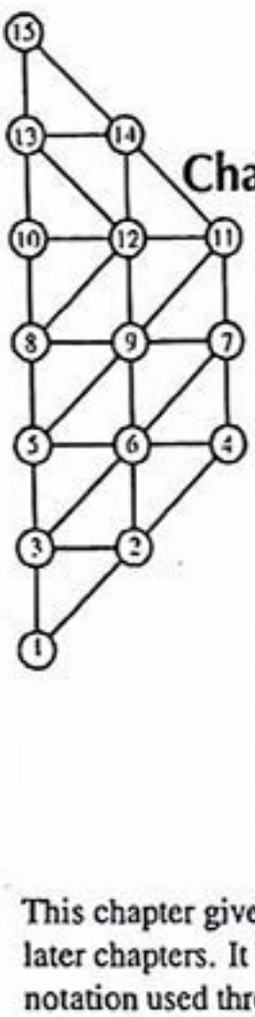

Chapter 1

Background in Linear Algebra

This chapter gives an overview of the relevant concepts in linear algebra that are useful in later chapters. It begins with a review of basic matrix theory and introduces the elementary notation used throughout the book. The convergence analysis of iterative methods requires a good level of knowledge in mathematical analysis and in linear algebra. Traditionally, many of the concepts presented specifically for these analyses have been geared toward matrices arising from the discretization of partial differential equations (PDEs) and basic relaxation-type methods. These concepts are now becoming less important because of the trend toward projection-type methods, which have more robust convergence properties and require different analysis tools. The material covered in this chapter will be helpful in establishing some theory for the algorithms and defining the notation used throughout the book.

1.1 Matrices

For the sake of generality, all vector spaces considered in this chapter are complex, unless otherwise stated. A complex $n \times m$ matrix A is an $n \times m$ array of complex numbers

$$a_{ij}, \quad i = 1, \ldots, n, \quad j = 1, \ldots, m.$$

The set of all $n \times m$ matrices is a complex vector space denoted by $\mathbb{C}^{n \times m}$. The main operations with matrices are the following:

- Addition: $C = A + B$, where A, B, and C are matrices of size $n \times m$ and

$$c_{ij} = a_{ij} + b_{ij}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, m.$$

- Multiplication by a scalar: $C = \alpha A$, where

$$c_{ij} = \alpha\, a_{ij}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, m.$$

- Multiplication by another matrix:

$$C = AB,$$

where $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times p}$, $C \in \mathbb{C}^{m\times p}$, and

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Sometimes, a notation with column vectors and row vectors is used. The column vector a_{*j} is the vector consisting of the jth column of A:

$$a_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Similarly, the notation a_{i*} will denote the ith row of the matrix A:

$$a_{i*} = (a_{i1}, a_{i2}, \ldots, a_{in}).$$

For example, the following could be written:

$$A = (a_{*1}, a_{*2}, \ldots, a_{*n})$$

or

$$A = \begin{pmatrix} a_{1*} \\ a_{2*} \\ \vdots \\ a_{m*} \end{pmatrix}.$$

The transpose of a matrix A in $\mathbb{C}^{m\times n}$ is a matrix C in $\mathbb{C}^{n\times m}$ whose elements are defined by $c_{ij} = a_{ji}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. It is denoted by A^T. It is often more relevant to use the transpose conjugate matrix denoted by A^H and defined by

$$A^H = \overline{A}^T = \overline{A^T},$$

in which the bar denotes the (element-wise) complex conjugation.

Matrices are strongly related to linear mappings between vector spaces of finite dimension. This is because they represent these mappings with respect to two given bases: one for the initial vector space and the other for the image vector space, or range, of A.

1.2 Square Matrices and Eigenvalues

A matrix is square if it has the same number of columns and rows, i.e., if $m = n$. An important square matrix is the identity matrix

$$I = \{\delta_{ij}\}_{i,j=1,\ldots,n},$$

where δ_{ij} is the Kronecker symbol. The identity matrix satisfies the equality $AI = IA = A$ for every matrix A of size n. The inverse of a matrix, when it exists, is a matrix C such that

$$CA = AC = I.$$

The inverse of A is denoted by A^{-1}.

