2

Sistemas de equações lineares

2.1 Matrizes

Vocabulário básico

As matrizes são “caixas” que contêm elementos numa forma ordenada.

Definição 2.1. *Dado qualquer $m \in \mathbb{N}$ e qualquer $n \in \mathbb{N}$, chamaremos* ***matriz de tamanho*** *$m \times n$ (ou dimensão $m \times n$) com entradas em $\mathbb{R}$ qualquer correspondência que permita determinar, para cada $i \in \{1, \ldots, n\}$ (a linha) e qualquer $j \in \{1, \ldots, n\}$ (a coluna), um* ***escalar***[1] *$a_{ij} \in \mathbb{R}$, que chamaremos entrada da matriz na linha i, coluna j.*

Denotamos o **conjunto das matrizes de tamanho** $m \times n$ com entradas em $\mathbb{R}$ por $M_{m\times n}(\mathbb{R})$.

Dar uma matriz é equivalente a dar uma coleção de $m \cdot n$ escalares ordenados numa tabela de tamanho $m \times n$, onde entendemos por tabela $m \times n$ o conjunto de todos os pares ordenados de números inteiros (i, j) com $1 \leq i \leq m$, $1 \leq j \leq n$, e por escalares ordenados na tabela qualquer aplicação que faz corresponder a cada (i, j) um elemento do conjunto de escalares, onde cada elemento da matriz corresponde a um índice “(i, j)” da tabela. Este elemento chama-se entrada correspondente ao índice (i, j), ou abreviadamente componente (i, j)-ésima da matriz. Uma matriz A de tamanho $m \times n$ pode ver-se, portanto, como uma aplicação:

$$
\begin{array}{rccc}
A\colon & \{(i,j)\colon 1 \leq i \leq m, 1 \leq j \leq n\} & \to & \mathbb{R} \\
 & (i,j) & \mapsto & A(i,j)
\end{array}
$$

Definição 2.2. *Dizemos que duas matrizes são* ***iguais*** *quando o seu tamanho é igual e as suas entradas são as mesmas na mesma ordem.*

Na representação escrita as entradas da matriz agrupam-se em tabelas com m linhas e n colunas, inseridas em parêntesis ou parêntesis retos, e denotam-se por letras maiúsculas. A identificação do duplo índice (i, j) correspondente a cada entrada é feita colocando o escalar associado aos índices (i, j) na posição correspondente da tabela –linha i, coluna j–.

Exemplo

Um exemplo de matriz de tamanho 5×3 *é*

$$
A = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 0 & 4 \\ 1 & 0 & \sqrt{2} \\ 31 & -4 & 0 \\ -3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \\ A_{41} & A_{42} & A_{43} \\ A_{51} & A_{52} & A_{53} \end{bmatrix}
$$

[1] Chamamos escalar qualquer elemento do conjunto dos números reais $\mathbb{R}$, mas a teoria de matrizes que pretendemos desenvolver é igualmente válida para matrizes de números complexos, inteiros, de polinómios..., conjuntos onde existam operações de adição e produto

Nesta matriz o elemento A_{23} é o situado na segunda linha, terceira coluna: o escalar 4.

Costuma-se usar letras maiúsculas para indicar matrizes e letras minúsculas para denotar escalares. Por isto para uma matriz A de tamanho $m \times n$, usaremos a_{ij} para designar as suas entradas. Esta notação pode resultar menos confusa do que A_{ij}, que, escrito com maiúsculas, parece indicar uma matriz e não um escalar.

Os pontos de $\mathbb{R}^n$ são sequências ordenadas de n escalares. A definição é formalmente equivalente à dada para matrizes de tamanho $1 \times n$ ou também para as de tamanho $n \times 1$.

Definição 2.3. *As matrizes de tamanho $m \times 1$ (portanto aquelas com uma única coluna) são conhecidas por **matrizes coluna** ou **vetores coluna**. As de tamanho $1 \times n$ são conhecidas por **matrizes linha** ou **vetores linha**.*

É frequente interpretar que as matrizes colunas são vetores, e representá-las com letras minúsculas (x, y, u, v, w,etc.).

Definição 2.4. *Quando A é uma matriz de tamanho $m \times n$, chamaremos linha i da matriz A, e denotaremos por $A_{i:}$, a matriz $1 \times n$ cujas entradas são $\begin{bmatrix} a_{i1} \; a_{i2} \; \ldots \; a_{in} \end{bmatrix}$.*

Chamaremos coluna j da matriz A, e denotamos por $A_{:j}$, a matriz $m \times 1$ cujas entradas são $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$.

