Chapter 3

Some Concepts form Linear Algebra

3.1 Introduction

This chapter contains a brief presentation of some basic results from linear algebra and, in particular, linear vector spaces and their role in optimization. In addition to a deeper understanding of the geometric properties of linear maps, this development has produced powerful computational tools for optimization. Among the most celebrated results in this setting is the classical projection theorem stating conditions for the existence, uniqueness and computation of solutions to minimum distance problems. This has numerous applications in optimization theory and optimal control. It may also invoked in an elegant reformulation of controllability and observability problems.

Most of the material in this chapter aims to address the problem of solving linear equations $Ax = b$ where x and b live in finite dimensional vector spaces like $\mathbf{R}^n$. While omitting some of the technical details, most of the development is done in an abstract setting that would allow an easy transition to the general case.

3.2 Fundamental Definitions and Properties

3.2.1 Vector Spaces and Subspaces

3.2.1 Definition: *A* **field** *is a set F with two operations '+' and '·' such that for any $a, b, c \in F$,:*

1. $a + b \in F$ *and* $a \cdot b \in F$.
2. $(F, +)$ *is an Abelian Group (with 0 denoting the identity).*
3. $(F - \{0\}, \cdot)$ *is an Abelian Group (with 1 denoting the identity).*
4. $\cdot$ *is distributive over addition (+), i.e.,* $a \cdot (b + c) = a \cdot b + a \cdot c$ *and* $(a + b) \cdot c = a \cdot c + b \cdot c$.

$(F, +)$ *is an* **Abelian Group** *if*

1. $(a + b) + c = a + (b + c)$ *(associativity).*
2. $a + b = b + a$ *(commutativity).*
3. *There exists an identity element 0 such that* $a + 0 = 0 + a = a$.
4. *For any* $a \in F$ *there exists an opposite element* $-a$ *such that* $a + (-a) = (-a) + a = 0$.

$\bigtriangledown\bigtriangledown$

Fields are the sets of scalars, e.g., $\mathbf{R}, \mathbf{C}$, rational functions. The operations '+', '·' are usually called addition and multiplication; often the '·' is omitted and ab simply means $a \cdot b$.

3.2.2 Definition: *A linear* **vector space** *(or, simply, vector space) over a field F, denoted by (X, F) is a set X together with two operations, addition $+ : X \times X \mapsto X$ and scalar multiplication $\cdot : F \times X \mapsto X$ such that for any $x, y \in X$ and $a, b \in F$:*

1. *$(X, +)$ is an Abelian Group, with 0_X denoting the zero vector.*
2. *$(a + b)x = ax + bx$ (distributivity).*
3. *$a(x + y) = ax + ay$ (distributivity).*
4. *$(ab)x = a(bx)$ (associativity).*
5. *$1x = x$, $0x = 0_X$.*

$\nabla\nabla$

When clear from the context, 0 is used to denote 0_X. Also, when clear from the context, the vector space can be denoted simply by X.

Examples of Vector Spaces include $(\mathbf{R}^n, \mathbf{R})$, $(\mathbf{C}^n, \mathbf{C})$. Other, more complicated, examples are $(l_\infty, \mathbf{R})$ (bounded sequences), $(C^0_{[a,b]}, \mathbf{R})$ (continuous functions on an interval $[a, b]$) etc.

3.2.3 Definition: *Let (X, F) be a vector space over the field F and $W \subset X$. (W, F) is a* **subspace** *of (X, F) if (W, F) is itself a vector space. Equivalently, $aw_1 + bw_2 \in W$, $\forall a, b \in F$, $\forall w_1, w_2 \in W$.* $\nabla\nabla$

3.2.4 Definition: *An* **affine space** *or* **variety** *is a set $V = W + x_0$ where W is a subspace of a vector space X and x_0 is a fixed element of X. The notation $W + x_0$ means all vectors that can be obtained as $w + x_0$ where $w \in W$.* $\nabla\nabla$

Examples of subspaces and varieties: Let $X = \mathbf{R}^2$. Then

$$W = \left\{ \begin{pmatrix} x_1 \\ 5x_1 \end{pmatrix} ; x_1 \in \mathbf{R} \right\}$$

is a subspace of X and

$$V = \left\{ \begin{pmatrix} x_1 \\ 5x_1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} ; x_1 \in \mathbf{R} \right\}$$

is a variety of X.

Let X be the vector space of continuous functions on $[0, 1]$. Then a subspace of X is the set of continuously differentiable functions on $[0, 1]$.

3.2.5 Definition: *Let X be a vector space and R, S subspaces of X. Then the following sets are subspaces of X*

$$R + S = \{r + s | r \in R, s \in S\}$$

$$R \cap S = \{v | v \in R \text{ and } v \in S\}$$

If $R \cap S = \{0\}$ then the set $J = R + S$ is called the **direct sum** *of R and S, denoted by $R \oplus S$. In such a case, for any $x \in J$ there exist unique $r \in R$, $s \in S$ such that $x = r + s$; the sets R, S are called complements with respect to J.* $\nabla\nabla$

Notice that the set $R \cup S$ is not necessarily a subspace of X.

3.2.6 Definition: *A set $\{x_1, x_2, \ldots x_n\}$ of elements of a vector space X is* **linearly independent** *if*

$$\sum_{i=1}^{n} a_i x_i = 0 \Rightarrow a_i = 0,\ i = 1, \ldots, n$$

where $a_i \in F$. $\nabla\nabla$

3.2.7 Definition: *Let $\{x_i\} = \{x_1, x_2, \ldots x_n\}$ be a set of elements of a vector space X (linearly independent or not). Then the* **span** *of this set is defined as*

$$span\{x_i\} = \{v | v = \sum_i a_i x_i,\ a_i \in F\}$$

$\bigtriangledown\bigtriangledown$

A span of a set of vectors is a subspace.