The *determinant* of a matrix may be defined in several ways. For simplicity, the following recursive definition is used here. The determinant of a 1×1 matrix (a) is defined as the scalar a. Then the determinant of an $n \times n$ matrix is given by

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} \det(A_{1j}),$$

where A_{1j} is an $(n-1) \times (n-1)$ matrix obtained by deleting the first row and the jth column of A. A matrix is said to be *singular* when $\det(A) = 0$ and *nonsingular* otherwise. We have the following simple properties:

- $\det(AB) = \det(A)\det(B)$.
- $\det(A^T) = \det(A)$.
- $\det(\alpha A) = \alpha^n \det(A)$.
- $\det(\bar{A}) = \overline{\det(A)}$.
- $\det(I) = 1$.

From the above definition of determinants it can be shown by induction that the function that maps a given complex value λ to the value $p_A(\lambda) = \det(A - \lambda I)$ is a polynomial of degree n; see Exercise 8. This is known as the *characteristic polynomial* of the matrix A.

Definition 1.1. *A complex scalar λ is called an eigenvalue of the square matrix A if a nonzero vector u of $\mathbb{C}^n$ exists such that $Au = \lambda u$. The vector u is called an eigenvector of A associated with λ. The set of all the eigenvalues of A is called the spectrum of A and is denoted by $\sigma(A)$.*

A scalar λ is an eigenvalue of A *if and only if* (iff hereafter) $\det(A - \lambda I) \equiv p_A(\lambda) = 0$. That is true iff λ is a root of the characteristic polynomial. In particular, there are at most n distinct eigenvalues.

It is clear that a matrix is singular iff it admits zero as an eigenvalue. A well-known result in linear algebra is stated in the following proposition.

Proposition 1.2. *A matrix A is nonsingular iff it admits an inverse.*

Thus, the determinant of a matrix determines whether or not the matrix admits an inverse.

The maximum modulus of the eigenvalues is called the *spectral radius* and is denoted by $\rho(A)$:

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

The *trace* of a matrix is equal to the sum of all its diagonal elements:

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

It can be easily shown that the trace of A is also equal to the sum of the eigenvalues of A counted with their multiplicities as roots of the characteristic polynomial.

Proposition 1.3. *If λ is an eigenvalue of A, then $\bar{\lambda}$ is an eigenvalue of A^H. An eigenvector v of A^H associated with the eigenvalue $\bar{\lambda}$ is called a left eigenvector of A.*

When a distinction is necessary, an eigenvector of A is often called a right eigenvector. Therefore, the eigenvalue λ as well as the right and left eigenvectors u and v satisfy the relations

$$Au = \lambda u, \quad v^H A = \lambda v^H$$

or, equivalently,

$$u^H A^H = \bar{\lambda} u^H, \quad A^H v = \bar{\lambda} v.$$

1.3 Types of Matrices

The choice of a method for solving linear systems will often depend on the structure of the matrix A. One of the most important properties of matrices is symmetry, because of its impact on the eigenstructure of A. A number of other classes of matrices also have particular eigenstructures. The most important ones are listed below:

- *Symmetric matrices:* $A^T = A$.
- *Hermitian matrices:* $A^H = A$.
- *Skew-symmetric matrices:* $A^T = -A$.
- *Skew-Hermitian matrices:* $A^H = -A$.
- *Normal matrices:* $A^H A = AA^H$.
- *Nonnegative matrices:* $a_{ij} \geq 0$, $i, j = 1, \ldots, n$ (similar definition for nonpositive, positive, and negative matrices).
- *Unitary matrices:* $Q^H Q = I$.

It is worth noting that a unitary matrix Q is a matrix whose inverse is its transpose conjugate Q^H, since

$$Q^H Q = I \quad \rightarrow \quad Q^{-1} = Q^H. \tag{1.1}$$

A matrix Q such that $Q^H Q$ is diagonal is often called orthogonal.

Some matrices have particular structures that are often convenient for computational purposes. The following list, though incomplete, gives an idea of these special matrices, which play an important role in numerical analysis and scientific computing applications.

- *Diagonal matrices:* $a_{ij} = 0$ for $j \neq i$. Notation:

$$A = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn}).$$

- *Upper triangular matrices:* $a_{ij} = 0$ for $i > j$.
- *Lower triangular matrices:* $a_{ij} = 0$ for $i < j$.

المحاضرة الأولى:

"الجلسة الأولى":

يعطي هذا الفصل لمحة عامة عن مفهوم ذات صلة في الجبر الخطي التي هي مفيدة في فصل لاحق.