Exemplo

A partir da matriz A antes dada podemos extrair as suas linhas e colunas. Exemplo de matriz linha e de matriz coluna: As matrizes $A_{2:}$, $A_{:3}$:

$$A_{2:} = \begin{bmatrix} 1 \; 0 \; 4 \end{bmatrix} \quad A_{:3} = \begin{bmatrix} 1 \\ 4 \\ \sqrt{2} \\ 0 \\ -1 \end{bmatrix}$$

Se temos uma matriz A de tamanho $m \times n$ e outra B de tamanho $m \times p$, podemos construir uma nova matriz que representamos por $\begin{bmatrix} A \; B \end{bmatrix}$, de tamanho $m \times (n+p)$, dada por:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & \vdots & \vdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1p} \\ b_{21} & \vdots & \vdots & b_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \ldots & b_{mp} \end{bmatrix}$$

$$\begin{bmatrix} A \; B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_{11} & b_{12} & \ldots & b_{1p} \\ a_{21} & \vdots & \vdots & a_{2n} & b_{21} & \vdots & \vdots & b_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_{m1} & b_{m2} & \ldots & b_{mp} \end{bmatrix}$$

De forma análoga, se temos uma matriz A de tamanho $m \times n$ e outra B de tamanho $q \times n$, podemos construir uma nova matriz que representamos por $\begin{bmatrix} A \\ B \end{bmatrix}$, de tamanho $(m+q) \times n$, dada por:

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & \vdots & \vdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & \vdots & \vdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{q1} & b_{q2} & \ldots & b_{qn} \end{bmatrix}$$

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \vdots & \vdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \\ b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & \vdots & \vdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{q1} & b_{q2} & \dots & b_{qn} \end{bmatrix}$$

Com estas definições, juntando matrizes de tamanhos apropriados, podemos construir **matrizes por blocos** :

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

onde A tem tamanho $m \times n$, B tem tamanho $m \times p$, C tem tamanho $q \times n$ e D tem tamanho $q \times p$.

Exemplo

Se consideramos as matrizes coluna $v_1 = \begin{bmatrix} 1 \ 2 \ 4 \ 3 \end{bmatrix}^t$ *e* $v_2 = \begin{bmatrix} 2 \ 2 \ 5 \ 0 \end{bmatrix}^t$*, então:*

$$\begin{bmatrix} v_1 \ v_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \\ 4 & 5 \\ 3 & 0 \end{bmatrix}$$

Se consideramos as matrizes:

$$A = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & -4 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \ 2 \ 3 \end{bmatrix}, \quad C = \begin{bmatrix} -1 \\ -1 \\ -2 \end{bmatrix}, \quad D = [7]$$

então podemos construir por blocos uma matriz 4×4 *dada por:*

$$\begin{bmatrix} A & C \\ B & D \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & -1 \\ 1 & 1 & -4 & -1 \\ 0 & 0 & 2 & -2 \\ 2 & 2 & 3 & 7 \end{bmatrix}$$

Nota 2.5. Em geral, uma matriz de tamanho $m \times n$ pode ser vista:

- Como uma coleção ordenada em tabela de $m \times n$ escalares $A = [a_{ij}]$
- Como uma coleção ordenada de n vetores coluna de dimensão $m \times 1$ (as colunas da matriz),
  $$A = \begin{bmatrix} v_1 \ v_2 \ \dots \ v_n \end{bmatrix} \text{ onde } v_j = A_{:j} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$
- ou como uma coleção ordenada de m vetores linha de dimensão $1 \times n$ (as linhas da matriz),
  $$A = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} \text{ onde } w_i = A_{i:} = \begin{bmatrix} a_{i1} \ a_{i2} \ \dots \ a_{in} \end{bmatrix}.$$

Definição 2.6. *Chamamos* ***matriz transposta*** *duma matriz A dada de tamanho $m \times n$ a matriz A^t de tamanho $n \times m$ cujas entradas são $(A^t)_{ij} = A_{ji}$, para cada $1 \leq i \leq n$, $1 \leq j \leq m$.*

Exemplo
A transposta duma matriz linha é uma matriz coluna.

$$\begin{bmatrix}1 & 4 & \sqrt{2} & 0 & -1\end{bmatrix}^t = \begin{bmatrix}1\\4\\\sqrt{2}\\0\\-1\end{bmatrix}$$

Exemplo
As seguintes matrizes são transpostas uma da outra:

$$A = \begin{bmatrix}3 & -2 & 1\\1 & 0 & 4\\1 & 0 & \sqrt{2}\\31 & -4 & 0\\-3 & 1 & -1\end{bmatrix}, \qquad A^t = \begin{bmatrix}3 & 1 & 1 & 31 & -3\\-2 & 0 & 0 & -4 & 1\\1 & 4 & \sqrt{2} & 0 & -1\end{bmatrix}$$

Exemplo
As matrizes coluna são escritas em forma mais simples se usamos a ideia de transposta, assim a segunda coluna $A_{:2}$ da matriz A anterior é:

$$A_{:2} = \begin{bmatrix}-2 & 0 & 0 & -4 & 1\end{bmatrix}^t$$

Propriedades da matriz transposta são:

$$(A^t)^t = A, \qquad (A^t)_{i:} = (A_{:i})^t, \qquad (A^t)_{:j} = (A_{j:})^t, \qquad \begin{bmatrix}A & B\\C & D\end{bmatrix}^t = \begin{bmatrix}A^t & C^t\\B^t & D^t\end{bmatrix}$$

Definição 2.7. *Uma matriz de tamanho $m \times n$ com $m = n$ diz-se que é uma **matriz quadrada**.*

Definição 2.8. *Numa matriz quadrada, dizemos que uma componente a_{ij} pertence à **diagonal principal** se corresponde a um índice $i = j$.*

Definição 2.9. *As matrizes que satisfazem $A^t = A$ conhecem-se por **matrizes simétricas**.*

Para que $A = A^t$, deve dar-se $m = n$ (a matriz A tem que ser quadrada) e também deve verificar-se $a_{ij} = a_{ji}$ para cada $1 \leq i, j \leq n$.