3.2.8 Definition: *A set $\{x_1, x_2, \ldots x_n\}$ of elements of a vector space X is a* **basis** *of X if it is a linearly independent set and its span is equal to X. The* **dimension** *of X is the number of elements in a basis and is independent of the particular basis used.* $\bigtriangledown\bigtriangledown$

$\mathbf{R}^3$ is a finite dimensional space of dimension 3. l_∞, C^0 are infinite dimensional spaces.
For sums of subspaces, $dim(R + S) = dim(R) + dim(S) - dim(R \cap S)$.

3.2.9 Definition: *Let (X, F), (Y, F) be vector spaces. A map (or transformation) $A : X \mapsto Y$ is* **linear** *if*

$$A(ax_1 + bx_2) = aA(x_1) + bA(x_2),\ \forall a, b \in F,\ \forall x_1, x_2 \in X$$

X is called the **domain** *of A and Y is the* **co-domain** *of A.* $\bigtriangledown\bigtriangledown$

So far, we have treated vectors as abstract objects. However, in view of the definition of a basis (its existence and construction is somewhat involved) we can associate the vector to its representation in terms of a given basis. For example, let $\{x_i\}$ be a basis for the vector space X and let $x \in X$ be represented by $x = \sum_i a_i x_i$. Then, if there is no ambiguity regarding the basis used, we can use the vector $[a_1, a_2, \ldots]$ to denote x (row or column, depending on the context). In this context, it follows that a linear transformation admits a matrix description, say A, where the i-th column of A are the coefficients of $A(x_i)$ with respect to the basis $\{y_i\}$. In principle, the linear transformation and its matrix representation should be assigned different symbols but this is not done here to avoid proliferation of notation.

3.2.10 Definition: *Let $A : X \mapsto Y$ be a linear map between two vector spaces X and Y. The* **null** *space of A is defined as*

$$\mathcal{N}(A) = \{x \in X : Ax = 0\}$$

The **range** *space of A is defined as*

$$\mathcal{R}(A) = \{y \in Y : y = Ax \text{ for some } x \in X\}$$

The **rank** *of A is defined as the dimension of $\mathcal{R}(A)$: $r(A) = dim\mathcal{R}(A)$.* $\bigtriangledown\bigtriangledown$

It is easy to show that both the null and range spaces are subspaces. Furthermore, notice that $\mathcal{R}(A) = span\{a_i\}$ where $\{a_i\}$ denotes the columns of A. An interesting property connecting dimensions is $dim\mathcal{N}(A) + dim\mathcal{R}(A) = \dim(X)$.

3.2.2 Inner products and norms

The following group of definitions and properties pertains to certain scalar-valued functions (functionals) of vectors. They serve to define geometrical properties in a vector space, such as orthogonality, alignment and distance.

3.2.11 Definition: *An* inner product *is a scalar-valued function of two elements of a vector space X, $< \cdot, \cdot >: X \times X \mapsto \mathbf{C}$ (or $\mathbf{R}$) with the following properties:*

1. *$\forall x \in X$, $< x, x > \in \mathbf{R}$, $< x, x > \geq 0$ and $< x, x >= 0 \Leftrightarrow x = 0$.*
2. *$\forall x, y \in X$, $< x, y >= \overline{< y, x >}$.*
3. *$\forall x, y \in X,\ a \in F$, $< x, ay >= a < x, y >$.*

4. $\forall x, y, z \in X$, $< x, y + x >=< x, y > + < x, z >$.

where $\bar{\cdot}$ denotes the complex conjugate. A vector space equipped with an inner product is called an inner product space (or pre-Hilbert) and is denoted by $(X, <, >)$. ▽▽

Common examples of inner poducts are $x^\top y$ in $\mathbf{R}^n$, $\bar{x}^\top y$ in $\mathbf{C}^n$, (other notations: $x^H y$, $x^* y$), $\int x^\top(t)y(t)\,dt$ in $(C^0, \mathbf{R})$ etc.

3.2.12 Lemma: *(Cauchy-Schwarz inequality) In an inner product space, $| < x, y > | \leq \sqrt{< x, x >}\sqrt{< y, y >}$, $\forall x, y \in X$. Equality holds iff $x = ay$ or $y = 0$.* ▽▽

Proof: For $y \neq 0$ and for any $a \in F$,

$$0 \leq < x - ay, x - ay >=< x, x > -a < x, y > -\bar{a} < y, x > +|a|^2 < y, y >$$

Taking $a = \frac{<x,y>}{<y,y>}$ the inequality follows. □□

3.2.13 Definition: *A* **norm** *on a vector space X is a functional $\|\cdot\| : X \mapsto \mathbf{R}$ with the following properties ($\forall x, y \in X, \forall a \in F$):*

1. *$\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$ (nonnegativity).*
2. *$\|ax\| = |a|\|x\|$ (homogeneity).*
3. *$\|y + x\| \leq \|x\| + \|y\|$ (triangle inequality).*

A vector space equipped with a norm is called a normed (vector) space and is denoted by $(X, \|\cdot\|)$. ▽▽

3.2.14 Corollary: *In an inner product space, $\|x\| \triangleq \sqrt{< x, x >}$ is a norm. It is referred to as the norm induced by the inner product and is often denoted by $\|\cdot\|_2$.* ▽▽

Notice that norms may be defined without the notion of an inner product. Norms themselves provide a sense of distance, i.e., $\|x\|$ can be interpreted as the distance of x from 0. As a distance the norm also induces a topology for the vector space (i.e., open sets, neighborhoods). However, a topology and a distance may be defined without the use of a norm.

