Álgebra Linear

Leis de uma Álgebra:

Uma lei interna T associa dois elementos x e y de um conjunto A à um terceiro elemento $z \in A$. Por exemplo: $x+y=z$ onde a lei interna é a adição $+$, ou $x\,y=z$, onde a lei interna é a multiplicação $\times$. Na nossa notação se afirma que $xT\,y=z \quad \rightarrow \quad x+y=z$ ou $xT\,y=z \quad \rightarrow \quad x\,y=z$.

A lei será associativa se: $(xT\,y)T\,z = xT(yT\,z)$ e será comutativa se $xT\,y = yT\,x$.
O elemento $a \in A$ será REGULAR se $xT\,a = yT\,a \quad \leftrightarrow \quad x=y$ e $aT\,x = aT\,y \quad \leftrightarrow \quad x=y$. O elemento $e \in A$ será unitário sobre a lei T se $xT\,e = eT\,x = x$.
O elemento $x \in A$ possui elemento inverso x' sobre a lei T se $\exists x' \in A / xT\,x' = x'T\,x = e$.

Teorema: se uma lei T possui elemento unitário, é associativa e $x \in A$ possui inversa, então a inversa é única e x é regular.

Provar por absurdo: supor que $\exists\ x' \neq x''$ inversos de x.

Nesse caso: $xT\,x' = e$ e $xT\,x'' = e$ além disso $x''T(xT\ x') = x''T\,e = x''$. Agora, como T é associativo então $x''T(xT\ x') = (x''T\,x)T\,x' = eT\,x' = x'$ o que implica que $x'' = x'$ em contradição com $x' \neq x''$. Logo a inversa é única.

Falta a regularidade: Seja $a = xT\,y$ e $b = xT\,z$. Queremos mostrar que se $a=b$ então $y=z$. Aplicando a inversa de x de ambos os lados temos que $x'T(xT\,y) = x'T(xT\,z)$ e usando a associatividade $(x'T\,x)T\,y = (x'T\,x)T\,z$ logo $eT\,y = eT\,z$ e, finalmente $y=z$ C.Q.D.

Álgebra com duas Leis. Vamos chamar a lei T de lei 1 e uma outra lei 2 de $\perp$. A lei $\perp$ será distributiva frente à lei T se $(xT\,y)\perp z = (x\perp z)T(y\perp z)$ ou $z\perp(xT\,y) = (z\perp x)T(z\perp y)$ para $\forall\, x,y,z \in A$. Exemplo: mostrar que a multiplicação é distributiva frente à adição mas o inverso não é verdade. $(x+y)z = x\,z + y\,z$ logo a mulitplicação é distributiva. Entretanto $x\,y + z \neq (x+z)(y+z)$.

GRUPO. Um conjunto G é um grupo se $\exists$ uma lei interna T com as seguintes propriedades:

1. T é associativa
2. T admite elemento unitário e
3. $\forall\, x \in G$ admite inversa $x' \in G$

Se, além da propriedade associativa, T for comutativa, o grupo é chamado de Abeliano.
CAMPOS. Seja G um grupo Abeliano de T com uma segunda lei $\perp$ associativa e distributiva frente à T. Seja e o elemento unitário de G e G^*, onde G^* é o conjunto de todos os elementos de G exceto e. Se $\perp$ é uma lei de grupo para G^* então G é um campo.

Exemplo: vamos considerar as leis $+$ e $\times$ para o conjunto dos números reais.

Primeira lei $+$:

$\left(x+y\right)+z=x+\left(y+z\right)$ associativa

$x+y=y+x$ comutativa

$x+e_{+}=e_{+}+x=x$ então $e_{+}=0$ é o elemento unitário da adição.

Inversa $x+\left(-x\right)=e_{+}$ e $-x$ é a inversa de x. Note que se o conjunto fosse o dos naturais não haveria inversa pois ele não inclui números negativos. Então $\square=G$ e $G^{*}=\left\{x\in\square\ /\ x\neq 0\right\}$.

Agora vamos analisar o comportamento da multiplicação frente ao G^{*}.