إنه ليس أكثر من مراجعة للنظرية الأساسية للمصفوفة وتقديم الترميز الذي سيتم استخدامه طوال الكتاب.

التحليل العددي للغرض الكاريزي يتطلب معرفة جيدة بالتحليل الرياضي والجبر الخطي. تقليدياً، إن العديد من المفاهيم المعتمدة بشكل محدود في هذه التحليلات تنشأ من أجل الارتقاء بالمصفوفات على المعادلات التفاضلية الجزئية ومسائل الشروط الحدية الأساسية.

هذه المفاهيم أصبحت الآن أقل أهمية بسبب الزيادة في المسائل الإحصائية والتي تتبع نهجات تقاربية مثيرة أكثر وتتطلب أدوات تحليلية مختلفة.

إن المادة التي سنغطيها في هذا الفصل ستكون مساعداً في تأسيس بعض الفرضيات حول الخوارزميات وتعريف الترميزات المستخدمة عبر الكتاب.

1-1 المصفوفات:

في الحالة العامة، كل الفضاءات الشعاعية المعتمدة في هذا الفصل هي حقيقية مالم ينص على خلاف ذلك.

المصفوفة المعتمدة من المرتبة $n \times m$ هي جدول من الأعداد المعتمدة حيث يأخذ بحيث من 1 حتى n و j يأخذ قيمة من 1 حتى m.

إن مجموعة كل المصفوفات من المرتبة $n \times m$ تمثل فضاء شعاعي معتمد يشار إليه بالرمز $\mathbb{R}^{n \times m}$.

العمليات الرئيسية على المصفوفة هي كالتالي:

- الجمع: $C = A + B$ حيث A, B, C هي مصفوفات من المرتبة $n \times m$.
- الضرب بعدد: حيث $C = \alpha A$ عندما $c_{ij} = \alpha a_{ij}$ $i = 1, 2, \ldots n$ ، $j = 1, 2, \ldots m$.

"الجلسة الثانية":

- الضرب بمصفوفة أخرى.

في بعض الأحيان يستخدم الترميز بالأسهم العمودية والأفقية.

إن الشعاع العمودي هو شعاع مكون من العدد من المرتبة n في المصفوفة A.

$$a_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

الشكل مشابه للرمز a_{i*} يشار إليه للدلالة على الصف من المرتبة i في المصفوفة A.

$$a_{i*} = (a_{i1}, a_{i2}, \dots a_{im})$$

على سبيل المثال يستطيع أن نكتب كالآتي: (مصفوفة سطرية أو عمودية)

$$A = (a_{*1}, a_{*2}, \dots, a_{*m})$$

$$A = \begin{pmatrix} a_{1*} \\ a_{2*} \\ \vdots \\ a_{n*} \end{pmatrix}$$

منقول المصفوفة A في الفضاء العقدي $\mathbb{C}$ من المرتبة $n \times m$ هو مصفوفة C في الفضاء العقدي $\mathbb{C}$ من المرتبة $m \times n$ $(\mathbb{C}^{m \times n})$ والتي عناصرها تعرف بالشكل:

$$c_{ij} = a_{ji} \quad , i = 1, \dots m, \; j = 1, \dots n$$

ونرمز بالشكل A^T.

في العادة هناك مهمة وثيقة أكثر في استخدام مرافق المنقول ومنقول المرافق ويرمز بالرمز A^H ويعرف بالعلاقات:

$$A^H = A^{-T} = \overline{A^T}$$

في هذا الخط نشير إلى مرافق المصفوفة (المرافق العقدي)

إن المصفوفات تتعلق بالتطبيقات الخطية بين الفضاءات الشعاعية منتهية البعد.

ذلك لأن تمثل هذه التطبيقات بالاعتماد على قاعدتين معطاتين.

الأولى للفضاء الشعاعي الأول (المنطلق)

والأخرى للمستقر الهدفي للفضاء الشعاعي أو مدى المصفوفة A.

1.2 المصفوفات المربعة والقيم الذاتية :

تكون المصفوفة مربعة إذا كان عدد أعمدتها مساوياً لعدد الصفوف مثلاً للتقطيع، إذا كان $n = m$.

مصفوفة مربعة خاصة هي المصفوفة الواحدية

$$I = \{ S_{ij} \}, \quad j = 1, \dots n$$

حيث S_{ij} هو الرمز kronecker .