3.2.15 Definition: *Two elements x, y of an inner product space are said to be* **orthogonal** *if $< x, y >= 0$. A set of elements $\{x_i\}$ is an* **orthogonal set** *if $< x_i, x_j >= 0$, $\forall i \neq j$; it is said to be* **orthonormal** *if, in addition, $< x_i, x_i >= 1$, $\forall i$.*
Two elements x, y of an inner product space are said to be **aligned** *if $< x, y >= \|x\|\|y\|$.* ▽▽

While orthogonality requires the notion of an inner product, alignment can be defined in more general terms between a normed space X and its "dual," X^*, the space of all bounded linear functionals on X. These concepts are further explored in Functional Analysis.

An important result that connects orthogonality and norms induced by an inner product is the Pythagorean theorem:

3.2.16 Theorem: *In an inner product space, let x, y be two orthogonal vectors. Then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.* ▽▽

Proof: $\|x + y\|^2 =< x + y, x + y >=< x, x > + < y, y > + < x, y > + < y, x >$. Observing that the last two terms are zero, the result follows. □□

3.2.17 Definition: *In an inner product space X, the orthogonal complement of a set $S \subseteq X$ is defined by*

$$S^\perp = \{x \in X | < x, s >= 0,\ \forall s \in S\}$$

▽▽

The orthogonal complement of a set is always a subspace. Furthermore, the following properties hold:

$$(R + S)^\perp = R^\perp \cap S^\perp \quad (R \cap S)^\perp = R^\perp + S^\perp$$

To show these properties, pick an element of the set in the left-hand side and show that it belongs to the right-hand side and vice-versa.

An important property of orthogonal complements is that they are closed, i.e., they contain all their limit points (or boundary or closure points). [1] In general, we have $S \subseteq S^{\perp\perp}$. Equality holds iff S is a closed subspace.

3.2.3 Banach and Hilbert Spaces

The following definitions are introduced to establish the terminology. While they are associated with important results, their deeper understanding is not required in the present development.

- A sequence $\{x_n\}$ in a normed space is said to be **Cauchy** if $\|x_n - x_m\| \to 0$ as $n, m \to \infty$.
- A normed vector space X is **complete** if every Cauchy sequence of elements in X has a limit in X. Such a space is called **Banach**.
- A complete pre-Hilbert space is called a **Hilbert** space. It is also a Banach space with norm induced by the inner product.

Some important consequences of these definitions are:

- In a Banach space a subset is complete iff it is closed. (Not necessarily true in general normed spaces.)
- In a normed vector space, any finite dimensional subspace is complete. (Finite dimensionality alleviates many technical problems!)
- Let R be a closed subspace of a Hilbert space X. Then $X = R \oplus R^{\perp}$.

3.2.4 Linear Maps

A linear map between two Hilbert spaces (e.g., $\mathbf{R}^n, \mathbf{C}^n$) induces a certain geometric structure of these spaces. That is, it allows the decomposition of the two spaces in subspaces related to the properties of the map.

3.2.18 Definition: *Let $A : X \mapsto Y$ be a linear map between Hilbert spaces. The adjoint map is defined as a map $A^* : Y \mapsto X$ such that*

$$< y, Ax >_Y = < A^* y, x >_X$$

for all $x \in X$, $y \in Y$. $\bigtriangledown\bigtriangledown$

The adjoint map is also linear. When X, Y are $\mathbf{R}^n, \mathbf{R}^m$ then $A^* = A^\top$, i.e., it is associated with the usual transpose of the matrix. For the complex case, the adjoint is the complex-conjugate transpose (A^H or A^*).

3.2.19 Lemma: *(Four fundamental subspaces, finite dimensional case) For a linear map $A : X \mapsto Y$, $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$, $\mathcal{N}(A)^\perp = \mathcal{R}(A^*)$.* $\bigtriangledown\bigtriangledown$

Proof: For the first equality, let $y \in \mathcal{R}(A)^\perp$. Then for all $x \in X$, $< y, Ax >= 0$. Hence, $< A^*y, x >= 0$ which implies that $A^*y = 0$ or $y \in \mathcal{N}(A^*)$. Hence, $\mathcal{R}(A)^\perp \subseteq \mathcal{N}(A^*)$.

Now, let $y \in \mathcal{N}(A^*)$ implying that $< A^*y, x >= 0$, for all $x \in X$. This implies that $< y, Ax >= 0$ and, therefore, $y \in \mathcal{R}(A)^\perp$. Hence, in view of the first part, $\mathcal{N}(A^*) \subseteq \mathcal{R}(A)^\perp$. Therefore $\mathcal{N}(A^*) = \mathcal{R}(A)^\perp$.

For the second equality, let $x \in \mathcal{N}(A)$. So, for all $y \in Y$, $< y, Ax >= 0 =< A^*y, x >$. Hence, $x \in \mathcal{R}(A^*)^\perp$ and $\mathcal{N}(A) \subseteq \mathcal{R}(A^*)^\perp$

Next, let $x \in \mathcal{R}(A^*)^\perp$. Then, for all $y \in Y$, $< x, A^*y >= 0 =< Ax, y >$. So, $Ax = 0$ and $x \in \mathcal{N}(A)$. Hence, $\mathcal{R}(A^*)^\perp \subseteq \mathcal{N}(A)$. The two imply that $\mathcal{R}(A^*)^\perp = \mathcal{N}(A)$.