$\left(x\,y\right)z=x\left(y\,z\right)$ associativa

$x\,y=y\,x$ comutativa

$\left(x+y\right)z=x\,z+y\,z$ distributiva frente à adição

$x\,e_{\times}=e_{\times}\,x=x$ então $e_{\times}=1$ é o elemento unitário da multiplicação. Logo a inversa será dada por $x\,x'=1$, ou seja, $x'=\dfrac{1}{x}$. Note que sem a exclusão do zero teríamos problema com a inversa do elemento unitário da adição pois $0'=\dfrac{1}{0}$.

Então □ é um CAMPO frente à adição e multiplicação. Note a necessidade de ampliar os conjuntos para a obtenção de grupos e campos. Partindo dos naturais □ e da operação $+$ foi necessário incluir os números negativos para a existência da inversa, chegando ao conjunto □ .

Já para a operação multiplicação foi-se obrigado a inlcuir o conjunto dos racionais □ para admitir inversas $x'=\dfrac{1}{x}$. Além disso, para admitir operações como $x\,x=y$ com $x\in G$ e $y\in G$ percebe-se a necessidade de inclusão dos irracionais e dos imaginários, caso $y<0$.

Matrizes

Uma matriz é uma tabela com m linhas e n colunas em que cada elemento é identificado pelos índices da linha e da coluna. Assim o elemento a_{ij} é encontrado no cruzamento da linha i com a coluna j, ou seja, o primeiro índice se refere à linha e o segundo à coluna.

$$\begin{array}{rc}
\begin{array}{r} coluna \rightarrow \\ linha \downarrow \end{array} & \begin{array}{ccccc} 1 & 2 & 3 & \cdots & n \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{array} \\
\begin{array}{r} 1 \rightarrow \\ 2 \rightarrow \\ 3 \rightarrow \\ \vdots \\ m \rightarrow \end{array} & \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}
\end{array}$$

Note que quando denotamos uma matriz $A_{m\times n}$ temos um bloco com $n \times m$ números. Uma analogia é com um circuito integrado, composto de milhões de elementos individuais. Outra analogia seriam as rotinas dos softwares que podem ser utilizadas como blocos individuais. Se existe uma álgebra para lidar com os blocos sem muita necessidade de descer ao nível dos elementos individuais, isso representa uma enorme economia de esforço intelectual. Assim, se for possível simplificar todo um conjunto de operação com matrizes apenas no nível mais alto, mais abstrato, o trabalho se torna muito mais produtivo. Por isso a álgebra de matrizes é uma ferramenta matemática das mais poderosas.

Exemplo: matriz Insumo-Produto. Para produzir q_i unidades do produto i é necessário usar $a_{i1}, a_{i2}, \cdots, a_{in}$ unidades dos insumos $1, 2, \cdots, n$. Nesse exemplo podemos escrever a matriz na forma:

$$\begin{array}{rcc}
\begin{array}{r} Insumo \rightarrow \\ produto \downarrow \end{array} & \begin{array}{ccccc} I_1 & I_2 & I_3 & \cdots & I_n \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{array} & \begin{array}{c} produto \\ \downarrow \end{array} \\
\begin{array}{r} produto_1 \rightarrow \\ produto_2 \rightarrow \\ produto_3 \rightarrow \\ \vdots \\ produto_m \rightarrow \end{array} & \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} & \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_m \end{bmatrix}
\end{array}$$

Vamos formalizar uma “ÁLGEBRA” de matrizes.

Dimensão: uma matriz de dimensão $m \times n$ [m por n] possui m linhas e n colunas. O elemento A_{ij} está no encontro da iésima linha com a jésima coluna.
Matriz NULA: $O_{m\times n} \,/\, O_{ij} = 0 \quad \forall i, j$, ou seja, todos os elementos são nulos.
Matriz COLUNA: $C_{m\times 1}$ tem apenas uma coluna e m linhas
Matriz LINHA: $L_{1\times n}$ tem apenas uma linha e n colunas.
Matriz QUADRADA: se $n = m$, A é uma matriz quadrada A_{mm} de ordem m.

