

Chapter 2

Mathematical Methods

A study of control systems dynamics and optimization has always needed a suitable mathematical language to formulate the problems, analyse possible behaviours and difficulties, develop algorithms and prove that these algorithms have the desired properties. This is also true of Iterative Control which exhibits all the properties and challenges of classical control problems with the added need to consider the effect of iteration on behaviours. The material in this chapter acts to remind readers of the mathematics needed for analysis of state space models in control theory and the essential structure of quadratic optimization problems. To this mix is added an introduction to the essentials of Hilbert spaces as a representation of signals, as a means of representing dynamical systems as operators on such spaces and as a means of creating a geometric approach to iterative optimization based control. The presentation aims to define both notation and explain concepts. Fuller details and proofs of the statements can be found in the references.

2.1 Elements of Matrix Theory

A $p \times q$ real (or complex) matrix A is an array of real (or complex) numbers of the form

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1q} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2q} \\ \vdots & & & & \vdots \\ A_{p1} & A_{p2} & A_{p3} & \cdots & A_{pq} \end{bmatrix} \quad (2.1)$$

The element in the i th row and j th column is denoted A_{ij} . If $q = 1$, the matrix is often called a vector. Block matrices can also be defined where the A_{ij} are $p_i \times q_j$ sub-matrices. In this case the dimensions of the matrix A are $\sum_{i=1}^p p_i \times \sum_{j=1}^q q_j$.

The following is essential for control theoretical purposes

1. The set of $p \times 1$ real (respectively complex) vectors is given the symbol \mathcal{R}^p (respectively \mathcal{C}^p).
2. A is said to be *square* if the number of rows is equal to the number of columns.
3. Addition of two $p \times q$ matrices A, B to form a $p \times q$ matrix C written $C = A + B$ is defined by the elements

$$C_{ij} = A_{ij} + B_{ij}, \quad \text{for } 1 \leq i \leq p, \quad 1 \leq j \leq q. \quad (2.2)$$

The $p \times q$ *zero matrix* is the matrix with all elements equal to zero.

4. Multiplication of A by a scalar λ produces a matrix $C = \lambda A$ where

$$C_{ij} = \lambda A_{ij}, \quad \text{for } 1 \leq i \leq p, \quad 1 \leq j \leq q. \quad (2.3)$$

5. Multiplication of a $p \times q$ matrix A by a $q \times r$ matrix B to produce a $p \times r$ matrix $C = AB$ is defined by the following computation of elements of C

$$C_{ij} = \sum_{k=1}^q A_{ik} B_{kj}, \quad \text{for } 1 \leq i \leq p, \quad 1 \leq j \leq r. \quad (2.4)$$

6. The *transpose* of a $p \times q$ matrix A is the $q \times p$ matrix A^T with

$$(A^T)_{ji} = A_{ij}, \quad \text{for } 1 \leq i \leq p, \quad 1 \leq j \leq q. \quad (2.5)$$

The act of taking the transpose of a product satisfies the rule $(AB)^T = B^T A^T$. If $A = A^T$ then A is said to be *symmetric*.

7. The *conjugate transpose* of a complex $p \times q$ matrix A is the $q \times p$ matrix A^* with

$$(A^*)_{ji} = \bar{A}_{ij}, \quad \text{for } 1 \leq i \leq p, \quad 1 \leq j \leq q, \quad (2.6)$$

where \bar{a} denotes the complex conjugate of a . The act of taking the conjugate transpose of a product satisfies the rule $(AB)^* = B^* A^*$. If $A = A^*$ then A is said to be *Hermitian*. If A is real then the conjugate transpose is simply the transpose and, if A is Hermitian, it is symmetric.

8. The determinant of a square $p \times p$ matrix is denoted $\det[A]$, $|A|$ or,

$$\det[A] = \begin{vmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1p} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2p} \\ \vdots & & & & \vdots \\ A_{p1} & A_{p2} & A_{p3} & \cdots & A_{pp} \end{vmatrix} \quad (2.7)$$

The determinant has many properties including

- a. the properties $\det[A] = \det[A^T]$ and $\det[\overline{A}] = \overline{\det[A]}$.
 - b. If both A and B are square $p \times p$ matrices, then $\det[AB] = \det[A] \det[B] = \det[BA]$.
9. A is said to be an *injection* or *one-to-one* if the only $p \times 1$ vector v satisfying the equation $Av = 0$ is the vector $v = 0$. In general, the set of vectors v such that $Av = 0$ is a vector subspace of \mathcal{R}^q (or \mathcal{C}^q) and is called the *kernel* or *null space* of A and denoted

$$\ker[A] = \{v : Av = 0\} \tag{2.8}$$

The subspace $\ker[A]$ is always $\neq \{0\}$ when $q > p$.

10. If, for every $p \times 1$ vector w , there exists a vector v such that $w = Av$, then A is said to be *onto* or a *surjection*. More generally, the set of all vectors w for which there exists a vector v such that $w = Av$ is called the *range* of A . It is a vector subspace of \mathcal{R}^p (or \mathcal{C}^p) and is denoted by

$$\mathcal{R}[A] = \{w : w = Av \text{ for some vector } v\} \tag{2.9}$$

A necessary condition for the range to be equal to \mathcal{R}^p (or \mathcal{C}^p as appropriate) is that $q \geq p$.

11. If A is both a surjection and an injection, it is said to be a *bijection* (or simply *nonsingular*). If A is a $p \times p$ square matrix, then it is a bijection if, and only if, it has non-zero determinant. If $\det[A] = 0$ then A is said to be *singular*.
12. A $p \times q$ matrix A is *invertible* if, and only if, it is a bijection. In particular, this requires that it is square ($p = q$) and it is equivalent to the statement that, for every vector w , there exists a unique vector v such that $w = Av$. The *inverse* of A is denoted by A^{-1} . It is a square $p \times p$ matrix of the form

$$A^{-1} = \frac{\text{adj}[A]}{\det[A]} \tag{2.10}$$

where the *adjugate* matrix $\text{adj}[A]$ has elements that are well defined polynomials in the elements of A .

13. For all invertible $p \times p$ matrices A , the inverse A^{-1} satisfies the equations

$$AA^{-1} = A^{-1}A = I_p \tag{2.11}$$

where I_p denotes the $p \times p$ *unit matrix* or *identity*

$$I_p = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.12}$$

For any $p \times q$ matrix B and $q \times p$ matrix C , the properties $I_p B = B$ and $C I_p = C$ hold true. Also $\det[I_p] = 1$ and hence $\det[A^{-1}] = (\det[A])^{-1}$.

Note: for notational simplicity, the subscript p on I_p is sometimes dropped to leave the symbol I . This should cause no confusion as matrix dimensions are usually clear from the context.

14. If A and B are $p \times p$ nonsingular matrices, then $(AB)^{-1} = B^{-1}A^{-1}$.
15. If T is square and nonsingular, then the mapping $A \mapsto T^{-1}AT$ is a *similarity transformation*. Both A and $T^{-1}AT$ have the same eigenvalues and $\det[A] = \det[T^{-1}AT]$.
16. For non-square $p \times q$ matrices A , other definitions of inverse play a role in matrix analysis. In particular, if $p \geq q$, a *left inverse* B of A is any matrix satisfying the condition $BA = I_q$. In a similar manner, if $p \leq q$, a *right inverse* B of A is any matrix satisfying the condition $AB = I_p$. A left inverse of A exists if, and only if, A has kernel $\{0\}$ and a right inverse exists, if and only if, $\mathcal{R}[A] = \mathcal{R}^p$ (or \mathcal{C}^p). If $p \neq q$, any left or right inverse is non-unique. If $p = q$, then they are unique and equal to the inverse A^{-1} . Specific examples of left, respectively right, inverses are given by, respectively,

$$B = (A^*A)^{-1}A^*, \quad B = A^*(AA^*)^{-1}. \quad (2.13)$$

17. Given the definition of the unit matrix, two useful relationships are as follows

- a. If A and B are, respectively, $p \times q$ and $q \times p$, then

$$\det[I_p + AB] = \det[I_q + BA]. \quad (2.14)$$

- b. If A has the partitioned form

$$A = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \quad (2.15)$$

with M_{11} square and nonsingular, then *Schur's Formula* is valid,

$$\det[A] = \det[M_{11}] \det[M_{22} - M_{21}M_{11}^{-1}M_{12}]. \quad (2.16)$$

The above algebraic properties of matrices are the basis of manipulation. For analysis purposes, a number of other properties and concepts are required and are summarized as follows

1. A finite set $\{H_j\}_{1 \leq j \leq M}$ of real (respectively complex) $p \times q$ matrices is said to be *linearly independent* if, and only if, the only real (respectively complex) scalars $\{a_j\}_{1 \leq j \leq M}$ satisfying the condition $\sum_{j=1}^M a_j H_j = 0$ are $a_j = 0$, $1 \leq j \leq M$.
2. The *rank* of a $p \times q$ matrix A is the maximum number of linearly independent columns of A regarded as $p \times 1$ vectors. A $p \times p$ matrix A is nonsingular if, and only if, it has rank equal to its dimension p .

3. The *characteristic polynomial* of a square $p \times p$ matrix A is defined by the determinant

$$\rho(s) = |sI_p - A| = \sum_{j=0}^p a_{p-j}s^j \quad \text{with } a_0 = 1. \quad (2.17)$$

It is a polynomial of degree p in the complex variable s with p , possibly complex, roots λ_j , $1 \leq j \leq p$ called the *eigenvalues* of A . If A is a real matrix, then the eigenvalues are either real or exist in complex conjugate pairs. More precisely, if λ is an eigenvalue, then its complex conjugate $\bar{\lambda}$ is also an eigenvalue. The *spectral radius* of A is defined by

$$r(A) = \max_{1 \leq j \leq p} |\lambda_j|. \quad (2.18)$$

4. A complex number λ is an eigenvalue of A if, and only if, there exists a non-zero solution vector $v \in \mathcal{C}^p$ solving the equation

$$Av = \lambda v \quad (2.19)$$

Such an *eigenvector* is not uniquely defined as, for example, it can be multiplied by any scalar and still be an eigenvector. If A has p linearly independent eigenvectors $\{v_j\}_{1 \leq j \leq p}$ then an *eigenvector matrix* E of A is defined to be the block matrix $E = [v_1, v_2, \dots, v_p]$. It is nonsingular and can be used to *diagonalize* A using the *similarity transformation*

$$E^{-1}AE = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_p \end{bmatrix} = \text{diag} [\lambda_1, \lambda_2, \dots, \lambda_p] \quad (2.20)$$

The diagonal matrix produced is often called the *diagonal canonical form* of A . A always has linearly independent eigenvectors if its p eigenvalues are distinct.

5. As $|sI - A| = |\bar{s}I - A^*|$, the eigenvalues of the conjugate transpose matrix are exactly the complex conjugates of the eigenvalues $\{\lambda_j\}$ of A . Suppose that the eigenvectors of A^* are denoted w_j , $1 \leq j \leq p$ and that $A^*w_j = \bar{\lambda}_j w_j$. This can be rewritten in the form $w_j^*A = \lambda_j w_j^*$ and w_j^* is termed a *left eigenvector* of A . If A has p linearly independent eigenvectors and associated eigenvector matrix E , then

$$E^{-1}A = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_p]E^{-1}. \quad (2.21)$$

Equating rows of the two sides of the equation indicates that the rows of E^{-1} are left eigenvectors of A and, as $E^{-1}E = I$, these left eigenvectors satisfy the conditions

$$w_i^*v_j = \delta_{ij}, \quad (2.22)$$

the *Kronecker delta* defined as $\delta_{ij} = 0$ whenever $i \neq j$ and unity otherwise.

6. If A is Hermitian, then its eigenvalues are all real valued. Its eigenvectors v_j form a set of p linearly independent vectors. It is always possible to scale these vectors so that they satisfy the *orthogonality condition*

$$v_i^* v_j = \delta_{ij} \quad (2.23)$$

Under these circumstances, $E^{-1} = E^*$ so that $E^* E = I$, an example of a *unitary matrix*. If E is real then it is an *orthogonal matrix*.

7. Almost every square matrix A can be diagonalized in the manner shown above. In some cases, diagonalization is not possible but, in such cases, a *Jordan canonical form* can be produced. More precisely, there exists an integer q and a nonsingular $p \times p$ matrix J with the property that

$$J^{-1} A J = \begin{bmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & J_q \end{bmatrix} = \text{blockdiag} [J_1, J_2, \dots, J_q] \quad (2.24)$$

where each *Jordan block* J_j , $1 \leq j \leq q$ has the structure of a $q_j \times q_j$ matrix as follows

$$J_j = \begin{bmatrix} \gamma_j & 1 & 0 & \cdots & 0 & 0 \\ 0 & \gamma_j & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \vdots & \\ 0 & 0 & 0 & \cdots & \gamma_j & 1 \\ 0 & 0 & 0 & \cdots & 0 & \gamma_j \end{bmatrix} \quad (2.25)$$

where each γ_j is an eigenvalue of A .