Taking orthogonal complements $\mathcal{N}(A)^\perp = \mathcal{R}(A^*)^{\perp\perp}$. In general the right-hand side is the closure of $\mathcal{R}(A^*)$ which is equal to $\mathcal{R}(A^*)$ itself if it is a closed subspace, e.g., when it is finite dimensional. $\square\square$

[1] $x \in X$ is a closure point of a set P if given any $\epsilon > 0$ there is $p \in P : \|x - p\| < \epsilon$. On the other hand, $P \subset X$ is open if for any $p \in P$ there exists an $\epsilon > 0$ such that all x satisfying $\|x - p\| < \epsilon$ are also members of P.

3.2.20 Definition: *A linear map $A : X \mapsto Y$ is said to be*

- **onto**, *if $\mathcal{R}(A) = Y$.*
- **one-to-one** *(1-1) if $\mathcal{N}(A) = \{0\}$ ($Ax_1 = Ax_2 \Rightarrow x_1 = x_2$)*
- **invertible** *if it is 1-1 and onto.*

$\nabla\nabla$

3.2.21 Properties: *For a linear map $A : X \mapsto Y$,*

- *A is onto iff $r(A) = dim(Y)$ (full row rank).*
- *A is 1-1 iff $r(A) = dim(X)$ (full column rank).*
- *A is invertible iff $r(A) = dim(X) = dim(Y)$.*
- *$r(A) = r(A^*)$*
- *if A is onto then there exists a right inverse of A, say A^{-R} such that $AA^{-R} = I$.*
- *if A is 1-1 then there exists a left inverse of A, say A^{-L} such that $A^{-L}A = I$.*
- *if A is invertible then $A^{-R} = A^{-L} = A^{-1}$.*

$\nabla\nabla$

For example, $A = (1, 2)$ is onto, with a right inverse $(3, -1)^\top$. On the other hand $A = (1, 2)^\top$ is 1-1 with a left inverse $(3, -1)$. But,

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

is neither.

In relation to norms, the gain or induced norm of a linear operators is defined as follows:

3.2.22 Definition: *For a linear map $A : X \mapsto Y$,*

$$\|A\|_{YX} = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Ax\|_Y$$

$\nabla\nabla$

Often, the notation is simplified to $\|A\|_{ip}$ or $\|A\|_p$ when the p-norm is used for both X and Y vector spaces. While matrix norms are not always easily computable, it follows that the induced 2-norm (i.e., with vector norms induced by the inner product) $\|A\|_2 = \bar{\sigma}(A) = \lambda_{max}^{1/2}(A^*A) = \lambda_{max}^{1/2}(AA^*)$ (the maximum singular value of A). The geometric interpretation of the maximum singular value is the radius of the smallest unit ball in Y that contains the map of the unit ball in X.

It can be shown that induced norms are also norms for matrices (i.e., making the matrix vector space, a normed space). However, the converse is not always true. In particular, an important property of induced norms is that they are **consistent**, that is $\|Ax\| \leq \|A\|\|x\|$ and $\|AB\| \leq \|A\|\|B\|$. This, together with the triangle inequality of norms offer the basic tools in various bounding procedures.

An example of a matrix norm that is consistent but not induced is the so-called Frobenius norm $\|A\|_F = (\sum_i \sum_j |a_{i,j}|^2)^{1/2}$. This norm is consistent with the usual Euclidean vector norm or 2-norm, induced by the inner product. On the other hand, the norm $\max_{i,j} |a_{i,j}|$ satisfies all the norm axioms but is not consistent.

The proof of this statement is by counterexample: Verify that the submultiplicative property does not hold for the case

$$A = B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

3.2.23 Other Matrix Properties:

- *A matrix A is called symmetric if $A^\top = A$ (Hermitian for the complex case where the transpose is replaced by complex conjugate transpose). Symmetric matrices have real eigenvalues and posess a complete orthonormal set of eigenvectors. They are diagonalizable by an orthogonal matrix.*
- *A matrix A is called orthogonal if $A^\top = A^{-1}$ (Unitary in the complex case). The eigenvalues of an orthogonal matrix have magnitude one. Orthogonal matrices define a rotations transformation and they are norm-preserving ($\|Ax\|^2 = x^\top A^\top Ax = \|x\|^2$).*
- *A matrix is skew-symmetric if $A^\top = -A$. For real vectors, $x^\top Ax = 0$.*
- *A matrix is normal if it commutes with its transpose ($AA^\top = A^\top A$). Normal matrices are all the matrices that can be diagonalized by an orthogonal matrix.*
- *A real matrix A is called* **positive definite** *if it is symmetric and $x^\top Ax > 0$ for all $x \neq 0$. When $x^\top Ax \geq 0$, $\forall x$ the matrix is called positive semi-definite. (Notation: $A = A^\top > 0$ and $A = A^\top \geq 0$ respectively.)*
- *A symmetric matrix is positive definite (semi-definite) iff all its eigenvalues are positive (nonnegative).*
- *For a symmetric matrix A the following inequality holds*

  $$\lambda_{min}(A)x^\top x \leq x^\top Ax \leq \lambda_{max}(A)x^\top x$$

  where $\lambda_{min}, \lambda_{max}$ denote the minimum and maximum eigenvalues respectively.

$\nabla\nabla$

3.2.5 Projections

Consider a decomposition of a vector space as $X = R \oplus S$. Then every element in X can be written uniquely as $x = r + s$ where $r \in R$, $s \in S$. A projection $\mathcal{P}_{R,S}$ is defined as a map $X \mapsto R$ by $\mathcal{P}_{R,S}x = r$. In general, these are called oblique projections (on R along S). Note that both subspaces R and S must be specified since the complement of a subspace is not unique. However, when the orthogonal complement is used ($S = R^\perp$) then the projection is called orthogonal and it is denoted by $\mathcal{P}_R$.