Função Delta de Kronecker: $\delta_{ij} = \begin{cases} 1 & se \quad i = j \\ 0 & se \quad i \neq j \end{cases}$

Matrizes quadradas especiais:

1. DIAGONAL: $D_{ij} = 0 \quad se\, i \neq j$, já D_{ii} pode tomar qualquer valor. Em termos da Delta de Kronecker pode ser escrita como: $D_{ij} = d_i\, \delta_{ij}$ onde d_i é o elemento da diagonal i. A δ_{ij} garante que todos os elementos fora da diagonal serão nulos.
2. IDENTIDADE: Matriz diagonal na qual $I_{ij}^{(n)} = \delta_{ij}$, onde n denota a ordem da matriz quadrada. Ou seja, é uma matriz com todos os elementos da diagonal iguais a UM e todos os elementos fora da diagonal iguais a ZERO.
3. Matriz TRIANGULAR SUPERIOR: $\Delta_{ij}^{+} = 0 \quad se \quad i > j$, ou seja, todos os elementos abaixo da diagonal são NULOS.
4. Matriz TRIANGULAR INFERIOR: $\Delta_{ij}^{-} = 0 \quad se \quad i < j$, ou seja, todos os elementos acima da diagonal são NULOS.
5. Matriz SIMÉTRICA: $S_{ij} = S_{ji}$
6. Matriz ANTISIMÉTRICA: $A_{ij} = -A_{ji}$

Note que matrizes representam um conjunto de números, logo é preciso definir as operações sobre esse conjunto.

Operação IGUALDADE: $A = B$ se, e somente se, $A_{ij} = B_{ij} \quad \forall ij$.

Os elementos da diagonal de uma matriz antisimétrica são nulos: $A_{ii} = -A_{ii} \quad \rightarrow \quad 2A_{ii} = 0 \quad \rightarrow \quad A_{ii} = 0$.

Operação ADIÇÃO: essa operação sobre ser efetuada sobre matrizes de mesma dimensões, $A_{m\times n}$ e $B_{m\times n}$, e é definida como $(A + B)_{ij} = A_{ij} + B_{ij}$.

Operação SUBTRAÇÃO: é a inversa da ADIÇÃO. $(A-B)_{ij} = A_{ij} - B_{ij}$

Propriedades da operação ADIÇÃO de matrizes:

a. $A+B=B+A$ pois $A_{ij}+B_{ij}=B_{ij}+A_{ij}$ ou seja é comutativa frente à adição.

b. $A+(B+C)=(A+B)+C$ pois $A_{ij}+(B_{ij}+C_{ij})=(A_{ij}+B_{ij})+C_{ij}$, ou seja, é distributiva.

c. $A+O=A$, a matriz nula é o elemento unitário da operação adição de matrizes.

Operação multiplicação por um escalar: $(aA)_{ij} = aA_{ij}$, ou seja, todos os elementos da matriz são multiplicados pelo escalar (número) a.

Propriedades da operação multiplicação por escalar:

1. $k(A+B)=kA+kB$ distributiva
2. $(k_1+k_2)A=k_1A+k_2A$
3. $0A=O$ é uma matriz nula
4. $k_1(k_2A)=(k_1k_2)A$

Operação **TRANSPOSTA** é denotada por $\overline{A}=A'$, e é definida por $\overline{A}_{ij}=A_{ji}$. Ou seja, essa operação troca linhas por colunas. Gosto da notação $\overline{A}$ porque podemos usá-la facilmente em equações como $\overline{A+B}$ que teria que usar parênteses na notação A', ou seja, $(A+B)'$. Além disso a aplicação da operação transposta duas vezes pode ser escrita como $\overline{\overline{A}}$ e na outra como A''. Matriz simétrica tem a propriedade: $\overline{S}=S$. Note que nesse caso a matriz S é obrigatoriamente quadrada.