إن المصفوفة الواحدية لها خاصية المساواة $AI = IA = A$ لكل مصفوفة A من المرتبة n

إن مقلوب المصفوفة A ، إن وجد ، هو مصفوفة C بحيث :

$$CA = AC = 1$$

نشير إلى مقلوب المصفوفة A بـ : A^{-1}.

" المحددة الثالثة " :

إن محدد المصفوفة يمكن أن يعرف بطرق عديدة :

للسهولة استخدم هنا الترتيب العمودي :

إن محدد المصفوفة ذات المرتبة 1x1 يعرف على أنه عدد a .

إذا " محدد المصفوفة من المرتبة nxn يعطى بالشكل :

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} \cdot a_{ij} \det(A_{ij})$$

حيث المصفوفة A_{ij} هي من المرتبة $(n-1) \times (n-1)$ تحصل عليها من حذف السطر الأول والعمود ذو الترتيب j في المصفوفة A .

نقال عن مصفوفة الأشاذة إذا كان محددها مساوياً الصفر ، أي $\det(A) = 0$

ونقال عكس ذلك غير شاذة ويحقق ذلك :

لدينا الخواص البسيطة التالية :

- $\det(AB) = \det(A) \cdot \det(B)$
- $\det(A^T) = \det(A)$
- $\det(\alpha A) = \alpha^n \det(A)$
- $\det(\bar{A}) = \overline{\det(A)}$
- $\det(I) = 1$

من التعريف العلوي للمحددات يمكن أن نكون من الواضح بالاستقراء الرياضي أن التابع الذي يربط كل مصفوفة مربعة A بالعدد $P_A(\lambda) = \det(A - \lambda I)$ هو كثير حدود من الدرجة n.

هذا نعرف بالحدودية المميزة للمصفوفة A.

تعريف 1.1:

العدد العقدي λ يدعى قيمة ذاتية للمصفوفة المربعة A إذا وجد شعاع غير صفري u من الفضاء $\mathbb{C}^n$ بحيث $Au = \lambda u$.

يدعى الشعاع u بالشعاع الذاتي للمصفوفة A مرتبط بالقيمة λ.

إن مجموعة كل القيم الذاتية للمصفوفة A يدعى طيف المصفوفة ويشار إليه بالرمز $\sigma(A)$.

العدد λ هو قيمة ذاتية للمصفوفة A إذا وفقط إذا:

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

هذا صحيح إذا وفقط إذا كان λ جذراً للمعادلة المميزة.

بالتعريف هناك على الأكثر n قيمة ذاتية مختلفة.

من الواضح أن تكون المصفوفة شاذة إذا وفقط إذا احتوت الصفر قيمة ذاتية λ.

نتيجة محصورة في الجبر الخطي نضيف إلى الادعاءات التالية:

الادعاء 1.2:

تكون المصفوفة A غير شاذة إذا وفقط إذا كان λ مختلف.

وهكذا، إن محدد المصفوفة يساوي جداء كل عناصر قطرها الرئيسي.

إن أثر المصفوفة يساوي مجموع كل عناصر قطرها الرئيسي.

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

"الصفة الرابعة":

إن من الواضح بسهولة أن أثر المصفوفة A تساوي أيضاً مجموع قيمها الذاتية محدودة بضاربها لجذور المعادلة المميزة.

الادعاء 1.3 :

إذا كانت λ قيمة ذاتية للمصفوفة A فإن مرافقها $\bar{\lambda}$ هو قيمة ذاتية للمصفوفة المرافقة A^H.

إن الشعاع الذاتي v للمصفوفة A^H المتعلق بالقيمة الذاتية $\bar{\lambda}$ يدعى شعاعاً ذاتياً يساري للمصفوفة A.

عندما نكون الاختلاف واضحاً فإن الشعاع الذاتي للمصفوفة A يدعى شعاعاً ذاتياً يمينياً.

لذا، إن القيمة الذاتية λ وأيضاً كل من الشعاع الذاتي اليميني واليساري يحققان العلاقات التالية:

$$Au = \lambda u \quad : \quad v^H A = \lambda v^H$$

أو بشكل مكافئ:

$$A^H v = \bar{\lambda} v \qquad u^H A^H = \bar{\lambda} u^H$$

المرجع المحاضرة الأولى