A projection is a linear operation and it is identity on the subspace onto which it projects. This observation gives rise to the following properties.

3.2.24 Properties: *For a linear map associated with a matrix P,*

- *P is a projection iff it is idempotent, i.e., $P^2 = P$.*
- *P is a projection iff $I - P$ is a projection.*
- *P is an orthogonal projection iff $P^2 = P = P^*$.*

$\nabla\nabla$

Typical examples of projections are:

- Assume A^*A is invertible (A is full column rank). Then $P = A(A^*A)^{-1}A^*$ is an orthogonal projection onto $\mathcal{R}(A)$.
- Assume AA^* is invertible (A is full row rank). Then $P = A^*(AA^*)^{-1}A$ is an orthogonal projection onto $\mathcal{R}(A^*) = \mathcal{N}(A)^\perp$.

3.2.6 Singular Value Decomposition

The singular value decomposition (SVD) is a canonical decomposition of a matrix, similar to the eigenvalue-eigenvector decomposition. In contrast to the latter, however, it exists for all matrices and it can be computed with numerically robust algorithms. It plays a significant role in the solution of linear systems of equations as it provides a numerically reliable method to compute a basis and projection matrices for the four fundamental subspaces.

3.2.25 Theorem: *Let $A \in \mathbf{R}^{m\times n}$ be a matrix of rank q. Then there exist orthogonal matrices $U \in \mathbf{R}^{m\times m}$, $V \in \mathbf{R}^{n\times n}$ such that $A = USV^\top$.*

S is a special matrix with block diagonal structure:

$$S = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$$

where $\Sigma = diag(\sigma_1, \ldots, \sigma_q)$ is a diagonal $q \times q$ matix. Its entries are positive and called the singular values of A. The rest of the matrix S is zero blocks of appropriate dimensions. $\bigtriangledown\bigtriangledown$

- The SVD of a matrix is not unique.
- The singular values of A are also the square-roots of the non-zero eigenvalues of $A^\top A$ and $AA^\top$.
- Even though the SVD can be defined in terms of eigenvalue-eigenvector decomposition of $A^\top A$ and $AA^\top$, these do not provide good numerical methods for its computation.
- Using the natural partitioning of S, the SVD of a matrix A can be written as

$$A = (U_1, U_2) \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^\top \\ V_2^\top \end{pmatrix}$$

  where U_1 is an $m \times q$ matrix and $V_1^\top$ is a $q \times n$ matrix. Then, $A = U_1 \Sigma V_1^\top$.
- $span(U_1) = \mathcal{R}(A)$, $span(U_2) = \mathcal{R}(A)^\perp$.
- $span(V_1) = \mathcal{N}(A)^\perp$, $span(V_2) = \mathcal{N}(A)$.
- $U_1 U_1^\top$ defines a projection on $\mathcal{R}(A)$.
- $U_2 U_2^\top = I - U_1 U_1^\top$ defines a projection on $\mathcal{R}(A)^\perp$.
- $V_1 V_1^\top = I - V_2 V_2^\top$ defines a projection on $\mathcal{N}(A)^\perp$.
- $V_2 V_2^\top$ defines a projection on $\mathcal{N}(A)$.

3.2.7 Optimization in Vector Spaces

One of the most powerful and important optimization principles is the projection theorem. Much of its power and simplicity hinges on the availability of an inner product and the induced special way distances are measured. In its simplest form, it states that the shortest distance from a point to a line is given by the perpendicular from the point to the line. Easily generalized to n-dimensional and infinite dimensional vector spaces, it provides a general tool to assess existence and uniqueness of solutions as well as to guide the development of computational procedures.

The optimization problem considered is this: Given a vector x in a pre-Hilbert space X and a subspace M in X, find the vector $m \in M$ closest to x in the sense that it minimizes $\|x - m\|$.

The fundamental questions on the uniqueness and characterization of the solution are addressed by the following preliminary theorem.

3.2.26 Theorem: *Let X be a pre-Hilbert space, M a subspace of X and $x \in X$ an arbitrary vector. If* *there is a vector $m_0 \in M$ such that $\|x-m_0\| \le \|x-m\|$, $\forall m \in M$, then m_0 is unique. A necessary and sufficient condition that $m_0 \in M$ is the unique minimizing vector in M is that the error vector $x - m_0$ is orthogonal to M.* $\bigtriangledown\bigtriangledown$

Proof: We show first that if m_0 is a minimizing vector, then $x - m_0$ is orthogonal to M. Suppose to the contrary, that there exists $m \in M$ not orthogonal to $x - m_0$. Without loss of generality, we can assume $\|m\| = 1$ and $< x - m_0, m >= \delta \neq 0$. Define $m_1 = m_0 + \delta m \in M$. Then

$$\begin{aligned} \|x-m_1\|^2 &= \|x-m_0-\delta m\|^2 \\ &= \|x-m_0\|^2 - < x-m_0, \delta m > - < \delta m, x-m_0 > + |\delta|^2 \\ &= \|x-m_0\|^2 - |\delta|^2 < \|x-m_0\|^2 \end{aligned}$$

Thus, m_0 cannot be a minimizer.