8. In all cases the determinant of A is computed from the product of all p eigenvalues

$$\det[A] = \prod_{j=1}^p \lambda_j. \quad (2.26)$$

A is singular if, and only if, it has a zero eigenvalue and hence $\ker[A] \neq \{0\}$.

9. In all cases, a square matrix A “satisfies its own characteristic equation”

$$\rho(A) = \sum_{j=0}^p a_{p-j} A^j = 0. \quad (2.27)$$

This statement is normally known as the *Cayley-Hamilton Theorem*. The result is the basis of many theoretical simplifications and insight exemplified by the easily proven fact that, if A is nonsingular, then $a_p \neq 0$ and

$$A^{-1} = -a_p^{-1} \sum_{j=1}^p a_{p-j} A^{j-1}. \tag{2.28}$$

That is, the inverse can be expressed in terms of powers of A and the coefficients in the characteristic polynomial.

Note: A related polynomial is the *minimum polynomial* of A which is the uniquely defined polynomial $\rho_{min}(s)$ of minimum degree that has the property $\rho_{min}(A) = 0$. The degree of the minimum polynomial is always less than or equal to p and is always equal to p if the eigenvalues of A are distinct.

10. More generally, the Cayley-Hamilton theorem implies useful facts about *functions of matrices*. If $f(s)$ is an analytic function of the complex variable s expressible as a power series $\sum_{j=0}^{\infty} f_j s^j$ with radius of convergence R , then the symbol $f(A)$ denotes the associated function of A defined by

$$f(A) = \sum_{j=0}^{\infty} f_j A^j \tag{2.29}$$

This series converges whenever the spectral radius $r(A) < R$. For example,

- a. the exponential function e^s has a power series expansion with $f_j = \frac{1}{j!}$. The corresponding matrix exponential is

$$e^A = \sum_{j=0}^{\infty} \frac{1}{j!} A^j = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \dots \tag{2.30}$$

- b. The function $(1 - s)^{-1} = \sum_{j=0}^{\infty} s^j$ has a radius of convergence $R = 1$. It follows that, if the spectral radius $r(A) < 1$, the matrix inverse $(I - A)^{-1}$ exists and has the convergent series expansion

$$(I - A)^{-1} = \sum_{j=0}^{\infty} A^j. \tag{2.31}$$

If A has a nonsingular eigenvector matrix E , then $A^j = E \text{diag}[\lambda_1^j, \dots, \lambda_p^j] E^{-1}$ and

$$f(A) = \sum_{j=0}^{\infty} f_j A^j = E \text{diag}[f(\lambda_1), \dots, f(\lambda_m)] E^{-1} \tag{2.32}$$

11. From the Cayley-Hamilton theorem it is easily seen that all powers A^j with $j \geq p$ can be expressed as a polynomial in A of degree less than or equal to $p - 1$. It follows that all functions $f(A)$ can be expressed as a polynomial in A of degree less than or equal to $p - 1$ by suitable choice of coefficients.

12. The *Spectral Mapping Theorem* states that, if A has eigenvalues λ_j , $1 \leq j \leq p$, and $r(A) < R$, then the eigenvalues of $f(A)$ are precisely $f(\lambda_j)$, $1 \leq j \leq p$.

The final group of useful properties are associated with the idea of positivity of *quadratic forms*

1. Suppose that A is a square, real, symmetric, $p \times p$ matrix and x an arbitrary $p \times 1$ vector in \mathcal{R}^p . Then the quadratic function $x^T A x$ is a *quadratic form*. If A is not symmetric, it can always be replaced by a symmetric matrix as $x^T A x \equiv x^T \left(\frac{A+A^T}{2} \right) x$.
2. If A is complex then the quadratic form is defined on \mathcal{C}^p as $x^* A x$. If A is Hermitian, then $x^* A x$ takes only real values.
3. A real matrix A is said to be *positive* if, and only if, $x^T A x \geq 0$ for all vectors $x \in \mathcal{R}^p$. If $x^T A x > 0$ whenever $x \neq 0$, then A is said to be *positive definite* and written in the form $A > 0$. If A is positive but not positive definite, it is *positive semi-definite* and written in the form $A \geq 0$. The expression $A \geq B$ (respectively $A > B$) is equivalent to $A - B \geq 0$ (respectively $A - B > 0$). Similar definitions are used for complex matrices and their associated quadratic forms.
4. Conditions for positivity for real, symmetric matrices include the following
 - a. A real, symmetric matrix A is positive if, and only if, all its eigenvalues satisfy the inequalities $\lambda_j \geq 0$, $1 \leq j \leq p$. It is positive definite if, and only if, all eigenvalues are strictly positive. Positive definite, symmetric matrices are hence always invertible.
 - b. If A and B are matrices and $A = B^T B$, then A is positive. A is positive definite if, and only if, $\ker[B] = \{0\}$.
 - c. A real, symmetric $p \times p$ matrix A is positive definite if and only if the *Principal Minors*

$$\begin{vmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1q} \\ A_{21} & A_{22} & A_{23} & \cdots & A_{2q} \\ \vdots & & & & \vdots \\ A_{q1} & A_{q2} & A_{q3} & \cdots & A_{qq} \end{vmatrix} > 0, \quad \text{for } 1 \leq q \leq p. \quad (2.33)$$

5. Positive, symmetric, real matrices A always have an orthogonal eigenvector matrix E and can be expressed in the form $A = E \text{diag}[\lambda_1, \dots, \lambda_p] E^T$ with $E^T E = I$.
6. A positive, symmetric, real matrix A has a unique symmetric, positive *square root* B such that $B^2 = A$. As with scalar square roots, B is usually denoted by the symbol $A^{\frac{1}{2}}$. It can be computed from the formula

$$A = E \text{diag}[\lambda_1^{\frac{1}{2}}, \dots, \lambda_p^{\frac{1}{2}}] E^T \quad (2.34)$$

2.2 Quadratic Optimization and Quadratic Forms

2.2.1 Completing the Square

The conceptual basis of much of optimization theory used in control systems algorithms has its origins in the simple ideas of minimization of quadratic functions of vectors in \mathcal{R}^p . This short section explains the basic ideas using a simple example and without the need for advanced mathematical methodologies more complicated than the matrix theory described above. The problem used to illustrate the ideas is the problem of minimizing the quadratic objective function

$$J(x) = x^T A x + 2b^T x + c. \quad (2.35)$$

where the $p \times p$ matrix A is real, symmetric and positive definite, b is real and $p \times 1$ and c is a real number. The solution is easily found by *completing the square* and verifying that

$$J(x) = (x + A^{-1}b)^T A(x + A^{-1}b) - b^T A^{-1}b + c. \quad (2.36)$$

The second two terms are independent of x . The fact that A is positive definite immediately yields the fact that the minimum value occurs when the first term is zero. The unique minimizing solution is hence

$$x_\infty = -A^{-1}b \quad \text{and} \quad J(x_\infty) = -b^T A^{-1}b + c. \quad (2.37)$$

Both can be computed using standard software if the matrices involved are of reasonable dimension and not ill-conditioned. Factors causing problems include:

1. Suppose that the eigenvalues of A are listed in order of ascending value $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p$ and A is written in its diagonal form $A = E \text{diag}[\lambda_1, \dots, \lambda_p] E^T$ with $E^{-1} = E^T$. The *condition number* of A is defined to be $c(A) = \frac{\lambda_p}{\lambda_1}$. It follows that the inverse of A has the structure

$$A^{-1} = E \text{diag}[\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_p^{-1}] E^T \quad (2.38)$$

The situation where the spread of the eigenvalues of A is large (that is, $c(A)$ is large) can be discussed by considering the case where λ_1 is very small. In such situations, small errors in characterizing this eigenvalue can lead to large changes in the computed solution $x = -A^{-1}b$.

2. These problems are exacerbated if the dimension p is large due to the number of floating point operations necessary in computer computation of A^{-1} .

In the quadratic problems considered in this text, similar quadratic objective functions will be considered but the “matrices” involved are replaced by operators associated with dynamical systems models and, in intuitive terms, have very high

(even infinite) dimensions and extremely large (possibly infinite) condition numbers. Formal solutions paralleling the algebraic constructions illustrated above hence have no immediate computational value, the core of the theoretical problem being that of developing control algorithms that use only feasible computational procedures that can be implemented using well-conditioned off-line algorithms and on-line feedback controllers.

2.2.2 Singular Values, Lagrangians and Matrix Norms

The *singular values* $0 \leq \sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_q$ of a real (respectively complex) $p \times q$ matrix A are real, positive numbers computed from the eigenvalues $0 \leq \lambda_1 \leq \dots \leq \lambda_q$ of the symmetric (respectively Hermitian) matrix $A^T A$ (respectively $A^* A$) by writing $\lambda_j = \sigma_j^2$, $1 \leq j \leq q$. The corresponding eigenvectors are often called *singular vectors*.

Associated with the matrix A is the notion of a matrix *norm*. As will be seen throughout this text, the idea of a norm is non-unique. What follows, therefore, is only an example that builds on the idea of singular values and illustrates the use of Lagrangian methods in optimization problems. The first step is the definition of a particular vector norm, the Euclidean norm, defined on vectors $x \in \mathcal{R}^q$ (respectively \mathcal{C}^q) by $\|x\| = \sqrt{x^T x}$ (respectively $\sqrt{x^* x}$). The Euclidean norm induces a norm $\|A\|$ on the matrix A by defining

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup\{\|Ax\| : \|x\| = 1\} \quad (2.39)$$

In particular, it follows that, for all vectors $x \in \mathcal{R}^q$ (respectively \mathcal{C}^q),

$$\|Ax\| \leq \|A\| \|x\|. \quad (2.40)$$

As a consequence, if A and B are $p \times q$ and $q \times r$ matrices, it follows that $\|ABx\| \leq \|A\| \|B\| \|x\|$ and hence

$$\|AB\| \leq \|A\| \|B\|. \quad (2.41)$$

Suppose now that A is real and vectors are in \mathcal{R}^q . From the above, the induced norm is the solution of an optimization problem with an equality constraint, namely

$$\|A\|^2 = \sup\{\|Ax\|^2 : \|x\| = 1\} = \sup\{x^T A^T A x : x^T x = 1\}. \quad (2.42)$$

The solution of this problem is computed by solving for the unique stationary point of the *Lagrangian*

$$\mathcal{L}[x, \lambda] = x^T A^T A x + 2\lambda^T (1 - x^T x) \quad (2.43)$$

where λ is the scalar *Lagrange Multiplier* for the single constraint $1 - x^T x = 0$. The stationary point of the Lagrangian is the solution of the two equations

$$1 - x^T x = 0 \quad \text{and} \quad A^T A x = \lambda x, \quad (2.44)$$

That is, λ is an eigenvalue of $A^T A$ and the largest value of the optimization objective function $\|Ax\|^2 = x^T A^T A x$ is simply the largest eigenvalue of $A^T A$. Hence

$$\|A\| = \sigma_q. \quad (2.45)$$

which provides a simple link between matrix norms and singular values. This relationship also holds for complex matrices operating on \mathcal{C}^q . Finally,

1. If A is $p \times p$ and nonsingular, then $\|A\| = \sigma_p$ and $\|A^{-1}\| = \sigma_1^{-1}$.
2. The smallest and largest singular values are denoted by $\underline{\sigma}(A) = \sigma_1$ and $\overline{\sigma}(A) = \sigma_p$ respectively.
3. The spectral radius is linked to matrix norms by the formula

$$r(A) = \lim_{k \rightarrow \infty} \|A^k\|^{1/k} \quad (2.46)$$

from which, for all $\varepsilon > 0$, there exists a real number $M_\varepsilon \geq 1$ such that

$$\|A^k\| \leq M_\varepsilon (r(A) + \varepsilon)^k \quad (2.47)$$

If A can be diagonalized by a nonsingular eigenvector matrix E , then it is possible to choose $\varepsilon = 0$ and $M_0 = \|E^{-1}\| \|E\|$.

2.3 Banach Spaces, Operators, Norms and Convergent Sequences

2.3.1 Vector Spaces

Matrices are just part of a more general approach to signal analysis based on *vector spaces* which are a mathematical generalization of the familiar three dimensional world that we live in. A real (respectively complex) vector space \mathcal{V} is a collection of objects (called vectors) with defined properties of vector addition and multiplication by real (respectively complex) scalars that satisfy the familiar relations

$$\begin{aligned} v_1 + v_2 &= v_2 + v_1 \\ v_1 + (v_2 + v_3) &= (v_1 + v_2) + v_3 \\ (\lambda_1 + \lambda_2)v &= \lambda_1 v + \lambda_2 v \\ \lambda(v_1 + v_2) &= \lambda v_1 + \lambda v_2 \end{aligned} \quad (2.48)$$

for all v, v_1, v_2, v_3 in \mathcal{V} and all scalars $\lambda, \lambda_1, \lambda_2$. The zero vector in \mathcal{V} is denoted by the symbol 0 . A *vector subspace* (or, more simply, a subspace) $\mathcal{U} \subset \mathcal{V}$ is any subset of \mathcal{V} that satisfies the properties defined above.