Propriedades da operação Transposta:

1. $\overline{\overline{A}}=A$, pois $\overline{A_{ij}}=A_{ji}$ logo $\left(\overline{\overline{A}}\right)_{ij}=\overline{A_{ji}}=A_{ij}$.
2. $\overline{A+B}=\overline{A}+\overline{B}$, pois $\left(\overline{A+B}\right)_{ij}=(A+B)_{ji}=A_{ji}+B_{ji}=\left(\overline{A}\right)_{ij}+\left(\overline{B}\right)_{ij}=\left(\overline{A}+\overline{B}\right)_{ij}$.
3. $\overline{kA}=k\,\overline{A}$
4. Se S é simétrica então $\overline{S}=S$, pois $\left(\overline{S}\right)_{ij}=S_{ji}=S_{ij}$
5. Se A é antisimétrica então $\overline{A}=-A$, pois $\left(\overline{A}\right)_{ij}=A_{ji}=-A_{ij}$

6. $B+\overline{B}=S$ é simétrica. $\overline{\left(B+\overline{B}\right)}=\overline{B}+\overline{\overline{B}}=\overline{B}+B=B+\overline{B}$, logo $\overline{\left(B+\overline{B}\right)}=\left(B+\overline{B}\right)$ é simétrica.

7. $B-\overline{B}=A$ é antisimétrica. $\overline{\left(B+\overline{B}\right)}=\overline{B}-\overline{\overline{B}}=\overline{B}-B=-\left(B-\overline{B}\right)$, logo $\overline{\left(B-\overline{B}\right)}=-\left(B-\overline{B}\right)$ é antisimétrica.

8. Toda matriz quadrada pode ser decomposta em uma matriz simétrica e uma antisimétrica. Prova:

Prova: $M=\frac{1}{2}M+\frac{1}{2}M=\frac{1}{2}M+\frac{1}{2}\overline{M}-\frac{1}{2}\overline{M}+\frac{1}{2}M=\frac{1}{2}\left(M+\overline{M}\right)+\frac{1}{2}\left(M-\overline{M}\right)$ logo $M=S+A$ onde $S=\frac{1}{2}\left(M+\overline{M}\right)$ e $A=\frac{1}{2}\left(M-\overline{M}\right)$

Operação MULTIPLICAÇÃO de matrizes:

Motivação. Suponha uma matriz que fornece as quantidades de vitaminas A, B e C nos alimentos I e II, do tipo:

$$\begin{array}{rc}
\text{vitamina} & \begin{matrix} A & B & C \\ \downarrow & \downarrow & \downarrow \end{matrix} \\
\begin{matrix} \text{alimento}\,I \rightarrow \\ \text{alimento}\,II \rightarrow \end{matrix} & \begin{pmatrix} 4 & 3 & 0 \\ 5 & 0 & 1 \end{pmatrix}
\end{array}$$

Quanto existe de cada vitamina em um prato com 5 unidades do alimento I e 2 do alimento II?

$\begin{pmatrix} 5 & 2 \end{pmatrix}\begin{pmatrix} 4 & 3 & 0 \\ 5 & 0 & 1 \end{pmatrix}=\begin{pmatrix} 5\times4+2\times5 & 5\times3+2\times0 & 5\times0+2\times1 \end{pmatrix}=\begin{pmatrix} 30 & 15 & 2 \end{pmatrix}$. Note que multiplicamos linha por coluna para obter a quantidade de cada uma das vitaminas em um prato combinando diferentes alimentos.

Só podemos, então, multiplicar duas matrizes A e B se o número de colunas de A for igual ao número de linhas de B. A operação **MULTIPLICAÇÃO** entre uma matriz de ordem $m\times n$ e outra de ordem $n\times p$ é definida por $(AB)_{ij}=\sum_{k=1}^{n}A_{ik}B_{kj}$, resultando em uma matriz de ordem $m\times p$. Note que o índice repetido da primeira matriz é o segundo, da coluna, e da segunda é o primeiro, o da linha. A regra para saber a dimensão da matriz multiplicação é cancelar a dimensão repetida, ou seja: $A_{m\times\underline{n}}B_{\underline{n}\times p}=C_{m\times p}$, ou seja, o n repetido desapareceu.