Next, we show that if the orthogonality condition holds, then m_0 is unique. For any $m \in M$ the Pythagorean theorem gives

$$\|x-m\|^2 = \|x-m_0+m_0-m\|^2 = \|x-m_0\|^2 + \|m_0-m\|^2$$

Thus, if $m \neq m_0$, then $\|x-m\| > \|x-m_0\|$. $\square\square$

The remaining question of existence requires a strengthening of the hypotheses. In particular, the problem lies in the closedness of the sets which would ensure the existence of a minimizer. Loosely speaking, without the closedness condition, the minimizer may be on the "boundary" of the set; then, a sequence $\{m_i\}$ may be constructed so that $\|x-m_i\|$ converges to the infimum of the distance $\|x-m\|$; but the sequence does not converge to an element in M.

3.2.27 Theorem: *(Classical Projection Theorem) Let H be a Hilbert space, M a closed subspace of H. Corresponding to any vector $x \in X$, there is a unique vector $m_0 \in M$ such that $\|x-m_0\| \le \|x-m\|$, $\forall m \in M$. A necessary and sufficient condition that $m_0 \in M$ is the unique minimizing vector in M is that the error vector $x-m_0$ is orthogonal to M.* $\bigtriangledown\bigtriangledown$

Proof: The uniqueness and orthogonality have been established by the previous theorem. To establish existence, define $\delta = \inf_{m\in M} \|x-m\| > 0$. (If the infimum is zero, then $x \in M$ and $m_0 = x$.) We wish to produce an $m_0 \in M$ for which $\|x-m_0\| = \delta$. For this purpose, let $\{m_i\}$ be a sequence of vectors in M such that $\|x-m_i\| \to \delta$. By the parallelogram law[2]

$$\|(m_j-x)+(x-m_i)\|^2 + \|(m_j-x)-(x-m_i)\|^2 = 2\|m_j-x\|^2 + 2\|x-m_i\|^2$$

Rearranging, we obtain

$$\|m_j-m_i\|^2 = 2\|m_j-x\|^2 + 2\|x-m_i\|^2 - 4\|x-(m_i+m_j)/2\|^2$$

For all i, j, the vector $(m_i+m_j)/2$ is in M, so $\|x-(m_i+m_j)/2\| \ge \delta$ and we obtain

$$\|m_j-m_i\|^2 = 2\|m_j-x\|^2 + 2\|x-m_i\|^2 - 4\delta^2$$

Since $\|x-m_i\| \to \delta$ as $i \to \infty$ we conclude that $\|m_i-m_j\| \to 0$ as $i, j \to \infty$. Therefore, $\{m_i\}$ is a Cauchy sequence and since M is a closed subspace of a complete space, it has a limit $m_0 \in M$. Finally, by the continuity of the norm, it follows that $\|x-m_0\| = \delta$. $\square\square$

Much of the usefulness of the Classical Projection Theorem comes with the availability of algorithms to perform the necessary projections. Its true value, however, is in the infinite dimensional case which would be difficult to handle with conventional arguments. It is an extremely valuable tool in approximation theory and Fourier series.

An important direct extension of the projection theorem is in the case of minimizing the distance to a convex set.[3]

[2]In an inner product space, $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$; proof by direct expansion of the norms in terms of the inner product.

[3]A set K in a vector space is **convex** if, given any $x, y \in K$, then all points on the line $ax + (1-a)y$ with $0 \le a \le 1$ are in K.

3.3 Solution of Linear Equations

The topic of this section is the solution of linear equations of the form $Ax = b$. Throughout the presentation A is a real matrix $m \times n$ and b is a real vector $m \times 1$. This fundamental problem arises very frequently in modeling and approximation. A slightly more general version occurs when the field (scalars) is the set of complex numbers. In such a case, the results are still valid with some minor changes (complex conjugate transpose instead of a transpose). The more interesting generalization is to the infinite dimensional case through the use of linear operators and their adjoints. While still similar in form, the results involve a fair amount of technical issues.

The existence and uniqueness of solutions for the basic problem can be easily addressed in the framework of the previous section. That is, for the system of equations $Ax = b$,

1. there exists a solution iff $b \in \mathcal{R}(A)$;
2. there exists a solution for all $b \in \mathbf{R}^m$ iff $\mathcal{R}(A) = \mathbf{R}^m$ (A is onto);
3. the solution is unique iff $\mathcal{N}(A) = \{0\}$ (A is 1-1);
4. there exists a unique solution for any $b \in \mathbf{R}^m$ iff A is invertible.
5. For the special case $Ax = 0$, there exist nontrivial solutions iff $r(A) < n$ (A is not 1-1).

Certainly, the computation of the solution, whenever it exists, is an important problem. However, in practice it is often the case that solutions do not exist (e.g., approximation problems). To address this issue, it is useful to embed the basic problem into a more general class of optimization problems, that is:

$$\min_{x \in \mathbf{R}^n} \|Ax - b\|$$

This is the celebrated **Least-Squares** problem, with its name justified by the selection of the norm (Euclidean or 2-norm). Clearly, when solutions exist, this minimization problem is equivalent to solving $Ax = b$. In view of the Classical Projection Theorem, this problem can be recast as

$$\min_{y \in \mathcal{R}(A)} \|y - b\|$$

which has a unique solution. Now, there exists a solution for $Ax = y$ but it may not be unique. Depending on the problem at hand, we may be interested in characterizing all solutions or selecting one that has certain additional properties. For the latter, it is convenient (and often possible) to express the additional desired properties as the solution of an optimization problem. That is, suppose that S is the set of solutions to $Ax = y$. Then we seek a vector x_* solving the problem:

$$\min_{x \in S} \|x - x_0\|$$

where x_0 is a given vector. When $x_0 = 0$, this problem is referred to as the **Minimum Norm** problem.