Definição 2.10. *Uma matriz quadrada A de tamanho $n \times n$ diz-se que é **triangular superior** se as suas entradas por baixo da diagonal são nulas, isto é se $a_{ij} = 0$ sempre que $i > j$.*

A representação na escrita destas matrizes é sempre da forma:

$$A = \begin{bmatrix}a_{11} & a_{12} & \dots & a_{1n}\\0 & a_{22} & \dots & \vdots\\\vdots & \vdots & \ddots & \vdots\\0 & 0 & \dots & a_{nn}\end{bmatrix}$$

Definição 2.11. *Uma matriz quadrada diz-se que é **triangular inferior** se as suas entradas por cima da diagonal são nulas, isto é, se $a_{ij} = 0$ sempre que $j > i$.*

A representação na escrita destas matrizes é sempre da forma:

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

Claramente, a transposta duma matriz triangular superior resulta uma matriz triangular inferior, e reciprocamente.

Definição 2.12. *Diremos que uma matriz é* ***diagonal*** *se as únicas entradas não nulas que contém estão na diagonal, isto é, se $a_{ij} = 0$ sempre que $i \neq j$.*

A representação na escrita duma matriz diagonal é:

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

Entre as matrizes diagonais, a mais destacada é a que tem entradas 1 na diagonal. Esta matriz é chamada **matriz identidade** de tamanho $n \times n$:

$$\mathrm{Id}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Operações aritméticas com matrizes

No conjunto de matrizes existem três operações relevantes: a soma, o produto com valores reais, e o produto de matrizes. O mais destacado e intrigante de todos é o produto de matrizes, que não é definido da maneira "evidente" (não é um produto das componentes correspondentes de cada matriz)

Definição 2.13. *Para qualquer matriz A e qualquer escalar (valor real) $\lambda \in \mathbb{R}$ chamamos produto $\lambda \cdot A$ a matriz de tamaho $m \times n$ cuja entrada (i, j) é o produto $\lambda \cdot a_{ij}$ da correspondente entrada na matriz A com o escalar λ*

$$(\lambda \cdot A)_{ij} = \lambda \cdot a_{ij}$$

Definição 2.14. *Se duas matrizes A, B têm* ***o mesmo tamanho*** *$m \times n$, chamamos soma $A + B$ das matrizes a matriz de tamanho $m \times n$ cuja entrada (i, j) é a soma $a_{ij} + b_{ij}$ das correspondentes entradas na matriz A e na matriz B.*

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

Definição 2.15. *Se as linhas da matriz A e as colunas da matriz B têm o mesmo tamanho (ou seja, A está formada por linhas de tamanho $1 \times n$ e B por colunas de tamanho $n \times 1$), chamamos* ***produto*** *das matrizes A, B a matriz cuja entrada (i, j) é o produto escalar da linha $A_{i:}$ de A com a coluna $B_{:j}$ de B (vistas como elementos de $\mathbb{R}^n$, e com o produto escalar natural de $\mathbb{R}^n$).*

$$(A \cdot B) = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + \ldots + a_{in} \cdot b_{nj} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

Assim quando A tem tamanho $m \times n$ e B tem tamanho $n \times q$ o produto $A \cdot B$ está definido sendo uma matriz de tamanho $m \times q$

Quando uma matriz A tiver o mesmo número de linhas do que colunas existem em B, diremos que **as matrizes A, B têm tamanhos compatíveis para o seu produto**.

O procedimento mais rápido para determinar manualmente a entrada (i, j) no produto $A \cdot B$ é "digital", usamos os nossos dedos indicadores, na matriz da esquerda avançamos nas entradas da linha i e na matriz da direita descemos nas entradas da coluna j, e somamos no nosso avanço o valor produto das duas entradas que estamos a indicar.

$$\begin{array}{cc} \begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix} & |\, i \end{array} \cdot \begin{array}{c} \underline{\quad j \quad} \\ \begin{bmatrix} \cdots\cdots & b_{1j} & \cdots \\ \cdots\cdots & b_{2j} & \cdots \\ \vdots\ \vdots & \vdots & \vdots \\ \cdots\cdots & b_{nj} & \cdots \end{bmatrix} \end{array} = \begin{array}{cc} \begin{array}{c} \underline{\quad j \quad} \\ \begin{bmatrix} \cdots\cdots & \cdot & \\ \vdots\ \ddots & & \vdots \\ \vdots & c_{ij} & \vdots \\ \cdots\cdots & \cdot & \ddots \end{bmatrix} \end{array} & |\, i \end{array}$$

$$c_{ij} = \begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix} \cdot \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}$$

As propriedades da soma e produto com números reais são similares às que já conhecemos para a soma e produto em $\mathbb{R}$:

Propriedades da adição de matrizes:

- Operação fechada: A soma de duas matrizes de $M_{m\times n}(\mathbb{R})$ é uma matriz de $M_{m\times n}(\mathbb{R})$. Não podem somar-se matrizes de tamanhos diferentes.
- Propriedade associativa: Para quaisquer matrizes $A, B, C \in M_{m\times n}(\mathbb{R})$,
  $(A + B) + C = A + (B + C)$.
- Propriedade comutativa: Para quaisquer matrizes $A, B \in M_{m\times n}(\mathbb{R})$,
  $A + B = B + A$.
- Existência de **elemento neutro**: A matriz $\mathbf{0}_{m\times n}$ definida por $\mathbf{0}_{ij} = 0$ para cada i, j, satisfaz
  $\mathbf{0} + A = A + \mathbf{0} = A$, para qualquer $A \in M_{m\times n}(\mathbb{R})$.
- Existência de **elemento oposto**: Para qualquer matriz $A \in M_{m\times n}(\mathbb{R})$, existe outra matriz $B \in M_{m\times n}(\mathbb{R})$ tal que
  $A + B = B + A = \mathbf{0}$.
  Este elemento B denotamo-lo por $-A$, e as suas entradas são os escalares $-a_{ij}$ opostos da correspondente entrada em A.