It is easily seen that \mathcal{R}^p (respectively \mathcal{C}^p) is a real (respectively complex) vector space. Also the set of real (respectively complex) $p \times q$ matrices is a real (respectively complex) vector space. Other examples and constructs of relevance to this text include

1. If \mathcal{V} is any real vector space, then its *complexification* \mathcal{V}^c is defined to be the complex vector space of all complex vectors $v = v_1 + iv_2$ with both v_1 and v_2 elements of \mathcal{V} . \mathcal{V}^c is sometimes written in the form

$$\mathcal{V}^c = \mathcal{V} \oplus i\mathcal{V}. \quad (2.49)$$

2. The space of infinite sequences (or time series) $\alpha = \{\alpha_0, \alpha_1, \alpha_2, \dots\}$ with $\alpha_j \in \mathcal{R}^p$ (or \mathcal{C}^p) is a vector space with addition $\gamma = \alpha + \beta$ and multiplication by scalars $\lambda\alpha$ defined by the equations, $\gamma_j = \alpha_j + \beta_j$ and $\lambda\gamma_j = \lambda\alpha_j$ for $j = 0, 1, 2, 3, \dots$. A number of subspaces are of particular relevance here including ℓ_∞ (the subspace of bounded sequences of scalars satisfying $\sup_{j \geq 0} |\alpha_j| < +\infty$) and ℓ_2 (the subspace of sequences of scalars satisfying $\sum_{j=0}^{\infty} |\alpha_j|^2 < +\infty$)
3. The space of real or complex valued continuous functions of a real variable t on an interval $a \leq t \leq b$ (denoted $[a, b] \subset \mathcal{R}$) is denoted by the symbol $C[a, b]$ and is a vector space with the usual definitions of addition and multiplication.
4. The space of all functions defined on $[a, b]$ and taking values in \mathcal{R}^p (respectively \mathcal{C}^p) is a real (respectively complex) vector space with the usual definitions of addition and multiplication. The real vector space $L_2^p[a, b]$ is the set of all real $p \times 1$ vector-valued functions f such that the Lebesgue integral

$$\|f\|^2 = \int_a^b \|f(t)\|^2 dt \quad (2.50)$$

is well defined and finite. If Q is any symmetric, positive definite $p \times p$ matrix, then an equivalent statement is that

$$\|f\|_Q^2 = \int_a^b f^T(t) Q f(t) dt \quad (2.51)$$

is well defined and finite. If $p = 1$, then the space is written $L_2[a, b]$.

The notion of *linear independence* of a set of vectors follows the example of matrix theory. More precisely, a set of vector $\{v_j\}_{1 \leq j \leq M}$ is linearly independent if, and only if, $\sum_{j=1}^M a_j v_j = 0$ implies that all scalars a_j are zero. A *basis* for \mathcal{V} is a linearly independent set $\{v_j\}_{1 \leq j \leq M}$ such that all elements $v \in \mathcal{V}$ can be written as a unique linear combination of the $\{v_j\}_{1 \leq j \leq M}$. If M is finite, then the space is said to be *finite dimensional* of *dimension* M . Otherwise it is infinite dimensional. The spaces \mathcal{R}^p and \mathcal{C}^p have dimension p whilst $C[a, b]$ and $L_2[a, b]$ are infinite dimensional. For infinite dimensional spaces, the statement is more clearly stated by saying that the basis set

$\{v_j\}_{j \geq 0}$ has the property that all finite subsets of this set are linearly independent and the set of all finite linear combinations is *dense* in \mathcal{V} . The concept of denseness is more fully described in Sect. 2.3.4.

Finally,

1. The space \mathcal{V} is said to be the sum of the vector subspaces $\{\mathcal{V}_j\}_{1 \leq j \leq q}$ of \mathcal{V} , written

$$\mathcal{V} = \mathcal{V}_1 + \mathcal{V}_2 + \cdots + \mathcal{V}_q, \quad (2.52)$$

if any vector $v \in \mathcal{V}$ can be written as a linear combination $v = \sum_{j=1}^q \alpha_j v_j$ with $v_j \in \mathcal{V}_j$, $1 \leq j \leq q$ and suitable choice of scalars $\{\alpha_j\}_{1 \leq j \leq q}$. It is a *direct sum decomposition* written

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_q \quad (2.53)$$

if, and only if, each $v \in \mathcal{V}$ can be written as a unique linear combination of elements of the subspaces.

2. A *product vector space* constructed from vector spaces $\{\mathcal{V}_j\}_{1 \leq j \leq q}$ is denoted by the *Cartesian Product* notation

$$\mathcal{V} = \mathcal{V}_1 \times \mathcal{V}_2 \times \cdots \times \mathcal{V}_q \quad (2.54)$$

and consists of the set of p -tuples $v = \{v_1, v_2, \dots, v_q\}$ with $v_j \in \mathcal{V}_j$, $1 \leq j \leq q$ and the same laws of composition as those defined for (finite) times series. An example of this notation is the real product space $L_2^p[a, b]$ defined by

$$L_2^p[a, b] = \underbrace{L_2[a, b] \times L_2[a, b] \times \cdots \times L_2[a, b]}_{p - \text{copies}} \quad (2.55)$$

It is sometimes convenient to identify $L_2^p[a, b]$ with the space of $p \times 1$ vectors f with elements consisting of real valued functions $f_j \in L_2[a, b]$, $1 \leq j \leq p$.

2.3.2 Normed Spaces

Measures of magnitude are important in applications of mathematics and are used extensively in this text as a means of algorithm design and analysis. For the vector space \mathcal{R}^p , the familiar measure is the Euclidean length of the vector defined as $\|v\| = \sqrt{v^T v}$. This is just an example of the more general concept of a *vector norm*. More precisely, if \mathcal{V} is a finite or infinite dimensional, real or complex vector space, then a norm on \mathcal{V} is a mapping from vectors v into real numbers $\|v\|$ with the properties that

$$\begin{aligned}
& \|v\| \geq 0 \\
\|v\| = 0 & \text{ if, and only if, } v = 0 \\
\|\lambda v\| & = |\lambda| \|v\| \\
\|v_1 + v_2\| & \leq \|v_1\| + \|v_2\|
\end{aligned} \tag{2.56}$$

for all vectors v, v_1, v_2 in \mathcal{V} and scalars λ . Note that the space \mathcal{V} possessing the norm is normally understood from the context but it is often useful to identify the space using a subscript such as $\|v\|_{\mathcal{V}}$ or by other means.

An example of a norm in $L_2^p[a, b]$ is

$$\|f\| = \left(\int_a^b e^{2\alpha t} f^T(t) Q f(t) dt \right)^{1/2} \tag{2.57}$$

where Q is a symmetric, real, positive-definite $p \times p$ matrix and α is any real scalar. Also, a norm in $C[a, b]$ can be defined by

$$\|f\| = \sup_{a \leq t \leq b} (e^{\alpha t} |f(t)|). \tag{2.58}$$

The $L_2^p[a, b]$ norm is also a norm on $C[a, b]$.

For real (or complex) $p \times q$ matrices, A , one choice of norm is the maximum singular value $\bar{\sigma}(A)$ of A whilst, if A is real, another is the so-called *Frobenius Norm* defined by the trace formula

$$\|A\| = \sqrt{\text{tr}[A^T A]} \quad (\text{the Frobenius Norm}). \tag{2.59}$$

or, more generally,

$$\|A\| = \sqrt{\text{tr}[W A^T Q A]} \quad (\text{the weighted Frobenius Norm}). \tag{2.60}$$

where Q and W are symmetric and positive definite matrices.

When endowed with a chosen norm $\|\cdot\|$, the space \mathcal{V} is called a *normed space*. The choice of norm is non-unique and the same underlying vector space, when given different norms, generates a new normed space. For \mathcal{R}^p , the following are norms for $v = [v_1, v_2, \dots, v_p]^T$,

$$\|v\| = \sqrt{v^T Q v} \quad (Q = Q^T > 0), \quad \|v\| = \max |v_j|, \quad \|v\| = \left(\sum_{j=1}^p |v_j|^q \right)^{1/q}, \quad q \geq 1. \tag{2.61}$$

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same underlying vector space are said to be *topologically equivalent* if, and only if, there exists scalars $0 < \beta_1 \leq \beta_2$ such that, for all $v \in \mathcal{V}$,

$$\beta_1 \|v\|_1 \leq \|v\|_2 \leq \beta_2 \|v\|_1 \tag{2.62}$$

All norms on a given finite dimensional space are topologically equivalent.

2.3.3 Convergence, Closure, Completeness and Banach Spaces

Given a normed space \mathcal{V} , an infinite sequence $\{v_j\}_{j \geq 0} = \{v_0, v_1, v_2, \dots\}$ is said to *converge in the norm topology* (or, more simply, to *converge*) to a limit vector $v \in \mathcal{V}$ if, and only if,

$$\lim_{j \rightarrow \infty} \|v - v_j\|_{\mathcal{V}} = 0 \quad (\text{written } \lim_{j \rightarrow \infty} v_j = v). \quad (2.63)$$

The nature of this convergence is defined by the norm used but it is easily seen that convergence with respect to one norm implies convergence with respect to any other topologically equivalent norm.

A subset $S \subset \mathcal{V}$ is said to be an *open* subset if, for every point $v \in \mathcal{V}$, the *Open Ball* $B(v; \delta)$ defined by

$$B(v; \delta) = \{w : w \in \mathcal{V}, \|v - w\| < \delta\} \quad (2.64)$$

lies in S for some choice of $\delta > 0$. S is said to be *closed* if it contains the limit points of all convergent sequences with elements in S . The *closure* of a subset S (denoted \bar{S}) is the set consisting of points in S plus the limits of all convergent sequences in S . One consequence of this is that the *Closed Ball* $B_c(v; \delta)$ defined by

$$B_c(v; \delta) = \{w : w \in \mathcal{V}, \|v - w\| \leq \delta\} \quad (2.65)$$

is the closure of the open ball $B(v; \delta)$. Subsets can be neither open nor closed. For real numbers, the symbols $[a, b]$, $[a, b)$, $(a, b]$ and (a, b) are used to denote the intervals, respectively,

$$\begin{aligned} \{t : a \leq t \leq b\} & \quad (\text{a closed interval}), \\ \{t : a \leq t < b\} & \quad (\text{a half open interval}), \\ \{t : a < t \leq b\} & \quad (\text{a half open interval}), \\ \{t : a < t < b\} & \quad (\text{an open interval}), \end{aligned} \quad (2.66)$$

A *Cauchy Sequence* $\{v_j\}_{j \geq 0}$ in \mathcal{V} is a sequence with the property that, for all $\varepsilon > 0$, there exists an integer n_ε such that $\|v_j - v_k\| < \varepsilon$ for all $j \geq n_\varepsilon$ and $k \geq n_\varepsilon$. That is, all points v_j in the sequence get “closer and closer together” as the index j increases. In general, not all Cauchy sequences converge. An example of this is the space $C[a, b]$ with the $L_2[a, b]$ norm. It is a simple matter to construct a sequence of continuous functions that converge in norm to a discontinuous function which, by definition, is not in $C[a, b]$. A normed space where all Cauchy sequences converge is said to be *complete* and it is said to be a *Banach Space*. For the purposes of this text, note that all normed spaces used are Banach Spaces unless otherwise stated including \mathcal{R}^p , \mathcal{C}^p , ℓ_2 and $L_2^p[a, b]$ and their Cartesian products for any p and $-\infty < a < b < +\infty$.

2.3.4 Linear Operators and Dense Subsets

All $p \times q$ real matrices are examples of linear operators between the real spaces \mathbb{R}^q and \mathbb{R}^p . The more general concepts of a linear operator/linear operators $\Gamma : \mathcal{V} \mapsto \mathcal{W}$ mapping a vector space \mathcal{V} into another vector space \mathcal{W} follows similar lines by satisfying the linearity assumptions that, for all v, v_1, v_2 in \mathcal{V} and scalars λ ,

$$\begin{aligned}\Gamma(v_1 + v_2) &= \Gamma v_1 + \Gamma v_2 \\ \Gamma(\lambda v) &= \lambda \Gamma v\end{aligned}\tag{2.67}$$

For example, if $\mathcal{V} = L_2^q[0, T]$ and $\mathcal{W} = L_2^p[0, T]$, T is finite and $H(t)$ is a $p \times q$ matrix with elements that are continuous in t , then the mapping $v \mapsto \Gamma v$ defined by the *Convolution Integral*

$$(\Gamma v)(t) = \int_0^t H(t-t')v(t')dt', \quad 0 \leq t \leq T,\tag{2.68}$$

is a well-defined linear operator. The *identity or unit operator* in \mathcal{V} is the linear operator $I : \mathcal{V} \mapsto \mathcal{V}$ defined by $Iv = v$ for all $v \in \mathcal{V}$. If both \mathcal{V} and \mathcal{W} are real vector spaces then a linear operator $\Gamma : \mathcal{V} \mapsto \mathcal{W}$ can be extended to a linear operator (again denoted by Γ) mapping the complexification \mathcal{V}^c into the complexification \mathcal{W}^c by the relation $\Gamma(u + iv) = \Gamma u + i\Gamma v$. Also two operators $\Gamma_1 : \mathcal{V} \mapsto \mathcal{V}$ and $\Gamma_2 : \mathcal{V} \mapsto \mathcal{V}$ are said to *commute* if

$$\Gamma_1 \Gamma_2 = \Gamma_2 \Gamma_1.\tag{2.69}$$

Linear operators can be associated with norms quite easily. More precisely, with the notation used above, suppose that both \mathcal{V} and \mathcal{W} are normed spaces, then the *operator norm of Γ* (induced by the norms in \mathcal{V} and \mathcal{W}) is defined to be

$$\|A\| = \sup_{v \neq 0} \frac{\|\Gamma v\|_{\mathcal{W}}}{\|v\|_{\mathcal{V}}} = \sup_{\|v\|=1} \frac{\|\Gamma v\|_{\mathcal{W}}}{\|v\|_{\mathcal{V}}}\tag{2.70}$$

If the norm is finite, the operator is said to be *bounded*. In all other cases, it is *unbounded*. The identity operator is bounded with induced norm $\|I\| = 1$.