Propriedades da operação **MULTIPLICAÇÃO** de matrizes:

a. Em geral $AB\neq BA$, ou seja, a operação multiplicação não comuta. Existem que casos em que a operação AB está definida mas a operação BA não. Exemplo: $A_{m\times n}B_{n\times p}$ está definida pois o número n de colunas de Aé igual ao de linhas de B, mas a operação $B_{n\times p}A_{m\times n}$ não se $p\neq m$. Para que as duas operações AB e BA sejam definidas é necessário que se A é $m\times n$ então Bé $n\times m$. Mesmo nesse caso os produtos $A_{m\times n}B_{n\times m}=C_{m\times m}$ e $B_{n\times m}A_{m\times n}=D_{n\times n}$, sequer possuem a mesma dimensão., logo não podem ser iguais. Se $m=n$ então as duas matrizes terão as mesmas dimensões, mas ainda assim os produtos AB e BA, genericamente, são diferentes. Podem ser iguais apenas em certas condições especiais. Note a diferença entre os dois produtos: $(AB)_{ij}=\sum_k A_{ik}B_{kj}$ e $(BA)_{ij}=\sum_k A_{kj}B_{ik}$

Propriedades da operação MULTIPLICAÇÃO.

1. Na álgebra de números reais sabemos que se $ab=0$, então $a=0$, ou $b=0$, ou ainda $a=0$ e $b=0$. No entanto no produto de matrizes é possível que $AB=O$ e $A\neq O$ e $B\neq O$. Veja o exemplo: $A=\begin{pmatrix}1&-1&1\\-3&2&-1\\-2&1&0\end{pmatrix}\neq O$ e $B=\begin{pmatrix}1&2&3\\2&4&6\\1&2&3\end{pmatrix}\neq O$ então:

$$AB=\begin{pmatrix}1-2+1&2-4+1&3-6+3\\-3+4-1&-6+8-2&-9+12-3\\-2+2&-4+4&-6+6\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}.$$

2. Multiplicação pela matriz identidade: $AI=IA=A$, ou seja, a matriz identidade é o elemento unitário frente à operação multiplicação de matrizes. Prova: $(AI)_{ij}=\sum_k A_{ik}I_{kj}=\sum_k A_{ik}\delta_{kj}=A_{ij}=(A)_{ij}$, logo $AI=A$. Por outro lado $(IA)_{ij}=\sum_k I_{ik}A_{kj}=\sum_k \delta_{ik}A_{kj}=A_{ij}=(A)_{ij}$, logo $IA=A$.
3. $A(B+C)=AB+AC$ é distributiva à esquerda.
4. $(A+B)C=AC+BC$ é distributiva à direita.

Prova:

$$\left[A(B+C)\right]_{ij}=\sum_k A_{ik}(B+C)_{kj}=\sum_k A_{ik}\left(B_{kj}+C_{kj}\right)=$$
$$=\sum_k A_{ik}B_{kj}+\sum_k A_{ik}C_{kj}=(AB)_{ij}+(AC)_{ij}$$

$$\left[(A+B)C\right]_{ij}=\sum_k (A+B)_{ik}C_{kj}=\sum_k\left(A_{ik}+B_{kj}\right)C_{kj}=$$
$$=\sum_k A_{ik}C_{kj}+\sum_k B_{ik}C_{kj}=(AC)_{ij}+(BC)_{ij}$$

5. $A(BC)=(AB)C$, é associativa.

$$\left[A(BC)\right]_{ij}=\sum_k A_{ik}(BC)_{kj}=\sum_k A_{ik}\sum_l B_{kl}C_{lj}=$$

$$=\sum_k\sum_l A_{ik}B_{kl}C_{lj}=\sum_l\left[\sum_k A_{ik}B_{kl}\right]C_{lj}=\sum_l(AB)_{il}C_{lj}=\left[(AB)C\right]_{ij}$$

6. $OA=O$ e $AO=O$
7. $\overline{AB}=\overline{B}\,\overline{A}$, na transposta do produto é preciso comutar a multiplicação.

Prova: $\left[\overline{AB}\right]_{ij}=\left[AB\right]_{ji}=\sum_k A_{jk}B_{ki}=\sum_k\left(\overline{B}\right)_{ik}\left(\overline{A}\right)_{kj}=\left[\overline{B}\,\overline{A}\right]_{ij}$

8. Se Q é uma matriz quadrada então $Q\overline{Q}=S$ é simétrica. Prova: $\overline{Q\overline{Q}}=\overline{\overline{Q}}\,\overline{Q}=Q\overline{Q}$ logo $\overline{\left(Q\overline{Q}\right)}=\left(Q\overline{Q}\right)$, logo $Q\overline{Q}=S$ é simétrica.