A existência de elemento oposto permitirá falar da operação de **diferença de matrizes**, denotaremos por $A - B$ simplesmente a soma de A com a matriz $-B$ oposta de B.

Com respeito da formação de matrizes com blocos temos:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} A+E & B+F \\ C+G & D+H \end{bmatrix} \text{ (para blocos do mesmo tamanho)}$$

Com respeito da transposição temos:

$$(A + B)^t = A^t + B^t$$

Propriedades do produto com escalares:

- Operação externa: O produto duma matriz de $M_{m\times n}(\mathbb{R})$ com um escalar é uma matriz de $M_{m\times n}(\mathbb{R})$.
- Propriedade associativa: Para qualquer matriz $A \in M_{m\times n}(\mathbb{R})$ e escalares $\lambda, \mu \in \mathbb{R}$,
  $(\lambda \cdot \mu) \cdot A = \lambda \cdot (\mu \cdot A)$.
- Propriedade distributiva com respeito à adição de escalares: Para qualquer matriz $A \in M_{m\times n}(\mathbb{R})$ e escalares λ, μ,
  $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$.
- Propriedade distributiva com respeito à adição de matrizes: Para quaisquer matrizes $A, B \in M_{m\times n}(\mathbb{R})$ e escalar λ,
  $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$.
- Existência de elemento identidade: A unidade $1 \in \mathbb{R}$ dos escalares satisfaz:
  $1 \cdot A = A$
  para qualquer matriz $A \in M_{m\times n}(\mathbb{R})$.
- O escalar nulo $0 \in \mathbb{R}$ verifica: $0 \cdot A = \mathbf{0}$ para qualquer $A \in M_{m\times n}(\mathbb{R})$ e para o $\mathbf{0} \in M_{m\times n}(\mathbb{R})$ elemento neutro da adição.
- Para qualquer $A \in M_{m\times n}(\mathbb{R})$ a matriz oposta (para a soma) é obtida através do produto com o oposto da unidade $-1 \in \mathbb{R}$:
  $-A = (-1) \cdot A$

Com respeito da formação de matrizes por blocos temos ainda:

$$\lambda \cdot \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \lambda \cdot A & \lambda \cdot B \\ \lambda \cdot C & \lambda \cdot D \end{bmatrix}$$

Com respeito da transposição temos:

$$(\lambda \cdot A)^t = \lambda \cdot A^t$$

O produto não está definido componente a componente, e tem assim uma interpretação e propriedades diferentes às que são conhecidas para o produto de números reais. Por exemplo, o produto não é comutativo, não podemos garantir $A \cdot B = B \cdot A$:

$$\begin{bmatrix} 3 & 1 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 7 & 11 \end{bmatrix} \neq \begin{bmatrix} 2 & 5 \\ 3 & 14 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 3 & 1 \\ -1 & 4 \end{bmatrix}$$

Propriedades do produto de matrizes:

- Operação externa: O produto duma matriz de $M_{m\times n}(\mathbb{R})$ com uma matriz de $M_{n\times p}(\mathbb{R})$ é uma matriz de tamanho $m\times p$.
- Propriedade associativa: Para quaisquer matrizes $A \in M_{p\times n}(\mathbb{R})$, $B \in M_{n\times q}(\mathbb{R})$, $C \in M_{q\times m}(\mathbb{R})$, o produto $A\cdot B \in M_{p\times q}(\mathbb{R})$ e $B\cdot C \in M_{n\times m}(\mathbb{R})$ satisfazem:
  $(A\cdot B)\cdot C = A\cdot (B\cdot C)$.
- Produto não comutativo: Quando duas matrizes $A \in M_{p\times n}(\mathbb{R})$, $B \in M_{n\times q}(\mathbb{R})$ satisfazem $A\cdot B = B\cdot A$, diz-se que A e B comutam. Nem todas as matrizes comutam. O produto de matrizes não é um produto comutativo.
- Propriedade distributiva respeito à soma de matrizes (na direita e na esquerda): Para quaisquer matrizes $A, B \in M_{p\times n}(\mathbb{R})$ e $C \in M_{n\times q}(\mathbb{R})$, verifica-se:
  $(A+B)\cdot C = A\cdot C + B\cdot C$
  Para quaisquer matrizes $A \in M_{p\times n}(\mathbb{R})$ e $B, C \in M_{n\times q}(\mathbb{R})$, verifica-se:
  $A\cdot (B+C) = A\cdot B + A\cdot C$.
- Relação com o produto com escalares: Para qualquer escalar $\lambda \in \mathbb{R}$ e matrizes $A \in M_{p\times n}(\mathbb{R})$, $B \in M_{n\times q}(\mathbb{R})$, verifica-se:
  $(\lambda\cdot A)\cdot B = A\cdot(\lambda\cdot B) = \lambda\cdot(A\cdot B)$.
- A matriz identidade Id_n satisfaz:
  $A\cdot \mathrm{Id}_n = A$, para qualquer escolha de $A \in M_{m\times n}(\mathbb{R})$.
  $\mathrm{Id}_n \cdot A = A$, para qualquer escolha de $A \in M_{n\times q}(\mathbb{R})$.