The definition of the operator norm implies that, for all $v \in \mathcal{V}$,

$$\|\Gamma v\|_{\mathcal{W}} \leq \|\Gamma\| \|v\|_{\mathcal{V}}\tag{2.71}$$

which implies the fact that boundedness of an operator is equivalent to its continuity. In addition, it is easily shown that, if $\Gamma_2 : \mathcal{V} \mapsto \mathcal{W}$ and $\Gamma_1 : \mathcal{W} \mapsto \mathcal{Z}$ are two bounded linear operators between normed spaces, then the composite operator $\Gamma_1 \Gamma_2 : \mathcal{V} \mapsto \mathcal{Z}$ defined by $(\Gamma_1 \Gamma_2)v = \Gamma_1(\Gamma_2 v)$ has a norm bound

$$\|\Gamma_1 \Gamma_2\| \leq \|\Gamma_1\| \|\Gamma_2\|.\tag{2.72}$$

In a similar manner, the operator sum $\Gamma_1 + \Gamma_2$ defined by $(\Gamma_1 + \Gamma_2)v = \Gamma_1v + \Gamma_2v$ and multiplication by scalars $(\lambda\Gamma)v = \lambda(\Gamma v)$ rules make the set $\mathcal{L}(\mathcal{V}; \mathcal{W})$ of bounded linear operators from \mathcal{V} into \mathcal{W} into a normed vector space in its own right. If \mathcal{V} and \mathcal{W} are Banach spaces, then so is $\mathcal{L}(\mathcal{V}; \mathcal{W})$.

The *kernel* and *range* of an operator $\Gamma : \mathcal{V} \rightarrow \mathcal{W}$ play a vital role in analysis and are defined (as for matrices) using

$$\begin{aligned} \ker[\Gamma] &= \{v \in \mathcal{V} : \Gamma v = 0\} \\ \mathcal{R}[\Gamma] &= \{w \in \mathcal{W} : w = \Gamma v \text{ for some } v \in \mathcal{V}\} \end{aligned} \quad (2.73)$$

The operator is *injective* (or *one-to-one*) if its kernel is the single point $\{0\}$ and it is *surjective* (or *onto*) if its range is \mathcal{W} . It is *bijective* if it is both injective and surjective. If Γ is bijective and \mathcal{W} is finite dimensional, then it has a bounded, linear inverse $\Gamma^{-1} : \mathcal{W} \mapsto \mathcal{V}$ defined by the relation $\Gamma^{-1}(\Gamma v) = v$ for all $v \in \mathcal{V}$ or, more simply $\Gamma^{-1}\Gamma = I$ where I is the identity in \mathcal{V} . If $\mathcal{V} = \mathcal{W}$, then, in addition, $\Gamma\Gamma^{-1} = I$.

The notion of inverse familiar in matrix theory also has relevance to the interpretation of the idea of *inverse operators* but with more technical complexity if \mathcal{V} and \mathcal{W} are infinite dimensional. One concept that is central to the discussion is that of a *dense subset* S_1 of a subset S of a normed space \mathcal{V} . More precisely, S_1 is dense in S if, and only if, for every point $v \in S$ and for every $\varepsilon > 0$, there exists a point $v_\varepsilon \in S_1$ such that $\|v - v_\varepsilon\| < \varepsilon$. In effect, every point in S has a point in S_1 arbitrarily close to it. Three observations are related to this

1. If Γ is bounded, then its kernel is closed.
2. Γ being bounded does not necessarily imply that its range is closed.
3. If \mathcal{W} is finite dimensional and Γ is bounded, Γ has a closed range.

The second observation can be illustrated by the case of a convolution operator (2.68) mapping $L_2^q[0, T]$ into $L_2^p[0, T]$. The range of Γ is, at best, the set of continuous $p \times 1$ matrix valued functions which is known to be dense in $L_2^p[0, T]$.

The important point that follows from the above is that the range of the operator in infinite dimensional spaces has more complex properties than those observed for matrices. The consequences of this fact are many in number and include the possibility that the inverse of a bounded operator may exist but be unbounded. If Γ has a bounded (and hence continuous) inverse, it is said to be a *Homeomorphism*. In such cases it is easily seen that $1 \leq \|\Gamma\| \|\Gamma^{-1}\|$.

For matrices, the eigenvalues of a matrix A are defined to be scalars λ which ensure that $\lambda I - A$ has no inverse. The idea of eigenvalues requires careful generalization to the case of linear operators. More precisely, suppose that \mathcal{V} is a complex Banach space, then the *Resolvent Set* of a linear operator $\Gamma : \mathcal{V} \mapsto \mathcal{V}$ is defined to be the set of complex numbers λ where $\lambda I - \Gamma$ is bijective. As a consequence of the *Open Mapping Theorem*, for any such λ , the *Resolvent Operator* $(\lambda I - \Gamma)^{-1}$ is bounded. Using this construction, the *spectrum* (denoted by $\text{spec}[\Gamma]$) of Γ is defined to be the complement of the Resolvent Set and hence is the set of complex numbers λ where $\lambda I - \Gamma$ does not have a bounded inverse. This definition includes the *eigenvalues* of Γ (the so-called *Point Spectrum* defined by the existence of non-zero

eigenvectors/eigenfunctions $v \in \mathcal{V}$ such that $\Gamma v = \lambda v$) but also other points in what are termed the *Continuous and Residual Spectrum*. In finite dimensional spaces, the residual and continuous spectra are empty and the results of matrix algebra describe the spectrum completely in terms of eigenstructure. Finally,

1. the *spectral radius* of a bounded operator $\Gamma : \mathcal{V} \mapsto \mathcal{V}$ is defined by

$$r(\Gamma) = \sup\{|\lambda| : \lambda \in \text{spec}[\Gamma]\}, \quad (2.74)$$

a definition that reduces to that for matrices if $\mathcal{V} = \mathcal{C}^p$. In particular, if \mathcal{V} is a Banach space, then

$$r(\Gamma) = \lim_{k \rightarrow \infty} \|\Gamma^k\|^{1/k} \quad (2.75)$$

which relates the spectral radius to powers of Γ . As $\|\Gamma^k\| \leq \|\Gamma\|^k$ for all $k \geq 1$, this implies that

$$r(\Gamma) \leq \|\Gamma\| \quad (2.76)$$

and, for all $\varepsilon > 0$, there exists a real number $M_\varepsilon \geq 1$ such that

$$\|\Gamma^k\| \leq M_\varepsilon (r(\Gamma) + \varepsilon)^k. \quad (2.77)$$

2. The ideas of *functions of operators* and the *Spectral Mapping Theorem*, easily proven for matrices, can be extended to bounded operators from Banach spaces into Banach spaces. More precisely, if $f(z)$ has a power series expansion $\sum_{j=0}^{\infty} f_j z^j$ with radius of convergence R , then $f(\Gamma)$ is defined to be the operator $\sum_{j=0}^{\infty} f_j \Gamma^j$ which is convergent if $r(\Gamma) < R$. The spectrum of $f(\Gamma)$ is simply $\{z : z = f(\eta), \eta \in \text{spec}[\Gamma]\}$ or, more compactly,

$$\text{spec}[f(\Gamma)] = f(\text{spec}[\Gamma]). \quad (2.78)$$

For example, the *Resolvent* $(\lambda I - \Gamma)^{-1}$ has the power series representation

$$(\lambda I - \Gamma)^{-1} = \sum_{j=0}^{\infty} \lambda^{-(j+1)} \Gamma^j \quad (2.79)$$

which is convergent if Γ has spectral radius strictly less than $|\lambda|$. A sufficient condition for this is that $\|\Gamma\| < |\lambda|$. The spectrum of the Resolvent is $\{z : z = (\lambda - \eta)^{-1}, \eta \in \text{spec}[\Gamma]\}$.

The following result relates the spectral radius to the iterative learning control studies in this text. More precisely, the spectral radius describes convergence in norm of a simple, but typical, iteration formula.

Theorem 2.1 (Convergence in Norm and the Spectral Radius) *Let \mathcal{V} be a Banach space. Then, given an arbitrary starting vector $v_0 \in \mathcal{V}$ and a bounded linear operator $\Gamma : \mathcal{V} \mapsto \mathcal{V}$, the sequence $\{v_j\}_{j \geq 0}$ generated by the iteration $v_{j+1} = \Gamma v_j, j \geq 0$, converges (in norm) to zero if (a sufficient condition)*

$$r(\Gamma) < 1. \quad (2.80)$$

A sufficient condition for this to be true is that $\|\Gamma\| < 1$.

Proof Note, using induction, $v_j = \Gamma^j v_0$ for all $j \geq 0$. Using the notation above, the assumptions make possible the selection of $\varepsilon > 0$ such that $r(\Gamma) + \varepsilon < 1$. It follows that, as required,

$$\|v_j\| = \|\Gamma^j v_0\| \leq \|\Gamma^j\| \|v_0\| \leq M_\varepsilon (r(\Gamma) + \varepsilon)^j \rightarrow 0 \text{ as } j \rightarrow \infty. \quad (2.81)$$

The theorem is proved as the norm condition is sufficient to ensure the required condition on the spectral radius. \square

Note the conceptual similarity of this results to the familiar results from discrete time, sampled data systems control where asymptotic stability of $x_{j+1} = Ax_j$ is equivalent to the condition $r(A) < 1$ and hence is equivalent to the poles of the systems transfer function being inside the unit circle of the complex plane.

2.4 Hilbert Spaces

2.4.1 Inner Products and Norms

Although Banach spaces play a role in some areas of Control Theory and Optimization, the addition of geometrical structures plays an important role in algorithm design. The relevant structure is that of a *Hilbert Space*. More precisely, let \mathcal{V} be a real (respectively complex) Banach space endowed with an associated *inner product* $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \mapsto \mathcal{R}$ (respectively \mathcal{C}) possessing the properties that, for all u, v, w in \mathcal{V} and real (respectively complex) scalars λ ,

$$\begin{aligned} \langle u, v \rangle &= \langle v, u \rangle \text{ (respectively } \langle u, v \rangle = \overline{\langle v, u \rangle}), \\ \langle u, v + w \rangle &= \langle u, v \rangle + \langle u, w \rangle, \\ \langle u, \lambda v \rangle &= \lambda \langle u, v \rangle, \\ \langle v, v \rangle &\geq 0 \text{ and} \\ \langle v, v \rangle &= 0 \text{ if, and only if, } v = 0. \end{aligned} \quad (2.82)$$

Suppose also that the norm in \mathcal{V} can be computed from the inner product using the formula

$$\|v\| = \sqrt{\langle v, v \rangle}, \quad (2.83)$$

then \mathcal{V} is said to be a real (respectively complex) Hilbert Space. Note that, if it is necessary to identify the space or some other aspect of the formulae arising in the theory, the norm and inner product may be given subscripts as an aide memoire to the reader. For example, to identify the space being considered, both $\|v\|$ and $\langle u, v \rangle$ can be written in the form $\|v\|_{\mathcal{V}}$ and $\langle u, v \rangle_{\mathcal{V}}$.

Examples of Hilbert spaces include:

1. The space \mathcal{R}^p is a Hilbert space with inner product $\langle u, v \rangle = u^T Q v$ where Q is any symmetric, positive definite $p \times p$ matrix.
2. If $-\infty < a < b < +\infty$, then $L_2^p[a, b]$ is a Hilbert space with inner product

$$\langle u, v \rangle = \int_a^b u^T(t) Q(t) v(t) dt \quad (2.84)$$

where $Q(t)$ is any piecewise continuous $p \times p$ matrix satisfying an inequality of the form

$$\alpha_1 I_p \leq Q(t) \leq \alpha_2 I_p, \quad \text{for all } t \in [a, b] \quad (2.85)$$

and some scalars $0 < \alpha_1 \leq \alpha_2$. For example, if $\alpha \geq 0$ and $Q(t) = e^{2\alpha t} Q$ with Q a constant, symmetric, positive definite matrix with eigenvalues $0 < q_1 \leq q_2 \leq \dots \leq q_p$, then the conditions are satisfied with $\alpha_1 = q_1 e^{\alpha a}$ and $\alpha_2 = q_p e^{\alpha b}$.

