H^{T} are similar. As a consequence, if λ is an eigenvalue of H so is also $-\lambda$. Since the entries of H are real numbers, and therefore the spectrum of this matrix is symmetric with respect to the real axis, the result follows. ◁

[12] For further reading, see e.g. [2].

Suppose now that the matrix (A.36) has no eigenvalues on the imaginary axis. Then, the matrix in question has exactly n eigenvalues in $\mathbb{C}^-$ and n eigenvalues in $\mathbb{C}^+$. As a consequence, there exist two complementary n-dimensional invariant subspaces of H: a subspace $\mathscr{V}^{\mathrm{s}}$ characterized by property that restriction of H to $\mathscr{V}^{\mathrm{s}}$ has all eigenvalues in $\mathbb{C}^-$, the *stable eigenspace*, and a subspace $\mathscr{V}^{\mathrm{a}}$ characterized by property that restriction of H to $\mathscr{V}^{\mathrm{a}}$ has all eigenvalues in $\mathbb{C}^+$, the *antistable eigenspace*. A situation of special interest in the subsequent analysis is the one in which the stable eigenspace (respectively, the antistable eigenspace) of the matrix (A.36) can be expressed in the form (A.35); in this case in fact, as observed before, it is possible to associate with this subspace a particular solution of the Riccati equation (A.33).

If there exists a matrix X^- such that

$$\mathscr{V}^{\mathrm{s}} = \mathrm{Im}\begin{pmatrix} I \\ X^- \end{pmatrix},$$

this matrix satisfies

$$A^{\mathrm{T}}X^- + X^-A + Q + X^-RX^- = 0$$

and the matrix $A+RX^-$ has all eigenvalues in $\mathbb{C}^-$. This matrix is the *unique* [13] solution of the Riccati equation (A.33) having the property that $A+RX$ has all eigenvalues in $\mathbb{C}^-$ and for this reason is called *the stabilizing* solution of the Riccati equation (A.33).

Similarly, if there exists a matrix X^+ such that

$$\mathscr{V}^{\mathrm{a}} = \mathrm{Im}\begin{pmatrix} I \\ X^+ \end{pmatrix},$$

this matrix satisfies

$$A^{\mathrm{T}}X^+ + X^-A + Q + X^+RX^+ = 0$$

and the matrix $A+RX^+$ has all eigenvalues in $\mathbb{C}^+$. This matrix is the unique solution of the Riccati equation (A.33) having the property that $A+RX$ has all eigenvalues in $\mathbb{C}^+$ and for this reason is called *the antistabilizing* solution of the Riccati equation (A.33).

The existence of such matrices X^- and X^+ is discussed in the following statement.

Proposition A.1. *Suppose the Hamiltonian matrix (A.36) has no eigenvalues on the imaginary axis and R is a (either positive or negative) semidefinite matrix.*

If the pair (A,R) is stabilizable, the stable eigenspace $\mathscr{V}^{\mathrm{s}}$ of (A.36) can be expressed in the form

$$\mathscr{V}^{\mathrm{s}} = \mathrm{Im}\begin{pmatrix} I \\ X^- \end{pmatrix}$$

[13] If $\mathscr{V}$ is an n-dimensional subspace of $\mathbb{R}^{2n}$, and $\mathscr{V}$ can be expressed in the form (A.35), the matrix X is necessarily unique.

in which X^- is a symmetric matrix, the (unique) stabilizing solution of the Riccati equation (A.33).

If the pair (A,R) is antistabilizable, the antistable eigenspace $\mathscr{V}^{\mathrm{a}}$ of (A.36) can be expressed in the form

$$\mathscr{V}^{\mathrm{a}} = \mathrm{Im}\begin{pmatrix} I \\ X^+ \end{pmatrix}$$

in which X^- is a symmetric matrix, the (unique) antistabilizing solution of the Riccati equation (A.33).

The following Proposition describes the relation between solutions of the algebraic Riccati equation (A.33) and of the algebraic Riccati *inequality*

$$A^{\mathrm{T}}X + XA + Q + XRX > 0\,. \tag{A.37}$$

Proposition A.2. *Suppose R is negative semidefinite. Let X^- (respectively X^+) be a solution of the Riccati equation (A.33) having the property that $\sigma(A+RX^-) \in \mathbb{C}^-$ (respectively, $\sigma(A+RX^+) \in \mathbb{C}^+$). Then, the set of solutions of*

$$A^{\mathrm{T}}X + XA + Q + XRX > 0\,,$$

is not empty and any X in this set satisfies $X < X^-$ (respectively, $X > X^+$).

References

1. P.S. Antsaklis, A.N. Michel, *Linear Systems*, McGraw Hill, 1997.
2. P. Lancaster, L. Rodman, *Algebraic Riccati Equations*, Oxford University Press, 1995.
3. M. Hautus. Linear matrix equations with applications to the regulator problem. In *I. Landau (ed.), Outils et modèles mathematiques pour lautomatique, lanalyse de systèmes et le traitement du signal*, C.N.R.S. Paris, **3**, pp. 399-412, 1983.