Com respeito da formação de matrizes com blocos temos:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} + \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} (A\cdot E + B\cdot G) & (A\cdot F + B\cdot H) \\ (C\cdot E + D\cdot G) & (C\cdot F + D\cdot H) \end{bmatrix} \text{ (para blocos compatíveis)}$$

Com respeito da transposição temos (cuidado com a mudança da ordem):

$$(A\cdot B)^t = B^t \cdot A^t$$

Assim quando temos matrizes decompostas em blocos, podemos aplicar sem problema o critério de "indicar os blocos" ao longo duma linha com o dedo indicador esquerdo, indicar os blocos ao longo duma coluna com o dedo indicador direito, e somar os vários produtos de dois blocos que estamos a indicar neste processo.

Nota 2.16. Se usamos matrizes e as operações disponíveis nas matrizes, obtemos uma aritmética de matrizes onde nem todas as propriedades são semelhantes às da aritmética de números reais. Muitas propriedades do produto de números reais deixam de ser agora válidas para o produto de matrizes.

Qualquer real diferente do zero tem um inverso: multiplicar com um determinado elemento (o inverso) produz a unidade $1 \in \mathbb{R}$. Nas matrizes a unidade para o produto é a matriz Id_n. No entanto muitas matrizes, para além da matriz nula, não vão ter inversa. De facto escrever:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \mathrm{Id}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

não é possível, seja qual for a, b, c, d (estamos a exigir $a+c = 1$ ao mesmo tempo que $a+c = 0$).

Na aritmética de matrizes, podemos usar **variáveis matriciais**, isto é, símbolos sem significado próprio que podem ser substituídos por matrizes concretas. As operações aritméticas aplicadas em matrizes e em variáveis matriciais levam a **expressões algébricas matriciais**, e equações que usam estas expressões. Assim podemos considerar a equação matricial $X\cdot X = B$ onde X é uma variável matricial.

Qualquer matriz A solução da igualdade anterior pode ser considerada uma raiz quadrada da matriz B. Contrariamente ao caso dos números reais, uma matriz B concreta pode ter mais do que duas raízes quadradas:

$$\begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 25 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 25 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & -5 \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 \\ 0 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 2.5 \\ 10 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 2.5 \\ 10 & 0 \end{bmatrix} = \dots$$

Entre as expressões mais simples com matrizes estão as expressões lineares $A \cdot X$, onde A é uma matriz conhecida e X é uma matriz incógnita.

Estudemos o problema de existência de inversa. Encontramos para começar que o produto não é comutativo, não é o mesmo perguntar se existe uma matriz B que satisfaz $B \cdot A = \mathrm{Id}$ do que perguntar se existe uma matriz B que satisfaz $A \cdot B = \mathrm{Id}$.

Chamaremos **equação matricial linear** qualquer equação $A \cdot X = B$, onde A, B são matrizes conhecidas e X é a matriz incógnita.

Duas equações matriciais lineares importantes são $A \cdot X = \mathrm{Id}$ e $A^t \cdot X = \mathrm{Id}$. A primeira tem como solução as matrizes X que são inversas à direita de X. A segunda tem como solução as matrizes X cuja transposta satisfaz $X^t \cdot A = \mathrm{Id}$, portanto aquelas onde X^t é inversa à esquerda de A.

Definição 2.17. *Dizemos que A é **invertível à direita** se a equação matricial $A \cdot X = \mathrm{Id}$ tem solução. Qualquer solução B é chamada uma **inversa à direita de** A*

*Dizemos que A é **invertível à esquerda** se a equação matricial $X \cdot A = \mathrm{Id}$ tem solução (equivalentemente, $A^t \cdot X^t = \mathrm{Id}$). Qualquer solução B é chamada **inversa à esquerda de** A*

*Dizemos que A é **invertível** se for invertível à esquerda e à direita.*

Observamos que A é invertível à esquerda se e só se A^t é invertível à direita, e que o problema de encontrar inversas fica reduzido ao problema de resolver equações matriciais lineares $A \cdot X = \mathrm{Id}$ ou $A^t \cdot X^t = \mathrm{Id}$

Exemplo

A matriz $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ *tem inversa à esquerda:*

$$\begin{bmatrix} 1 & 0 & -2 \\ 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathrm{Id}_2$$

Esta inversa ajuda a resolver por exemplo uma equação do tipo $X \cdot A = C$. Por exemplo:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{33} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 0 & 1 \\ -1 & 2 \end{bmatrix}$$

tem como solução a matriz:

$$X = \begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 0 & 1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -2 \\ 1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & -3 & -5 \\ 9 & -5 & -13 \\ 1 & -1 & -1 \\ 1 & -2 & 0 \end{bmatrix}$$

Com efeito, esta matriz verifica:

$$X \cdot A = \begin{bmatrix} 4 & -3 & -5 \\ 9 & -5 & -13 \\ 1 & -1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \left(\begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 0 & 1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -2 \\ 1 & -1 & -1 \end{bmatrix} \right) \cdot \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 0 & 1 \\ -1 & 2 \end{bmatrix} \cdot \left(\begin{bmatrix} 1 & 0 & -2 \\ 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \right) =$$

$$= \begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 0 & 1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 4 & 5 \\ 0 & 1 \\ -1 & 2 \end{bmatrix} = C$$

o que prova que existe uma solução, mas não que esta solução seja a única.