Finally, the inner product has a number of useful additional properties including the *Cauchy-Schwarz Inequality* which takes the form, for any u, v in \mathcal{V}

$$|\langle u, v \rangle| \leq \|u\| \|v\| \quad (2.86)$$

with equality holding if, and only if, v is a multiple of u . That is, if and only if, $v = \lambda u$ for some scalar λ .

Also the inner product allows the introduction of ideas of orthogonality. More precisely, two vectors u, v in \mathcal{V} are said to be *orthogonal* if, and only if, $\langle u, v \rangle = 0$, a definition that is consistent with that used for Euclidean geometry in \mathcal{R}^p . The *orthogonal complement* of a vector subspace S of \mathcal{V} is denoted S^\perp where

$$S^\perp = \{v \in \mathcal{V} : \langle v, u \rangle = 0 \text{ for all } u \in S\} \quad (2.87)$$

S^\perp is a closed subspace. If, in addition, S is a closed subspace, then \mathcal{V} has the direct sum decomposition

$$\mathcal{V} = S \oplus S^\perp = \{w = u + v : u \in S, v \in S^\perp\}. \quad (2.88)$$

If $\{v_j\}_{j \geq 1}$ is a basis for \mathcal{V} and $\langle v_j, v_k \rangle = 0$ whenever $j \neq k$, then the basis set is an *orthogonal basis*. If, by suitable scaling (notably replacing each v_j by the normalized vector $v_j / \|v_j\|$), the basis set is said to be an *orthonormal basis* with the defining property that, for all $j \geq 1$ and $k \geq 1$,

$$\langle v_j, v_k \rangle = \delta_{jk} \quad \text{where} \quad (2.89)$$

the symbol δ_{jk} is the *Kronecker Delta* defined by $\delta_{jk} = 0$ if $j \neq k$ and is unity otherwise. Under these conditions, any vector $v \in \mathcal{V}$ has the form

$$v = \sum_{j=1}^{\infty} \alpha_j v_j \quad \text{with} \quad \alpha_j = \langle v_j, v \rangle, \text{ for } j \geq 1, \text{ and } \|v\|^2 = \sum_{j=1}^{\infty} |\alpha_j|^2 < \infty. \quad (2.90)$$

Finally, the ideas of inner products can be applied to some normed spaces that are not complete. Examples include the space $C[a, b]$ endowed with the inner product (and induced norm) used for the Hilbert space $L_2[a, b]$ or, more generally, a dense, but not complete, subspace of a Hilbert space. Such spaces are said to be *Pre-Hilbert Spaces*. The geometry of such spaces is identical to that of Hilbert spaces but results that rely on the convergence of Cauchy sequences (and hence the existence of limits) need to be carefully considered.

2.4.2 Norm and Weak Convergence

The convergence of sequences in Hilbert spaces is defined as in any normed space but it is often called *convergence in norm*, *convergence in the norm topology* or, more simply, *norm convergence*. This is because, in Hilbert spaces, another useful definition of convergence is that of *Weak Convergence*. More precisely, a sequence $\{v_j\}_{j \geq 0}$ in a Hilbert space \mathcal{V} is said to *converge weakly* to a vector $v_\infty \in \mathcal{V}$ if, and only if,

$$\lim_{j \rightarrow \infty} \langle f, v_\infty - v_j \rangle = 0 \quad \text{for all } f \in \mathcal{V}. \quad (2.91)$$

The Cauchy-Schwarz inequality immediately indicates that convergence in norm to v_∞ implies weak convergence to that vector. However, weak convergence of a sequence does not imply, necessarily, its convergence in norm.

The limit need only be valid on a dense subset of \mathcal{V} as,

Theorem 2.2 *Using the notation above, suppose that the sequence $\{v_j\}_{j \geq 0}$ is bounded in the sense that there exists a real scalar M such that $\|v_j\| \leq M$ for all $j \geq 0$. Suppose also that, for some dense subset $S \subset \mathcal{V}$*

$$\lim_{j \rightarrow \infty} \langle f, v_\infty - v_j \rangle = 0 \quad \text{for all } f \in S. \quad (2.92)$$

Then, $\{v_j\}_{j \geq 0}$ converges weakly to v_∞ .

Proof First note that $\|v_\infty - v_j\| \leq \|v_\infty\| + \|v_j\| \leq \|v_\infty\| + M$. Next write, for any $f_\varepsilon \in \mathcal{V}$,

$$\langle f, v_\infty - v_j \rangle = \langle f - f_\varepsilon, v_\infty - v_j \rangle + \langle f_\varepsilon, v_\infty - v_j \rangle. \quad (2.93)$$

Let $\varepsilon > 0$ be arbitrary and choose $f_\varepsilon \in S$ so that $\|f - f_\varepsilon\| < \varepsilon$, then the inequality

$$|\langle f, v_\infty - v_j \rangle| \leq \|f - f_\varepsilon\| \|v_\infty - v_j\| + |\langle f_\varepsilon, v_\infty - v_j \rangle| \quad (2.94)$$

indicates that

$$\limsup_{j \rightarrow \infty} |\langle f, v_\infty - v_j \rangle| \leq \varepsilon (\|v_\infty\| + M) \quad (2.95)$$

which proves the result as the left hand side is independent of $\varepsilon > 0$ which is arbitrary. \square

Another useful property of weak convergence is that of guaranteed convergence of subsequences. This is stated as follows and is usually associated with the notion of *weak compactness of the closed unit ball* $B(0; 1)$ in \mathcal{V} .

Theorem 2.3 (Ascoli's Theorem) *If \mathcal{V} is a Hilbert space and $S = \{v_j\}_{j \geq 0}$ is an infinite but bounded sequence of vectors, then S has a subsequence $S_1 = \{v_j\}_{k_j \geq 0}$ that converges weakly to some vector v_∞ in \mathcal{V} .*

The result states that *all* bounded sequences in any Hilbert space contain weakly convergent subsequences. If the convergent subsequence is removed from the original sequence, it will leave either a finite set (a situation which implies that the sequence itself converges weakly to v_∞) or an infinite sequence. In the second case, the remaining sequence is also bounded and hence (by Ascoli's Theorem) it, too, has a subsequence converging weakly to some (possibly different) weak limit $\hat{v}_\infty \in \mathcal{V}$. It is concluded that there is a possibility that S has many subsequences with different weak limits.

The essential property needed for weak convergence is boundedness of the sequence. The following result provides some insight into possibilities.

Theorem 2.4 (Weak Convergence and Operator Norms) *Any iteration $v_{j+1} = \Gamma v_j$, $j \geq 0$, in a real Hilbert space \mathcal{V} where $\Gamma : \mathcal{V} \rightarrow \mathcal{V}$ is a bounded linear operator with norm $\|\Gamma\| \leq 1$ generates a sequence $\{v_j\}_{k_j \geq 0}$ that is bounded in norm and has weakly convergent subsequences. If $\|\Gamma\| < 1$, then the sequence converges in norm to zero.*

Proof As $v_j = \Gamma^j v_0$, it follows that $\|v_j\| \leq \|\Gamma\|^j \|v_0\|$ which proves the result as the sequence is always bounded by $\|v_j\| \leq \|v_0\|$. Ascoli's Theorem then indicates the existence of a weak limit of some subsequence. Convergence in norm if $\|\Gamma\| < 1$ follows from the definitions. \square

Iterations of the form $v_{j+1} = \Gamma v_j$ appear regularly in this text. Theorems 2.1 and 2.4 above provide two conditions for some form of convergence. It is worth noting that the values of the spectral radius or norm of the operator Γ are central to the stated results. As $r(\Gamma) \leq \|\Gamma\|$, the use of the spectral radius will produce the best prediction of some form of convergence. This is particularly apparent in the case when $r(\Gamma) < 1 < \|\Gamma\|$ when the norm cannot be used to prove convergence

but the use of the spectral radius indicates convergence in norm to zero. The more difficult, but important, case that also plays a role in iterative control is the case when $r(\Gamma) = \|\Gamma\| = 1$ when weak convergence of subsequences is guaranteed by Theorem 2.4 but convergence in norm is not covered by either result.

2.4.3 Adjoint and Self-adjoint Operators in Hilbert Space

Suppose that $\Gamma : \mathcal{V} \mapsto \mathcal{W}$ is a bounded linear operator mapping a real or complex Hilbert space \mathcal{V} into a real or complex Hilbert space \mathcal{W} . The *Adjoint Operator* $\Gamma^* : \mathcal{W} \mapsto \mathcal{V}$ is the uniquely defined bounded linear operator mapping \mathcal{W} into \mathcal{V} and satisfying the identity, for all $u \in \mathcal{W}$ and $v \in \mathcal{V}$,

$$\langle u, \Gamma v \rangle_{\mathcal{W}} = \langle \Gamma^* u, v \rangle_{\mathcal{V}} \quad (2.96)$$

There are many general links between an operator and its adjoint. These include the additive, multiplicative and inversion rules

$$\begin{aligned} (\Gamma_1 + \Gamma_2)^* &= \Gamma_1^* + \Gamma_2^*, \\ (\Gamma_1 \Gamma_2)^* &= \Gamma_2^* \Gamma_1^* \text{ and, if } \mathcal{V} = \mathcal{W}, \quad (A^{-1})^* = (A^*)^{-1} \end{aligned} \quad (2.97)$$

(when the inverse exists). The cases of real and complex Hilbert spaces need a little care as, for $\Gamma : \mathcal{V} \mapsto \mathcal{W}$ and any scalar λ , the adjoint $(\lambda \Gamma)^* = \lambda \Gamma^*$ if \mathcal{V} is a real Hilbert space but equal to $\bar{\lambda} \Gamma^*$ if \mathcal{V} is a complex Hilbert space. Also

$$(\Gamma^*)^* = \Gamma. \quad (2.98)$$

A result that plays a role in the following text expresses the adjoint of a map into a product space in terms of adjoints of operators associated with each component.

Theorem 2.5 (The Adjoint of a Map into a Product Hilbert Space) *Let $\mathcal{V}, \mathcal{W}_1, \dots, \mathcal{W}_p$ be real Hilbert spaces and define the product Hilbert space $\mathcal{W}_1 \times \dots \times \mathcal{W}_p$ to be the product space with inner product and induced norm defined by*

$$\begin{aligned} \langle (w_1, \dots, w_p), (z_1, \dots, z_p) \rangle_{\mathcal{W}_1 \times \dots \times \mathcal{W}_p} &= \sum_{j=1}^p \langle w_j, z_j \rangle_{\mathcal{W}_j} \\ \text{and} \quad \|(w_1, \dots, w_p)\|_{\mathcal{W}_1 \times \dots \times \mathcal{W}_p}^2 &= \sum_{j=1}^p \|w_j\|_{\mathcal{W}_j}^2. \end{aligned} \quad (2.99)$$

Let the operator $G : \mathcal{V} \rightarrow \mathcal{W}_1 \times \dots \times \mathcal{W}_p$ be linear and bounded. Then G can be represented by the mapping, for all $v \in \mathcal{V}$,

$$Gv = (G_1 v, G_2 v, \dots, G_p v) \quad (2.100)$$

where $G_j : \mathcal{V} \rightarrow \mathcal{W}_j$ is linear and bounded. The adjoint map $G^* : \mathcal{W}_1 \times \cdots \times \mathcal{W}_p \rightarrow \mathcal{V}$ is the bounded linear operator defined by the relation,

$$G^*(w_1, w_2, \dots, w_p) = G_1^*w_1 + G_2^*w_2 + \cdots + G_p^*w_p, \quad (2.101)$$

where, for $1 \leq j \leq p$, $G_j^* : \mathcal{W}_j \rightarrow \mathcal{V}$ is the adjoint of G_j .