3.3.1 Characterization of all solutions

The problem here is to characterize all solutions to $Ax = b$, assuming that there exists at least one solution.

3.3.1 Proposition: *Suppose that x_0 is a solution to $Ax = b$. Then a vector x is a solution iff it belongs to the variety $x_0 + \mathcal{N}(A)$.* ▽▽

Proof: (if): Suppose that $x \in x_0 + \mathcal{N}(A)$, i.e., $x = x_0 + x_n$ with $x_n \in \mathcal{N}(A)$. Then $Ax = A(x_0 + x_n) = Ax_0 + Ax_n = Ax_0 = b$.

(only if): Suppose $Ax = b$. Then $Ax - Ax_0 = b - b = 0$ so $A(x - x_0) = 0$ implying that $x - x_0 = x_n \in \mathcal{N}(A)$. □□

This simple argument provides a very important tool. Its usefulness can be seen, for example, in optimization. Suppose that we want to minimize a functional $f(x)$ subject to the constraint $Ax = b$. Let N be a matrix whose columns are a basis for $\mathcal{N}(A)$. Then, given a particular solution x_0, the constraint becomes $x = x_0 + Ny$ where y is a $dim\mathcal{N}(A)$-dimensional vector. Substituting, we get the equivalent unconstrained optimization problem to minimize $f(x_0 + Ny) = g(y)$ with respect to y. Furthermore, since the relationship between x and y is affine, the form of f is left essentially unaffected (e.g., quadratic forms in x remain quadratic in y).

3.3.2 The Minimum Norm problem

Let A be an $m \times n$ real matrix with $m < n$ and A being onto (full row rank). We are interested in computing the minimum norm solution to $Ax = b$.

Since A is onto, then a solution exists for any $b \in \mathbf{R}^m$. In view of the parametrization of all solutions, we are interested in solving the optimization problem

$$\min_{m \in \mathcal{N}(A)} \|x_0 - m\|$$

where x_0 is a particular solution. (The desired x would then be $x_0 - m$.) From the Classical Projection Theorem we have that the solution to this problem, say m_0, exists, is unique and $x_0 - m_0$ should be orthogonal to $\mathcal{N}(A)$. In other words, $x_0 - m_0$ is the projection of x_0 onto $\mathcal{N}(A)^\perp$. This projection can be computed easily by

$$\mathcal{P}_{\mathcal{N}(A)^\perp}(x_0) = A^\top (AA^\top)^{-1} A x_0$$

Notice that A being onto ensures the invertibility of $AA^\top$ (the proof of that is not trivial but it is straightforward). Finally, since x_0 is a solution, we have that $Ax_0 = b$ from which we get an expression that is independent of x_0.

Thus, the solution to the minimum norm problem is

$$\begin{aligned} x_{MN} &= \arg\min_{Ax=b} \|x\| = A^\top (AA^\top)^{-1} b \\ \|x_{MN}\| &= \min_{Ax=b} \|x\| = (b^\top (AA^\top)^{-1} b)^{1/2} \end{aligned}$$

Observe that even though we began by considering a particular solution, the final answer involves only the problem data (A, b).

3.3.3 The Least Squares problem

Let A be an $m \times n$ real matrix with $n < m$ and A being 1-1 (full column rank). We are interested in computing the least squares solution to $Ax = b$.

Here, a solution may not exist in a strict sense since, in general, $\mathcal{R}(A) \subset \mathbf{R}^m$ and an arbitrary $b \in \mathbf{R}^m$ may not be in $\mathcal{R}(A)$. For this reason, we seek to minimize the difference between the left- and right-hand side. To solve this problem, let $y = Ax$ so $y \in \mathcal{R}(A)$, whereby the least-squares optimization problem becomes

$$\min_{y \in \mathcal{R}(A)} \|y - b\|$$

From the Classical Projection Theorem we have that there exists a unique minimizer y_0 and such that $y_0 - b$ is orthogonal to $\mathcal{R}(A)$. [4]

Next, in order to compute y_0 and the corresponding x_0, we observe that y_0 is the projection of b onto $\mathcal{R}(A)$. Thus,

$$y_0 = \mathcal{P}_{\mathcal{R}(A)}(b) = A(A^\top A)^{-1} A^\top b$$

and since $Ax_0 = y_0$ it follows that $x_0 = (A^\top A)^{-1} A^\top b$; this solution is also unique since A is 1-1.

[4] In general, the uniqueness of y_0 does not imply uniqueness of an x_0 such that $Ax_0 = y_0$. This is true in our special case since A was assumed to be 1-1.

In conclusion, the least squares solution can be computed by the following simple formula:

$$\begin{aligned} x_{LS} &= \arg\min_{x\in\mathbf{R}^n} \|Ax-b\| = (A^\top A)^{-1}A^\top b \\ \|Ax_{LS}-b\| &= \min_{x\in\mathbf{R}^n} \|Ax-b\| = [b^\top(I-A(A^\top A)^{-1}A^\top)b]^{1/2} \end{aligned}$$

where for the latter, we used the fact $\mathcal{P}_{\mathcal{R}(A)^\perp} = I - \mathcal{P}_{\mathcal{R}(A)}$.

An interesting observation in this problem is that regardless of A, $\mathcal{R}(A^\top) = \mathcal{R}(A^\top A)$ as long as $\mathcal{R}(A)$ is closed. To show this, decompose the domain of $A^\top$ to $\mathcal{N}(A^\top) \oplus \mathcal{N}(A^\top)^\perp$ and use $\mathcal{N}(A^\top)^\perp = \mathcal{R}(A)$, provided that the latter is closed. This implies that solutions to $A^\top Ax = A^\top b$ exist. It can be shown that these solutions are also minimizers for the origninal least-squares problem and vice-versa. The set of equations $A^\top Ax = A^\top b$ is referred to as the **normal equations**.