Exemplo

A matriz $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix}$ *admite inversa à direita, e não só uma, como vemos no exemplo:*

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1/2 & -1 \\ 2 & -1 \\ -1/2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathrm{Id}_2$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 2 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathrm{Id}_2$$

Proposição 2.18. *Se $A \in M_{m\times n}(\mathbb{R})$ tem inversa à direita e inversa à esquerda, ambas são a mesma matriz. Esta é chamada* ***a matriz inversa de*** *A, e denotada por A^{-1}*

Prova: Como se pode ver com as seguintes operações:

$$\left.\begin{array}{l} B \cdot A = \mathrm{Id}_n \\ A \cdot \bar{B} = \mathrm{Id}_m \end{array}\right\} \Rightarrow \left.\begin{array}{l} (B \cdot A) \cdot \bar{B} = \mathrm{Id}_n \cdot \bar{B} = \bar{B} \\ B = B \cdot \mathrm{Id}_m = B \cdot (A \cdot \bar{B}) \end{array}\right\} \Rightarrow B = B \cdot (A \cdot \bar{B}) = (B \cdot A) \cdot \bar{B} = \bar{B}$$

Para uma matriz A, se existe alguma inversa à direita e alguma inversa à esquerda de A, então todas as inversas pela direita coincidem com todas as inversas pela esquerda. Portanto a inversa à direita ou à esquerda deve ser única, e vamos representá-la por A^{-1}.

Mais ainda, se encontramos uma solução de $A \cdot X = \mathrm{Id}$ (inversa à direita de A), procurar se existe inversa à esquerda supõe simplesmente comprovar se esta solução X é única e satisfaz $X \cdot A = \mathrm{Id}$. Se a inversa à direita não for única, não podem existir inversas à esquerda, e es a inversa à direita não satisfaz $X \cdot A = \mathrm{Id}$, não podem existir inversas à esquerda.

Nota 2.19. A existência de oposto para a soma permite falar da diferença de matrizes. No entanto não podemos falar na divisão de matrizes C/A porque iríamos encontrar duas alternativas: como o produto $C \cdot A^{-1}$, ou como o produto $A^{-1} \cdot C$, que em geral não produzem o mesmo resultado.

Entre as propriedades do cálculo de inversa temos:

$$(\lambda A)^{-1} = \frac{1}{\lambda} \cdot A^{-1}, \qquad (A \cdot B)^{-1} = B^{-1} \cdot A^{-1}, \qquad (A^t)^{-1} = (A^{-1})^t$$

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & (-A^{-1}BC^{-1}) \\ 0 & C^{-1} \end{bmatrix}$$

As únicas matrizes invertíveis são as matrizes quadradas (e nem todas as matrizes quadradas), mas para provar este resultado temos que esperar pelo capítulo onde resolveremos equações lineares matriciais por redução Gaussiana.

Norma matricial e número de condição

Uma questão relevante agora é: Se queremos fazer um produto $A \cdot x$, sendo que x é introduzido em forma aproximada como x^*, será que o erro dos dados de entrada se propaga de maneira descontrolada nos valores de saída?.

Proposição 2.20. *Fixemos uma norma $\| \cdot \|_p$ nos espaços vetoriais $\mathbb{R}^n$, $\mathbb{R}^m$.*

Para qualquer matriz $A \in M_{m\times n}(\mathbb{R})$, consideremos os seguintes valores:

$$M(A) = \max_{\|u\|_p=1} \|A \cdot u\|_p$$

$$m(A) = \min_{\|u\|_p=1} \|A \cdot u\|_p$$

Então o produto com A para vetores arredondados irá propagar o erro absoluto dos dados de entrada como segue:

$$\Delta_p(A \cdot x^*, A \cdot x) \leq M(A) \cdot \Delta_p(x^*, x)$$

sendo o erro absoluto Δ_p medido com a norma-p, e sendo a igualdade atingida para determinadas aproximações x^* de x

O produto com A para vetores arredondados irá propagar o erro relativo dos dados de entrada como segue:

$$\delta_p(A \cdot x^*, A \cdot x) \leq \frac{M(A)}{m(A)} \cdot \delta_p(x^*, x)$$

sendo novamente o erro relativo medido com a norma-p, e sendo a igualdade atingida para determinadas aproximações x^* de x

O valor $M(A)$ é chamado **norma da matriz** A, e serve como número de condição associado à matriz A para erros absolutos. O valor $M(A)/m(A)$ é o **número de condição** associado à matriz A para erros relativos.

Demonstração. Para qualquer x^* aproximação de x com erro absoluto $\epsilon \geq 0$ temos $x^* - x = \epsilon \cdot u$, onde u é um vetor unitário ($\|u\|_p = 1$)

O erro absoluto de $A \cdot x^*$ como aproximação de $A \cdot x$ é então: $\|A \cdot x^* - A \cdot x\|_p = \|A \cdot (x^* - x)\|_p = \|A \cdot \epsilon u\|_p = \epsilon \cdot \|A \cdot u\|_p$

Como chamamos $M(A)$ precisamente o máximo dos valores $\|A \cdot u\|$ quando u é unitário, deduzimos o resultado indicado no enunciado. As aproximações que pior se comportam para o produto com A são então aquelas onde $x^* - x$ seja um múltiplo do vetor u unitário que gera o valor máximo de $\|A \cdot u\|_p$.