Proof The characterization of G in terms of the G_j follows easily from the linearity of G . The adjoint of G is identified from the equation

$$\begin{aligned} \langle (w_1, \dots, w_p), Gv \rangle_{\mathcal{W}_1 \times \cdots \times \mathcal{W}_p} &= \sum_{j=1}^p \langle w_j, G_j v \rangle_{\mathcal{W}_j} \\ &= \sum_{j=1}^p \langle G_j^* w_j, v \rangle_{\mathcal{V}} = \langle \sum_{j=1}^p G_j^* w_j, v \rangle_{\mathcal{V}}. \end{aligned} \quad (2.102)$$

The theorem is proved by comparing this with $\langle G^*(w_1, w_2, \dots, w_p), v \rangle_{\mathcal{V}}$. \square

An operator $\Gamma : \mathcal{V} \mapsto \mathcal{V}$ is *self adjoint* if, and only if, $\Gamma = \Gamma^*$. If Γ is self adjoint then $\langle u, \Gamma u \rangle_{\mathcal{V}}$ is always real. Γ is then said to be *positive* if $\langle u, \Gamma u \rangle \geq 0$ for all $u \in \mathcal{V}$, *positive definite* if it is positive and $\langle u, \Gamma u \rangle = 0$ if, and only if, $u = 0$ and *positive semi-definite* if it is positive but there exists a non-zero u such that $\langle u, \Gamma u \rangle = 0$. Positive commuting operators have special properties as follows:

Theorem 2.6 *If Γ_1, Γ_2 and Γ_3 are linear, bounded, self-adjoint, positive, commuting operators mapping a Hilbert space \mathcal{V} into itself, then*

$$\begin{aligned} \Gamma_1 \geq 0 \ \& \ \Gamma_2 \geq 0 \Rightarrow \Gamma_1 \Gamma_2 \geq 0 \\ \Gamma_1 \geq \Gamma_2 &\Rightarrow \Gamma_1 \Gamma_3 \geq \Gamma_2 \Gamma_3 \end{aligned} \quad (2.103)$$

The form of the adjoint operator depends on the spaces used and, in particular, on the form of inner product used. For example, matrix algebra proves that,

Theorem 2.7 (Adjoint of a Matrix Operator) *let \mathcal{V} be \mathbb{R}^p with inner product $\langle \hat{v}, v \rangle_{\mathcal{V}} = \hat{v}^T R v$ (where $R = R^T > 0$) and \mathcal{W} be \mathbb{R}^q with inner product $\langle \hat{w}, w \rangle_{\mathcal{W}} = \hat{w}^T Q w$ (where $Q = Q^T > 0$). Γ is a $p \times q$ real matrix with adjoint Γ^* satisfying, for all u and w ,*

$$w^T Q \Gamma v = (\Gamma^* w)^T R v \quad \text{and hence} \quad \Gamma^* = R^{-1} \Gamma^T Q. \quad (2.104)$$

In particular, when $R = I_p$ and $Q = I_q$, the adjoint is simply the transpose of the matrix Γ .

Note: it is conventional to use the $*$ notation to denote the adjoint operator but it is also often used to denote the complex conjugate transpose of a matrix. There is a possibility of confusion from time to time but careful attention to the context of the analysis should easily resolve any ambiguity.

The operator $\Gamma^* \Gamma$ is self adjoint. It is also positive as $\langle u, \Gamma^* \Gamma u \rangle = \langle \Gamma u, \Gamma u \rangle = \|\Gamma u\|^2 \geq 0$. As a consequence,

Theorem 2.8 (Invertibility and Positivity of Operators) *With the above notation,*

1. *The operator $\Gamma^* \Gamma : \mathcal{V} \rightarrow \mathcal{V}$ is positive definite if, and only if, $\ker[\Gamma] = \{0\}$.*
2. *The operator $\Gamma : \mathcal{V} \rightarrow \mathcal{W}$ has a bounded inverse if, and only if, there exists a constant $\alpha > 0$ such that $\Gamma^* \Gamma \geq \alpha I_{\mathcal{V}}$ and $\Gamma \Gamma^* \geq \alpha I_{\mathcal{W}}$.*

These properties link invertibility to positivity as follows,

Theorem 2.9 (Invertibility and Positivity of Operators) *An operator $\Gamma : \mathcal{V} \rightarrow \mathcal{V}$ has a bounded inverse on the Hilbert space \mathcal{V} if there exists a real number $\varepsilon_0 > 0$ such that*

$$\Gamma + \Gamma^* \geq \varepsilon_0^2 I. \quad (2.105)$$

Proof Noting that, for any real scalar λ ,

$$\begin{aligned} 0 &\leq (I - \lambda \Gamma^*)(I - \lambda \Gamma) = I - \lambda(\Gamma + \Gamma^*) + \lambda^2 \Gamma^* \Gamma \\ &\leq (1 - \lambda \varepsilon_0^2)I + \lambda^2 \Gamma^* \Gamma. \end{aligned} \quad (2.106)$$

Exactly the same relationship for $\Gamma \Gamma^*$ is obtained using $(I - \lambda \Gamma)(I - \lambda \Gamma^*)$ so that the positivity condition of the previous result holds by choosing λ so that $\alpha = \lambda^{-2}(\lambda \varepsilon_0^2 - 1) > 0$. \square

The operator norms $\|\Gamma\|$ and $\|\Gamma^*\|$ are related by the expression,

$$\|\Gamma\| = \|\Gamma^*\| \quad (2.107)$$

and the range and kernels of the operators satisfy the orthogonality relations

$$\begin{aligned} (a) \quad \mathcal{R}[\Gamma^*]^\perp &= \ker[\Gamma] \text{ and hence } \mathcal{R}[\Gamma]^\perp = \ker[\Gamma^*] \\ (b) \quad \mathcal{R}[\Gamma] &= \ker[\Gamma^*]^\perp \text{ and hence } \mathcal{R}[\Gamma^*] = \ker[\Gamma]^\perp \end{aligned} \quad (2.108)$$

from which the *Projection Theorem in Hilbert space* (Theorem 2.17) gives

$$\mathcal{V} = \overline{\mathcal{R}[\Gamma^*]} \oplus \ker[\Gamma] \quad \text{and} \quad \mathcal{W} = \overline{\mathcal{R}[\Gamma]} \oplus \ker[\Gamma^*] \quad (2.109)$$

A following result provides an important property of a Hilbert space in terms of the range of an operator and the kernel of its adjoint. The result has close links to the above but, more formally,

Theorem 2.10 (Denseness and the Orthogonal Complement of the Kernel) *Suppose that $\Gamma : \mathcal{V} \mapsto \mathcal{W}$ where \mathcal{V} and \mathcal{W} are Hilbert spaces. Then the range space $\mathcal{R}[\Gamma^*]$ is dense in \mathcal{V} if, and only if, $\ker[\Gamma] = \{0\}$.*

Proof If $\ker[\Gamma] = \{0\}$, suppose that $\mathcal{R}[\Gamma^*]$ is not dense. It follows that its closure is a proper closed subspace S of \mathcal{V} with an orthogonal complement S^\perp containing non-zero vectors. Let $v \in S^\perp$ be non-zero and write $\langle v, \Gamma^* w \rangle = 0$ for all $w \in \mathcal{W}$. It follows that $\langle \Gamma v, w \rangle = 0$ for all $w \in \mathcal{W}$ so that (choosing $w = \Gamma v$) $\Gamma v = 0$

which contradicts the assumption that $\ker[\Gamma] = \{0\}$. Next suppose that $\mathcal{R}[\Gamma^*]$ is dense. It follows that the condition $\langle v, \Gamma^*w \rangle = 0$ for all $w \in \mathcal{W}$ implies that $v = 0$ which trivially leads to $\mathcal{R}[\Gamma^*]^\perp = \ker[\Gamma] = \{0\}$ as required. \square

The properties of the range spaces of an operator and its adjoint are also connected as follows:

Theorem 2.11 (Closed Range Theorem) *If $\Gamma : \mathcal{V} \rightarrow \mathcal{W}$ is a bounded linear operator between Hilbert spaces, then Γ has a closed range in \mathcal{W} if, and only if, the range of the adjoint Γ^* is closed in \mathcal{V} .*

The norm of a self adjoint operator $\Gamma : \mathcal{V} \rightarrow \mathcal{V}$ is related to the values taken by an associated quadratic form. More precisely its norm can be computed from the parameters

$$\begin{aligned} a &= \inf\{\langle u, \Gamma u \rangle : u \in \mathcal{V} \text{ and } \|u\| = 1\} \\ \text{and } b &= \sup\{\langle u, \Gamma u \rangle : u \in \mathcal{V} \text{ and } \|u\| = 1\} \\ \text{to be } \|\Gamma\| &= \max\{|a|, |b|\} \\ \text{and, in particular, } \|\Gamma\| &= r(\Gamma). \end{aligned} \quad (2.110)$$

This expression can be written in the form, where I is the identity operator,

$$aI \leq \Gamma \leq bI \quad (2.111)$$

which forms the basis of the theorem

Theorem 2.12 (Invertibility of Self Adjoint Operators) *With the notation used above, suppose that $\Gamma : \mathcal{V} \rightarrow \mathcal{V}$ is self adjoint. Then the spectrum of Γ contains only real numbers in the closed interval $[a, b]$. In particular, Γ has a bounded inverse if $ab > 0$.*

A useful relationship valid when Γ is self adjoint and positive is

$$\|\Gamma\| = \sup\{\langle u, \Gamma u \rangle : u \in \mathcal{V} \text{ and } \|u\| = 1\} \quad (2.112)$$

which immediately yields the result that, for any bounded $\Gamma : \mathcal{V} \rightarrow \mathcal{W}$,

$$\|\Gamma^* \Gamma\| = \|\Gamma\|^2. \quad (2.113)$$

A useful consequence of this is that

Theorem 2.13 (Norm of a Matrix Operator) *Using the notation and assumptions of Theorem 2.7,*

$$\|\Gamma\|^2 = \sup\{\langle u, \Gamma^* \Gamma u \rangle : u \in \mathcal{V} \text{ and } \|u\| = 1\} = r(\Gamma^* \Gamma) \quad (2.114)$$

where $r(\Gamma^*\Gamma)$ is the spectral radius of $\Gamma^*\Gamma = R^{-1}\Gamma^T Q\Gamma$. Moreover, for all choices of Q and R , $\|\Gamma\|$ is the largest singular value of $\Gamma_{QR} = Q^{\frac{1}{2}}\Gamma R^{-\frac{1}{2}}$ and

$$\|\Gamma^*\| = \|\Gamma\|. \quad (2.115)$$

Proof The proof follows from Theorem 2.7 using Lagrange multiplier techniques to solve 2.112 regarded as a function minimization problem. Next, it is easy to see that $\det(\lambda I_p - \Gamma\Gamma^*) = \lambda^{p-q} \det(\lambda I_q - \Gamma^*\Gamma)$ so that the eigenvalues of $\Gamma^*\Gamma$ and $\Gamma\Gamma^*$ differ, at most, only by a number of zero eigenvalues. Finally, the eigenvalues of $\Gamma^*\Gamma$ are the squares of the singular values of $Q^{\frac{1}{2}}\Gamma R^{\frac{1}{2}}$ as $\Gamma^*\Gamma = R^{-\frac{1}{2}} \left[\left(Q^{\frac{1}{2}}\Gamma R^{-\frac{1}{2}} \right)^T Q^{\frac{1}{2}}\Gamma R^{-\frac{1}{2}} \right] R^{\frac{1}{2}}$. That is $\Gamma^*\Gamma$ and $\Gamma_{QR}^T \Gamma_{QR}$ are related by a similarity transformation. \square

If $\Gamma : \mathcal{Y} \mapsto \mathcal{Y}$ is positive and self adjoint there exists a unique positive, self adjoint operator $\hat{\Gamma} : \mathcal{Y} \mapsto \mathcal{Y}$ with the property that $\Gamma = \hat{\Gamma}\hat{\Gamma}$. For this reason, $\hat{\Gamma}$ is said to be the unique *positive square root* of Γ and is written $\hat{\Gamma} = \Gamma^{1/2}$. The bounded, positive, self-adjoint linear operator $\Gamma^{1/2}$ has the properties that it commutes with every operator that commutes with Γ and

$$\Gamma^{1/2} = (\Gamma^{1/2})^* \geq 0, \quad \ker[\Gamma^{1/2}] = \ker[\Gamma] \quad \text{and} \quad \mathcal{R}[\Gamma] \subset \mathcal{R}[\Gamma^{1/2}] \quad (2.116)$$

so that, in particular, $\Gamma^{1/2}$ is positive definite if, and only if, Γ is positive definite.

The spectrum of a self adjoint operator Γ lies in the closed ball $B_c(0; r(\Gamma)) \subset B_c(0; \|\Gamma\|)$. Using (2.110), the spectrum of $\Gamma - \frac{b+a}{2}I$ lies in the closed ball $B_c(0, \frac{b-a}{2})$ and hence, using the spectral mapping Theorem, the spectrum of Γ lies in the shifted closed ball $\frac{b+a}{2} + B_c(0, \frac{b-a}{2})$. In particular,

Theorem 2.14 (Invertibility of $(I + \Gamma)^{-1}$) *Suppose that $\Gamma : \mathcal{Y} \rightarrow \mathcal{Y}$ where \mathcal{Y} is a real Hilbert space is bounded, self adjoint and positive. Then $I + \Gamma$ is a bijection and the inverse operator $(I + \Gamma)^{-1}$ is well-defined and bounded.*

Proof Using the discussion preceding this result, $a = 0$ and $b = \|\Gamma\|$ and hence

$$\text{spec}[\Gamma] \subset \frac{\|\Gamma\|}{2} + B_c(0, \frac{\|\Gamma\|}{2}). \quad (2.117)$$

The proof is now complete as -1 is not in the spectrum of Γ . \square

Note: Operators of this type play a central role in Iterative Algorithms.

Finally, useful conditions for $\Gamma_1\Gamma_2$ to be self adjoint can be stated as follows

Theorem 2.15 (When is $\Gamma_1\Gamma_2$ self adjoint?) *Suppose that the two self-adjoint operators Γ_1 and Γ_2 map a real Hilbert space \mathcal{Y} into itself and that Γ_2 is positive definite. Then the product $\Gamma_1\Gamma_2$ is self-adjoint if the inner product in \mathcal{Y} is replaced by the new inner product*

$$\langle u, v \rangle_0 = \langle u, \Gamma_2 v \rangle_{\mathcal{Y}}. \quad (2.118)$$

The two topologies are equivalent if there exists a real scalar $\varepsilon_0^2 > 0$ such that $\Gamma_2 \geq \varepsilon_0^2 I$.