3.3.4 The general problem: Solution through SVD

In the previous discussion we have developed some simple expressions for the solution of special-case least squares and minimum norm problems. The assumptions were concerned with the nonsingularity of a certain matrix that may not be satisfied in the general case. A more critical issue is that with actual data these assumptions may be formally satisfied but the correpsonding matrix can be near-singular or badly-conditioned. This would increase the sensitivity of the solution to numerical round-off errors and/or perturbations in the data. To address these issues, a numerically more reliable solution can be obtained via the SVD of the matrix A.

Let A be an $m \times n$ matrix and consider the problem

$$\begin{aligned} \min_x &\quad \|x\| \\ \text{s.t.} &\quad \|Ax-b\| = \min_x \|Ax-b\| \end{aligned}$$

Using the SVD, write $A = USV^\top$, and define $w = V^\top x$. Since V is invertible, the above minimization problem is equivalent to

$$\min_w \|USw-b\| = \min_w \|U(Sw-U^\top b)\| = \min_w \|Sw-(U^\top b)\|$$

where the last equality follows from the fact that U is orthogonal (unitary) and, therefore, norm-preserving. Furthermore, notice that since V is also orthogonal (unitary), $\|w\| = \|x\|$ and any minimization of $\|w\|$ is equivalent to the minimization of $\|x\|$.

Next, let q be the rank of A and define the vector $c = U^\top b$. Bringing in the special structure of S, the minimization problem becomes

$$\min_w \left\| \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_q \\ w_{q+1} \\ \vdots \\ w_n \end{pmatrix} - \begin{pmatrix} c_1 \\ \vdots \\ c_q \\ c_{q+1} \\ \vdots \\ \vdots \\ c_m \end{pmatrix} \right\|$$

where Σ is a $q \times q$ matrix. Clearly, the least squares solution is

$$w_1 = c_1/\sigma_1,\ \dots,\ w_q = c_q\sigma_q,\ \{w_{q+1},\dots w_n\} \text{ arbitrary}$$

and the corresponding least squares error is $(|c_{q+1}|^2 + \cdots + |c_m|^2)^{1/2}$. Furthermore, the minimum norm least squares solution is obtaind by setting $w_{q+1} = \cdots = w_n = 0$.

Thus, translating the result back to the original coordinates, we have

$$x_{LSMN} = V \begin{pmatrix} \Sigma^{-1}[U^\top b]_1^q \\ 0 \\ \vdots \\ 0 \end{pmatrix} = [V]_1^q \left(\ \Sigma^{-1}[U^\top b]_1^q \ \right)$$

where $[U^\top b]_1^q$ denotes the first q elements of the vector $U^\top b$ and $[V]_1^q$ denotes the first q columns of V.

The practical usefulness of this approach is in the ability to set a threshold on the singular values of the matrix A that will be inverted. Thus the sensitivity of the solution is reduced at the expense of a slight increase in the error that can be decided a priori.

Finally, it should be mentioned that variants of this approach can be constructed that do not require the SVD of the whole matrix, something that can be important in large problems (i.e., large m).

3.3.5 Weighted Least Squares Minimum Norm Problems

We conclude this chapter with a few comments on weighted versions of the basic Least Squares Minimum Norm problem. Such a case may arise in applications where, for example, certain directions in the approximation should be emphasized. Similarly, in the minimum norm problem there may be directions (components of x) that is desirable to keep small.

Both of these cases can be handled by using **weighted norms**, defined as follows:

$$\|x\|_Q = \sqrt{x^\top Q x}$$

For the above expression to constitute a norm Q must be a positive definite matrix ($Q = Q^\top > 0$). Similarly, one can define Q-orthogonality via the inner product $x^\top Q y$ etc. Degenerate cases where $Q \geq 0$ can also occur but they require somewhat different techniques and are not considered here. A standard way to handle such problems is through a coordinate transformation via the square-root of Q. For a positive (semi) definite matrix the square root exists and is also positive (semi) definite and satisfies $\sqrt{Q}^\top \sqrt{Q} = Q$. Now, define the new coordinates as $\bar{x} = \sqrt{Q}x$, where $\|\bar{x}\| = \|x\|_Q$. In the new coordinates, the weighted norm becomes the standard Euclidean norm and orthogonality works in the usual sense.

As an application of this approach consider the weighted LSMN problem

$$\begin{array}{ll} \min_x & \|x\|_R \\ \text{s.t.} & \|Ax - b\|_Q = \min\limits_x \|Ax - b\|_Q \end{array}$$

Define the new coordinates for the domain and co-domain as $\bar{x} = \sqrt{R}x$ and $\bar{y} = \sqrt{Q}y$. So, $\|x\|_R = \|\bar{x}\|$ and $\|Ax - b\|_Q = \|\sqrt{Q}A\sqrt{R}^{-1}\bar{x} - \sqrt{Q}b\|$. With an obvious notation the weighted problem is now equivalent to the original one, but in new coordinates:

$$\begin{array}{ll} \min_{\bar{x}} & \|\bar{x}\| \\ \text{s.t.} & \|\bar{A}\bar{x} - \bar{b}\| = \min\limits_{\bar{x}} \|\bar{A}\bar{x} - \bar{b}\| \end{array}$$

Of course, once the solution is found, it should be converted to the original coordinates through the inverse transformation. Even though this is not hard, it is usually possible (and preferable) to manipulate the expressions so that the solution is computed directly in the original coordinates.