Com respeito do erro relativo, basta observar que as aproximações x^* de x com erro relativo δ são vetores $x^* = x + \delta \cdot \|x\|_p \cdot u$ para u unitário. Portanto:

$$\delta_p(A \cdot x^*, A \cdot x) = \frac{\|A \cdot x^* - A \cdot x\|_p}{\|A \cdot x\|_p} = \frac{\|A \cdot (x^* - x)\|_p}{\|A \cdot x\|_p} = \frac{\|A \cdot (\delta \cdot \|x\|_p \cdot u)\|_p}{\|A \cdot x\|_p} =$$
$$= \delta \cdot \frac{\|x\|_p \cdot \|A \cdot u\|_p}{\|A \cdot x\|_p} = \delta \cdot \frac{\|A \cdot u\|_p}{\|A \cdot (x/\|x\|_p)\|_p}$$

como $x/\|x\|_p$ é um vetor unitário e u é outro vetor unitário, o fator que multiplica com δ será no máximo o valor $M(A)/m(A)$. As aproximações que pior se comportam para o erro relativo são aquelas onde x seja um múltiplo do vetor unitário v que produz o mínimo possível de $\|A \cdot v\|_p$, e onde $x^* - x$ seja múltiplo do vetor unitário u que produz o máximo possível de $\|A \cdot u\|_p$ $\square$

O valor $M(A)$ indicado no resultado acima resulta ter as propriedades exigidas às normas. Este valor é chamado norma-p da matriz e proporciona mecanismos para medir o "tamanho" da matriz, e portanto para calcular erros se usamos uma matriz como aproximação de outra. Temos assim o conceito de **norma matricial**

Definição 2.21. *Fixemos uma norma $\|\cdot\|_p$ para medições em vetores de $\mathbb{R}^n$. Isto permite medir a norma dum vetor coluna (que é identificado com um elemento de $\mathbb{R}^n$).*

Para qualquer matriz de tamanho $m \times n$, $A \in M_{m\times n}$, chamamos **norma-p da matriz** *o valor:*

$$\|A\|_p = \max_{\|u\|_p=1} \|A \cdot u\|_p$$

Isto é, a norma-p de A mede até que ponto conseguimos esticar vetores unitários ao multiplicarmos com A.

Como já foi indicado no caso dos vetores de $\mathbb{R}^n$, a existência duma norma no conjunto de matrizes irá permitir medir até que ponto uma matriz A^* é uma boa aproximação duma matriz A. Em particular, podemos falar do erro absoluto e erro relativo de A^* como aproximação de A.

Ao considerarmos uma das normas-p nos espaços $\mathbb{R}^n$, a correspondente norma matricial satisfaz:

- É definido-positiva: $\|A\| > 0, \qquad \forall A \in M_{m\times n}(\mathbb{R}), \quad A \neq 0_{m\times n}$
- É homogénea: $\|\lambda \cdot A\| = |\lambda| \cdot \|A\|$ $\forall A \in M_{m\times n}(\mathbb{R})$, $\forall \lambda \in \mathbb{R}$
- É sub-aditiva: $\|A + B\| \leq \|A\| + \|B\|$, $\forall A, B \in M^{m\times n}(\mathbb{R})$
- É sub-multiplicativa: $\|A \cdot B\| \leq \|A\| \cdot \|B\|$, $\forall A \in M_{m\times n}(\mathbb{R})$, $B \in M_{n\times q}(\mathbb{R})$
- Em particular $\|A \cdot x\| \leq \|A\| \cdot \|x\|$, $\forall A \in M_{m\times n}(\mathbb{R})$, $x \in M_{n\times 1}(\mathbb{R})$
- O valor $M(A) = \max\limits_{\|u\|_p=1} \|A \cdot u\|_p$ coincide com $\|A\|_p$ e se A for invertível com inversa A^{-1}, o valor $m(A) = \min\limits_{\|u\|_p=1} \|A \cdot u\|_p$ coincide com $1/\|A^{-1}\|_p$

Em particular, o **número de condição para o erro relativo** associado a uma matriz invertível A é:

$$\kappa_p(A) = \|A\|_p \cdot \|A^{-1}\|_p$$

Mais ainda, a norma irá permitir definir a matriz limite duma sucessão de matrizes $A(k)$, e a ordem de convergência (linear, quadrática, etc). A noção de convergência e a ordem de convergência novamente não depende da norma escolhida. Se há convergência de ordem $r > 1$ duma sucessão de matrizes quando trabalhamos com uma norma, existirá convergência com a mesma ordem $r > 1$, quando trabalhamos com outra norma.

Nota 2.22. No caso da norma $\|\cdot\|_2$ em $\mathbb{R}^n$ (norma euclidiana), determinar a norma matricial associada não resulta nada simples. A norma duma matriz coincide com a raiz quadrada do maior auto-valor da matriz $A^t \cdot A$. O conceito de auto-valor fica muito para além dos objetivos desta sebenta, e o seu cálculo preciso não em nada simples, exige encontrar raízes dum polinómio.