Proof The bilinear form $\langle u, v \rangle_0$ satisfies all the requirements of an inner product and its associated norm $\| \cdot \|_0$. The self-adjoint property follows as

$$\langle u, \Gamma_1 \Gamma_2 v \rangle_0 = \langle u, \Gamma_2 \Gamma_1 \Gamma_2 v \rangle_{\mathcal{Y}} = \langle \Gamma_1 \Gamma_2 u, \Gamma_2 v \rangle_{\mathcal{Y}} = \langle \Gamma_1 \Gamma_2 u, v \rangle_0 \quad (2.119)$$

Finally, the existence of $\varepsilon_0^2 > 0$ ensures the topological equivalence of the two norms follows as

$$\varepsilon_0^2 \|u\|_{\mathcal{Y}}^2 \leq \langle u, \Gamma_2 u \rangle_{\mathcal{Y}} = \|u\|_0^2 \leq \|\Gamma_0\| \|u\|_{\mathcal{Y}}^2. \quad (2.120)$$

□

Note: This result plays a role in the analysis of the convergence and robustness of many of the algorithms in the following chapters.

2.5 Real Hilbert Spaces, Convex Sets and Projections

The structure of real Hilbert spaces provides a powerful set of results related to optimization. These results are expressed in terms of *Projection onto Convex Sets*. A convex set $S \subset \mathcal{V}$ in a real Hilbert space \mathcal{V} is any set satisfying the condition that, for any two points u, v in S , the vector

$$w = \lambda u + (1 - \lambda)v \in S \quad \text{for all } \lambda \in [0, 1]. \quad (2.121)$$

(where λ is a real number). The vector w is said to be a *convex combination* of u and v . The *Convex Hull* of a set $S \subset \mathcal{V}$ is the smallest convex set containing S .

Suppose that v_0 is an arbitrary point of \mathcal{V} and consider the problem of finding the point in a convex set S that is closest to v_0 . This problem can be written formally as the solution (if it exists) of the optimization problem

$$v_1 = \arg \min \{ \|v_0 - v\| : v \in S \} \quad (2.122)$$

That is, v_1 is the vector $v \in S$ that minimizes the norm $\|v_0 - v\|$ and hence is the nearest point in S to v_0 . For visualization purposes, v_1 can be thought of as the projection of v_0 onto the set S . In general, it is possible that no such point exists but, for many problems in practice, a solution does exist. The most useful theorem characterizing the existence of v_1 and its relationship to v_0 is as follows:

Theorem 2.16 (Minimum Distance to a Closed Convex Set) *Suppose that S is a closed, convex set in the real Hilbert space \mathcal{V} . If $v_0 \in \mathcal{V}$, then the optimization Problem (2.122) has a unique solution $v_1 \in S$. A necessary and sufficient condition for v_1 to be that solution is that*

$$\langle v - v_1, v_1 - v_0 \rangle_{\mathcal{V}} \geq 0 \text{ for all } v \in S. \quad (2.123)$$

A particular case of interest is when S is a closed vector subspace of \mathcal{V} .

Theorem 2.17 (The Projection Theorem in Hilbert Space) *Suppose that S is a closed vector subspace in the real Hilbert space \mathcal{V} . If $v_0 \in \mathcal{V}$, then the optimization Problem (2.122) has a unique solution $v_1 \in S$. A necessary and sufficient condition for v_1 to be that solution is that the following orthogonality condition is true,*

$$\langle v - v_1, v_1 - v_0 \rangle_{\mathcal{V}} = 0 \text{ for all } v \in S. \quad (2.124)$$

In particular, as $v - v_1 \in S$ is arbitrary, the condition reduces to

$$\langle v, v_1 - v_0 \rangle_{\mathcal{V}} = 0 \text{ for all } v \in S \quad (2.125)$$

which is simply the requirement that $v_1 - v_0$ is orthogonal to every vector in S .

Proof The existence and uniqueness of v_1 follows from the previous theorem as does the requirement that $\langle v - v_1, v_1 - v_0 \rangle_{\mathcal{V}} \geq 0$ for all $v \in S$. Suppose that there exists a vector $v \in S$ such that $\langle v - v_1, v_1 - v_0 \rangle_{\mathcal{V}} > 0$, then, noting that $\hat{v} = -v + 2v_1 \in S$, a simple computation indicates that $\langle \hat{v} - v_1, v_1 - v_0 \rangle_{\mathcal{V}} < 0$ contradicting the assumption that v_1 solves the problem. \square

The case of S being a vector subspace gives rise to the notion of a *Projection Operator*. More precisely, using the notation of the Projection Theorem 2.17, the computation $v_0 \mapsto v_1$ defines a mapping $P_S : \mathcal{V} \rightarrow S$. The orthogonality condition also indicates that the mapping is linear and hence P_S is a linear operator called the *Orthogonal Projection Operator onto S* . It is bounded as, writing $v_1 = P_S v_0$, using the orthogonality condition, and noting that $0 \in S$,

$$\begin{aligned} \|v_0\|^2 &= \|(v_0 - v_1) + v_1\|^2 = \|v_0 - v_1\|^2 + 2\langle v_0 - v_1, v_1 \rangle + \|v_1\|^2 \\ &= \|v_0 - v_1\|^2 + \|v_1\|^2 \geq \|v_1\|^2 \end{aligned} \quad (2.126)$$

so that, together with the observation that $P_S v_0 = v_0$ if, and only if, $v_0 \in S$, gives

$$\|P_S\| = 1. \quad (2.127)$$

From the definitions $P_S^2 = P_S$ and hence $\ker[I - P_S] = S$. In addition, any vector $v_0 \in \mathcal{V}$ has a unique decomposition of the form $v_0 = v_1 + (v_0 - v_1) = P_S v_0 + (I - P_S)v_0$ where $v_1 = P_S v_0$ is orthogonal to S and hence $(I - P_S)v_0$. As a consequence, \mathcal{V} has a direct sum decomposition of the form

$$\mathcal{V} = S \oplus S^\perp \text{ where } S = \ker[I - P_S] \text{ and } S^\perp = \ker[P_S]. \quad (2.128)$$

In particular, for any u and v in \mathcal{V} , it follows that $\langle P_S u, v \rangle = \langle P_S u, P_S v \rangle = \langle u, P_S v \rangle$ so that P_S is self adjoint and positive (but not strictly positive).

Finally, another form of convex set that plays a useful role in the analysis of linear control systems is that of a *Linear Variety*, namely a convex set S that is constructed by a translation of a subspace \mathcal{W} . The resultant set is denoted by $S = a + \mathcal{W}$ where $a \in \mathcal{V}$ defines the translation and

$$S = \{v : v = a + w \text{ for some } w \in \mathcal{W}\}. \quad (2.129)$$

Note that the choice of a is not unique as it can be replaced by any vector $a + w_0$ with $w_0 \in \mathcal{W}$. As $v - v_0 = (v - a) - (v_0 - a)$ and $v - a \in \mathcal{V}$, the solution of the problem in this case can be expressed in the form

$$v_1 = a + P_{\mathcal{W}}(v_0 - a) \quad (2.130)$$

A useful example of a closed linear variety is the set

$$S = \{u : r = Gu + d\} = u_0 + \mathcal{W}, \quad \text{with } \mathcal{W} = \ker[G] \quad (2.131)$$

where $G : \mathcal{U} \rightarrow \mathcal{Y}$ is linear and bounded, \mathcal{U} and \mathcal{Y} are real Hilbert spaces, $r \in \mathcal{Y}$, $d \in \mathcal{Y}$ and $u_0 \in \mathcal{U}$ is any point (assumed to exist) satisfying $r = Gu_0 + d$. Another example of a closed linear variety is that of a *closed Hyperplane* defined by taking $\mathcal{Y} = \mathcal{R}$ and G as the map $G : u \mapsto \langle \alpha, u \rangle$ (for some $\alpha \in \mathcal{U}$) and setting

$$S = \{u : \langle \alpha, u \rangle = c\} \quad (2.132)$$

where both $\alpha \in \mathcal{U}$ and the real number c are specified. If $u_0 \in S$, then $S = \{u : \langle \alpha, u - u_0 \rangle = 0\}$ which identifies the set of vectors $u - u_0$ as that of all vectors orthogonal to α . A *Separating Hyperplane* separating two sets S_1 and S_2 in a real Hilbert space \mathcal{V} is a hyperplane of the above type where $\langle \alpha, u \rangle \geq c$ for all $u \in S_1$ and $\langle \alpha, u \rangle \leq c$ for all $u \in S_2$ or vice versa. An example of a separating hyperplane is obtained from the result describing the minimum distance from $v_0 \in \mathcal{V}$ to a closed convex set S . More precisely, suppose that v_0 is not in S . From Theorem 2.122, the hyperplane $\langle v - v_1, v_1 - v_0 \rangle = 0$ is a hyperplane that separates the point set $\{v_0\}$ from S . Separating hyperplanes are not unique as the hyperplane $\langle v - \lambda v_1 - (1 - \lambda)v_0, v_1 - v_0 \rangle = 0$ is also a separating hyperplane when $\lambda \in [0, 1]$.

2.6 Optimal Control Problems in Hilbert Space

Quadratic optimization problems for linear systems play a central role in this text. In general, signals will be regarded as being vectors in suitable real Hilbert spaces. Operators will be used to represent systems behaviour by providing a linear relationship between system output signals and its input signals. The minimization of

objective functions created using quadratic functions of signal norms is the chosen mechanism for creating new control algorithms with known properties. The use of this level of abstraction will be seen to provide solutions that cover the solution of a wide range of problems of interest ranging from the case of continuous dynamics to sampled data systems to multi-rate systems and many other useful situations. The details are provided in the following chapters but the general form of the solution can be derived using the mathematical methods already described in this chapter.

A system will be described as a mapping between a set \mathcal{U} of *input signals* u and a set \mathcal{Y} of resultant *output signals* y . Both \mathcal{U} and \mathcal{Y} are taken to be normed vector spaces. A *Linear System* is characterized by a bounded, linear operator $G : \mathcal{U} \rightarrow \mathcal{Y}$ and the input to output mapping is defined by a relation of the form

$$y = Gu + d \quad (2.133)$$

where d represents the output behaviour when the input $u = 0$ (and hence, typically, behaviours due to initial condition and disturbances). A general form of *Linear, Quadratic, Optimal Control Problem* can be defined as the computation of the input that minimizes the *Objective Function* (often called a *Performance Index* or *Performance Criterion*)

$$J(u) = \|r - y\|_{\mathcal{Y}}^2 + \|u_0 - u\|_{\mathcal{U}}^2 \quad (2.134)$$

subject to the constraint that y and u are linked by $y = Gu + d$. The vectors $u_0 \in \mathcal{U}$ and $r \in \mathcal{Y}$ are assumed to be known and the problem is interpreted as an attempt to reduce the variation of the output from the specified signal r whilst not using input signals that deviate too much from u_0 . The relative weighting of these two objectives is reflected in the choice of norms in \mathcal{Y} and \mathcal{U} .

The solution to this problem when both \mathcal{U} and \mathcal{Y} are real Hilbert spaces is particularly valuable. Denote the adjoint operator of G by G^* and use the notation

$$\begin{aligned} e = r - y = r - Gu - d \quad \text{and} \quad e_0 = r - y_0 \\ \text{where} \quad y_0 = Gu_0 + d \end{aligned} \quad (2.135)$$

is the output response to the input u_0 .

Theorem 2.18 (Solution of the Optimal Control Problem) *With the problem definition given above, the input-output pair (y, u) minimizing the objective function (2.134) subject to the constraint (2.133) is given by the implicit formulae*

$$u = u_0 + G^*e \quad (2.136)$$

As a consequence,

$$e = (I + GG^*)^{-1}e_0 \quad \text{and hence} \quad u = u_0 + G^*(I + GG^*)^{-1}e_0. \quad (2.137)$$

In particular, the minimum value of $J(u)$ can be computed to be

$$\min J(u) = \langle e_0, (I + GG^*)^{-1}e_0 \rangle_{\mathcal{Y}}. \quad (2.138)$$

Proof Two alternative proofs are provided in the next two subsections. \square

The proofs depend on the material in the previous sections. In particular, they depend on the algebraic properties of inner products, the properties of the adjoint operator, the Projection Theorem and the identities

$$\begin{aligned} G^*(I + GG^*)^{-1} &= (I + G^*G)^{-1}G^*, & G(I + G^*G)^{-1}G^* &= GG^*(I + GG^*)^{-1}, \\ &\text{and } (I + GG^*)^{-1} + GG^*(I + GG^*)^{-1} &= I. \end{aligned} \quad (2.139)$$

The inverses $(I + GG^*)^{-1}$ and $(I + G^*G)^{-1}$ exist and are bounded due to the positivity of GG^* and G^*G and the resultant lower bounds $I + GG^* \geq I > 0$ and $I + G^*G \geq I > 0$.