Existe um método, no entanto, para saber se $\|A\|_2 > \lambda$ (equivalentemente, se $\lambda^2\mathrm{Id} - A^t \cdot A$ tem só auto-valores positivos). Para uma matriz simétrica $S = A^t \cdot A$, determinar se $\|S\|_2^2 > \lambda^2$ é possível como segue: Basta determinar se $M_\lambda = \lambda^2\mathrm{Id} - S$ é definido-positiva, e como veremos isto acontece se e só se esta matriz simétrica M_λ tem uma decomposição de Cholesky.

Concluímos que:

$$\|A\|_2 > \lambda \Leftrightarrow \lambda^2\mathrm{Id} - A^tA \text{ tem decomposição de Cholesky}$$

No entanto, para a norma-∞ e para a norma-1 o cálculo é simples:

Proposição 2.23. *Se A é uma matriz composta por linhas $v_1, \ldots, v_m$, e se estas linhas têm norma-1 dada por $S_1 = \|v_1\|_1$, $S_2 = \|v_2\|_1$, …, $S_m = \|v_m\|_1$ então a norma$-\infty$ da matiz A está dada por:*

$$\|A\|_\infty = \max(S_1, S_2, \ldots, S_m)$$

Se A é uma matriz composta por colunas $v_1, \ldots, v_n$, e se estas colunas têm norma-1 dada por $S_1 = \|v_1\|_1$, $S_2 = \|v_2\|_1$, …, $S_n = \|v_n\|_1$ então a norma$-\infty$ da matiz A está dada por:

$$\|A\|_1 = \max(S_1, S_2, \ldots, S_n)$$

Exemplo

Consideremos a matriz $A = \begin{bmatrix} 1 & -2 & 5 \\ -2 & -1 & 8 \end{bmatrix}$

As linhas têm norma-1 *dada por* $S_1 = |1| + |-2| + |5| = 8$, $S_2 = |-2| + |-1| + 8 = 11$, *portanto* $\|A\|_\infty = \max(8, 11) = 11$

As colunas têm norma-1 *dada por* $S_1 = |1| + |-2| = 3$, $S_2 = |-2| + |-1| = 3$, $S_3 = |5| + |8| = 13$, *portanto* $\|A\|_1 = 13$

Nota 2.24. O produto de matrizes pode estar mal condicionado. Pensemos no seguinte exemplo

$$A = \begin{bmatrix} 1 & -6 & 7 & -9 \\ 1 & -5 & 0 & 0 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 1 & -5 \end{bmatrix}, \quad x = \begin{bmatrix} 604 \\ 121 \\ 24 \\ 5 \end{bmatrix}, \quad x^* = \begin{bmatrix} 605 \\ 120 \\ 25 \\ 4 \end{bmatrix}, \quad x - x^* = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

As matrizes x^*, x escolhidas são bastante similares. O erro relativo de x^* comparado com x é simples de calcular se usamos a norma-∞:

$$\delta(x^*, x) = \frac{\|x^* - x\|_\infty}{\|x\|_\infty} = \frac{1}{604}$$

No entanto o produto $A \cdot x$ resulta ser muito diferente de $A \cdot x^*$:

$$A \cdot x = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad A \cdot x^* = \begin{bmatrix} 24 \\ 5 \\ -5 \\ 5 \end{bmatrix}, \quad A \cdot x^* - A \cdot x = \begin{bmatrix} 23 \\ 6 \\ -6 \\ 6 \end{bmatrix}$$

Sendo x^* muito similar a x, os produtos com A resultam muito diferentes um do outro. Medido o erro relativo com a norma-∞ temos:

$$\delta(A \cdot x^*, A \cdot x) = \frac{\|A \cdot x^* - A \cdot x\|_\infty}{\|A \cdot x\|_\infty} = \frac{23}{1}$$

Apesar de que x^* tem um erro relativo pequeno com respeito de x (erro relativo $1/604$), a matriz $A \cdot x^*$ tem um erro relativo grande com respeito de $A \cdot x$ (erro relativo 23), o erro relativo do dado de entrada fica multiplicado com 23×604, quase catorze mil vezes o erro relativo original.

As matrizes que têm esta caraterística, puderem ampliar de maneira forte o erro relativo dos dados de entrada, são chamada **matrizes mal condicionadas**. Nestas matrizes muitos cálculos arredondados podem introduzir grandes erros devidos aos arredondamentos de ponto flutuante. Aprenderemos mais sobre estas matrizes nas próximas secções.

Este não é um problema que nos deva preocupar se trabalhamos com valores exatos.

2.2 Resolução de Sistemas de Equações lineares: Métodos diretos

Dar uma equação exige indicar uma aplicação $f\colon X \to Y$ e um ponto $b \in Y$ do conjunto de destino. A equação determinada por esta aplicação f e ponto b é o problema de determinar os elementos $x \in X$ no conjunto inicial cuja imagem por f é o elemento b fixado.

Assim falaremos da equação $f(x) = b$, onde b é o que chamamos **termo independente da equação** e f é o que chamamos **função que define a equação**, função que se aplica em elementos do conjunto X, representados através duma **variável**, normalmente denotada x e chamada **variável incógnita** da equação.

Uma solução da equação $f(x) = b$ é qualquer ponto $x \in X$ cuja imagem por f seja o elemento $b \in Y$