2.6.1 Proof by Completing the Square

Note that $G(u - u_0) = -(e - e_0)$ and consider the following inner product

$$\begin{aligned} \gamma &= \langle u - u_0 - G^*(I + GG^*)^{-1}e_0, (I + G^*G)(u - u_0 - G^*(I + GG^*)^{-1}e_0) \rangle_{\mathcal{U}} \\ &\geq \|u - u_0 - G^*(I + GG^*)^{-1}e_0\|_{\mathcal{U}}^2 \geq 0, \end{aligned} \quad (2.140)$$

noting that it is equal to zero if, and only if $u - u_0 - G^*(I + GG^*)^{-1}e_0 = 0$. If this condition is valid then, operating on the equation with G gives $e = e_0 - GG^*(I + GG^*)^{-1}e_0 = (I + GG^*)^{-1}e_0$ which would prove (2.137) and hence (2.136). It remains to prove therefore that $\gamma = 0$.

The inner product can be written as

$$\begin{aligned} \gamma &= \langle u - u_0 - (I + G^*G)^{-1}G^*e_0, (I + G^*G)(u - u_0 - (I + G^*G)^{-1}G^*e_0) \rangle_{\mathcal{U}} \\ &= \langle u - u_0, (I + G^*G)(u - u_0) \rangle_{\mathcal{U}} - 2\langle u - u_0, G^*e_0 \rangle_{\mathcal{U}} + \langle G^*e_0, (I + G^*G)^{-1}G^*e_0 \rangle_{\mathcal{U}} \\ &= \|u - u_0\|_{\mathcal{U}}^2 + \|e - e_0\|_{\mathcal{Y}}^2 + 2\langle e - e_0, e_0 \rangle_{\mathcal{Y}} + \langle G^*e_0, (I + G^*G)^{-1}G^*e_0 \rangle_{\mathcal{U}} \\ &= \|u - u_0\|_{\mathcal{U}}^2 + \|e\|_{\mathcal{Y}}^2 + \|e_0\|_{\mathcal{Y}}^2 - 2\langle e, e_0 \rangle_{\mathcal{Y}} + 2\langle e - e_0, e_0 \rangle_{\mathcal{Y}} \\ &\quad + \langle G^*e_0, (I + G^*G)^{-1}G^*e_0 \rangle_{\mathcal{U}} \\ &= J(u) + \langle G^*e_0, (I + G^*G)^{-1}G^*e_0 \rangle_{\mathcal{U}} - \|e_0\|_{\mathcal{Y}}^2 \\ &= J(u) + \langle e_0, G(I + G^*G)^{-1}G^*e_0 \rangle_{\mathcal{Y}} - \|e_0\|_{\mathcal{Y}}^2 \\ &= J(u) + \langle e_0, (GG^*(I + G^*G)^{-1} - I)e_0 \rangle_{\mathcal{Y}} \\ &= J(u) - \langle e_0, (I + GG^*)^{-1}e_0 \rangle_{\mathcal{Y}} \end{aligned} \quad (2.141)$$

the second term being independent of (y, u) . It follows that $J(u)$ is minimized if, and only if, $\gamma = 0$. This first proof of Theorem 2.18 is now complete. \square

2.6.2 Proof Using the Projection Theorem

An alternative derivation of the solution uses the Projection Theorem in the product space $\mathcal{Y} \times \mathcal{U}$ of input/output pairs (y, u) regarded as a real Hilbert space with inner product (and associated norm) defined using

$$\begin{aligned} \langle (z, v), (y, u) \rangle_{\mathcal{Y} \times \mathcal{U}} &= \langle z, y \rangle_{\mathcal{Y}} + \langle v, u \rangle_{\mathcal{U}} \\ \|(y, u)\|_{\mathcal{Y} \times \mathcal{U}} &= \sqrt{\|y\|_{\mathcal{Y}}^2 + \|u\|_{\mathcal{U}}^2}. \end{aligned} \quad (2.142)$$

With this notation, $J(u) = \|(r, u_0) - (y, u)\|_{\mathcal{Y} \times \mathcal{U}}^2$ and the optimal control problem is that of finding the pair (y_1, u_1) that solves the minimum norm problem

$$\begin{aligned} (y_1, u_1) &= \arg \min \{ \|(r, u_0) - (y, u)\|^2 : (y, u) \in S \} \\ \text{where } S &= \{(y, u) : y = Gu + d\} \end{aligned} \quad (2.143)$$

is a linear variety in $\mathcal{Y} \times \mathcal{U}$. It is closed as any sequence $\{(y_j, u_j)\}_{j \geq 0}$ in S converging (in the norm topology) to a point (y, u) has the property that $\{y_j\}_{j \geq 0}$ converges to y in \mathcal{Y} and $\{u_j\}_{j \geq 0}$ converges to u in \mathcal{U} . Also, as $y_j = Gu_j + d$, $j \geq 0$,

$$\begin{aligned} 0 \leq \|y - Gu - d\|_{\mathcal{Y}} &= \|(y - y_j) - G(u - u_j) + (y_j - Gu_j - d)\|_{\mathcal{Y}} \\ &\leq \|(y - y_j) - G(u - u_j)\|_{\mathcal{Y}} + \|y_j - Gu_j - d\|_{\mathcal{Y}} \\ &= \|(y - y_j) - G(u - u_j)\|_{\mathcal{Y}} \\ &\leq \|y - y_j\|_{\mathcal{Y}} + \|G\| \|u - u_j\|_{\mathcal{U}} \end{aligned} \quad (2.144)$$

which tends to zero as $j \rightarrow \infty$. Hence $y = Gu + d$ which proves that $(y, u) \in S$.

Applying the projection theorem, the solution (y_1, u_1) of the optimal control problem satisfies

$$\langle (z, v) - (y_1, u_1), (y_1, u_1) - (r, u_0) \rangle_{\mathcal{Y} \times \mathcal{U}} = 0 \quad \text{for all } (z, v) \in S. \quad (2.145)$$

This equation is just

$$\langle z - y_1, y_1 - r \rangle_{\mathcal{Y}} + \langle v - u_1, u_1 - u_0 \rangle_{\mathcal{U}} = 0 \quad (2.146)$$

Using the equations $y_1 = Gu_1 + d$, $z = Gv + d$ and $e = r - y_1$ then gives

$$\langle v - u_1, -G^*e \rangle_{\mathcal{Y}} + \langle v - u_1, u_1 - u_0 \rangle_{\mathcal{U}} = \langle v - u_1, u_1 - u_0 - G^*e \rangle_{\mathcal{U}} = 0 \quad (2.147)$$

for all $v \in \mathcal{U}$. Choosing $v = 2u_1 - u_0 - G^*e$, it follows that $\|u_1 - u_0 - G^*e\|^2 = 0$ which proves the result using the same algebraic manipulations as those used in the previous subsection and the computation of the minimum value of the objective function as follows,

$$\|r - y_1\|_{\mathcal{Y}}^2 + \|u_0 - u_1\|_{\mathcal{U}}^2 = \langle e, (I + GG^*)e \rangle_{\mathcal{Y}} = \langle e_0, (I + GG^*)^{-1}e_0 \rangle_{\mathcal{Y}}. \quad (2.148)$$

This completes the second proof of Theorem 2.18. \square

2.6.3 Discussion

The solution of the optimal control problem described above provides a formal approach to the solution of a wide class of problems following the process summarized as the steps,

1. Identify the vector space \mathcal{U} from which the inputs signals are to be chosen.
2. Choose an inner product and norm for \mathcal{U} that ensures that it is a real Hilbert space and also reflects the physical importance of signals.
3. Identify the vector space \mathcal{Y} containing the outputs signals.
4. Choose an inner product and norm for \mathcal{Y} that ensures that it is a real Hilbert space and also reflects the physical importance of signals.
5. Characterize the system as a bounded linear mapping G from \mathcal{U} into \mathcal{Y} and identify the form of its adjoint operator G^* .
6. Write the defining relationship for the optimal solution in the implicit form $u = u_0 + G^*e$ with $e = r - y$ and find a causal representation of this controller that can be implemented in real life.

This process could apply to any problem satisfying the assumptions but the devil is in the detail. The main problem is that expressed in the last step, namely the conversion of the implicit relationship between u and e into a useable computation. In later chapters (see for example, Sects. 3.10 and 4.7), this idea will be linked, in the special cases of linear state space models, to *Two Point Boundary Value Problems* and the use of *Riccati equations*. More generally, the computations suffer from additional complexities and high dimensionality. Even the simplest cases present challenges. For example, let $\mathcal{U} = \mathbb{R}^q$ with inner product $\langle \hat{v}, v \rangle_{\mathcal{U}} = \hat{v}^T R v$ (where $R = R^T > 0$) and $\mathcal{Y} = \mathbb{R}^p$ with inner product $\langle \hat{w}, w \rangle_{\mathcal{Y}} = \hat{w}^T Q w$ (where $Q = Q^T > 0$). The operator G is a $p \times q$ real matrix with adjoint (Eq. (2.104)) defined by the $q \times p$ matrix $G^* = R^{-1}G^T Q$. Rather than using the implicit relationship, the direct computation of u can be undertaken using

$$u = u_0 + G^*(I + GG^*)^{-1}e_0 = u_0 + R^{-1}G^T Q(I_p + GR^{-1}G^T Q)^{-1}e_0 \quad (2.149)$$

This is a feasible approach to finding the solution and may work well in many cases but, if the dimensions p and q are large, the calculation of the inverse matrix could be challenging particularly if $I + GR^{-1}G^T Q$ is ill-conditioned. This example does have relevance to the topics in the text associated with the description of discrete time systems in supervector form (see Sect. 4.7). Simplifications are possible in this case

as the elements in G have a structural pattern that makes the inversion implicit in $G^*(I + GG^*)^{-1}$ equivalent to the solution of a Two Point Boundary Value Problem which is solved using Riccati equation and associated simulation methods.

2.7 Further Discussion and Bibliography

The chapter has reviewed material that plays a role in the development of the techniques used in this text. Matrices form the computational core of the algorithms and, although many readers will be familiar with basic matrix algebra, an understanding of the structures and the analysis tools available for matrix methodologies is helpful as is an understanding of the way that matrices are useful in the interpretation of high (but finite) dimensional problems using simple geometrical insights generated from the familiar three-dimensional world. There are many texts that cover the material required ranging from undergraduate engineering texts such as [60, 105, 106] to essentially mathematical texts that approach the topics using both algebraic concepts and the ideas of finite dimensional vector spaces [45, 46, 53] within which matrices are representations of operators using a specified basis set. Many teaching texts on control theory and control engineering also have a summary of the necessary material [4, 39, 43, 63, 71, 81]. The material is essentially the same but differing perspectives and different levels of abstraction are used. It is useful to note that there are links between matrices and transfer function descriptions using Toeplitz matrices, details of which can be found in [51]. For the purposes of this text, an understanding of, and fluency with, the algebraic structures and analysis tools will help the reader to “see through the symbols” and concentrate more usefully on the form and meaning of the system properties used and the nature of the algorithms described. An understanding of the algebraic and computational aspects of matrix theory will form the basis for any computational software required for the exploitation of the material and also in ensuring that data formats fit the necessary matrix structures.

Relevant techniques from functional analysis are also summarized in the chapter. This material will be less familiar to many readers but, in its simplest form, it can be regarded as a generalization of matrix theory to cover a wider range of problems. In particular, matrices are replaced by operators between, possibly infinite dimensional, signal spaces and the geometry of the three dimensional world is generalized to higher, possibly infinite, dimensions. The underpinning ideas of vector spaces endowed with norms to measure signal magnitude and the notion of bounded linear operators between such spaces mirror the familiar notion of a system as a device that maps input signals into output signals and the measurement of signal magnitude using measures such as least square values or maximum magnitudes. Although much of this work can be viewed at the algebraic level as being very similar to matrix (or even transfer function) methodologies, the technical details associated with the ideas do contain many subtle issues that need to be considered at the mathematical level. These take many forms but, perhaps the most important are those of existence of solutions to defined problems, the convergence of infinite sequences, the

introduction of the notion of adjoint operators and their properties, convexity and the idea of projection onto convex sets. There are many texts that provide the mathematical background for these topics including general texts on analysis [101] and functional analysis such as [12, 31, 52, 107] and more specialist texts on Hilbert space theory such as [54], operator theory [32, 33] and optimization algorithms using functional analytic methods [69]. A reference to Open Mapping and Closed Graph Theorems is found in [73] and an extensive analysis of projection methodologies is found in [36]. In the author's experience, the choice of text to suit the needs of a particular reader depends upon that reader and his or her background and preferred way of thinking.

Finally, the content of the text is mathematical in its chosen language and much of the supporting mathematics of Laplace and Z -transforms is used extensively in control engineering texts. Some of the more advanced tools needed can be found in texts on classical and multivariable control (see Sect. 1.5), mathematical systems theory [59, 104] and specialist texts [112] and papers [75] on geometric systems theory and decoupling theory [37, 47].