9. A multiplicação de duas matrizes triangulares superiores é triangular superior.

10. A multiplicação de duas matrizes triangulares inferiores é triangular inferior.

Matriz triangular superior é tal que $\Delta^+_{ij}=0 \quad se \quad i>j$. Nesse caso, se $i>j$ então:

$$\left(\Delta^+T^+\right)_{ij}=\sum_k\Delta^+_{ik}T^+_{kj}=\sum_{k=1}^{j}\left[\Delta^+_{ik}\right]T^+_{kj}+\sum_{k=j+1}^{n}\Delta^+_{ik}\left[T^+_{kj}\right]=\sum_{k=1}^{j}0T^+_{kj}+\sum_{k=j+1}^{n}\Delta^+_{ik}0$$

Já se $i<j$ então:

$$\left(\Delta^+T^+\right)_{ij}=\sum_k\Delta^+_{ik}T^+_{kj}=\sum_{k=1}^{i}\left[\Delta^+_{ik}\right]T^+_{kj}+\sum_{k=i+1}^{j}\Delta^+_{ik}\left[T^+_{kj}\right]+\sum_{k=i+1}^{j}\Delta^+_{ik}\left[T^+_{kj}\right]=$$

$$=\sum_{k=i+1}^{j}\Delta^+_{ik}\left[T^+_{kj}\right]\neq 0$$

Logo trata-se de uma matriz triangular superior.

Matriz triangular inferior é tal que $\Delta^-_{ij}=0 \quad se \quad i<j$, nesse caso se $i<j$ então:

$$\left(\Delta^-T^-\right)_{ij}=\sum_k\Delta^-_{ik}T^-_{kj}=\sum_{k=1}^{i}\Delta^-_{ik}\left[T^-_{kj}\right]+\sum_{k=i+1}^{n}\left[\Delta^-_{ik}\right]T^-_{kj}=\sum_{k=1}^{i}\Delta^-_{ik}0+\sum_{k=i+1}^{n}0T^-_{kj}=0$$

Já se $i>j$ então

$$\left(\Delta^{-}T^{-}\right)_{ij}=\sum_{k}\Delta_{ik}^{-}T_{kj}^{-}=\sum_{k=1}^{j}\Delta_{ik}^{-}T_{kj}^{-}+\sum_{k=j+1}^{i}\Delta_{ik}^{-}T_{kj}^{-}+\sum_{k=i+1}^{n}\Delta_{ik}^{-}T_{kj}^{-}=\sum_{k=j+1}^{i}\Delta_{ik}^{-}T_{kj}^{-}\neq 0$$

Logo trata-se de uma matriz triangular inferior.

VETORES.

Vetores são matrizes linha ou coluna que denotamos por $\vec{A}'=\begin{pmatrix}A_1 & A_2 & \cdots & A_n\end{pmatrix}$ ou $\vec{B}=\begin{pmatrix}B_1\\ B_2\\ \vdots\\ B_n\end{pmatrix}$.

Produto escalar ou produto interno. Vamos multiplicar a matriz linha $\vec{A}'=\begin{pmatrix}A_1 & A_2 & \cdots & A_n\end{pmatrix}$ pela matriz coluna $\vec{B}=\begin{pmatrix}B_1\\ B_2\\ \vdots\\ B_n\end{pmatrix}$. O resultado $\vec{A}'\vec{B}=\begin{pmatrix}A_1 & A_2 & \cdots & A_n\end{pmatrix}\begin{pmatrix}B_1\\ B_2\\ \vdots\\ B_n\end{pmatrix}=\sum_i A_iB_i$ se chama PRODUTO ESCALAR e é denotado por $\vec{A}\cdot\vec{B}$. Note que se $\vec{B}=\vec{A}$ então $\vec{A}\cdot\vec{A}=A_1^2+A_2^2+\cdots+A_n^2$ e chamamos a norma de $\vec{A}$ $\left\|\vec{A}\right\|=\sqrt{\vec{A}\cdot\vec{A}}=\sqrt{A_1^2+A_2^2+\cdots+A_n^2}$. A interpretação geométrica do produto escalar pode ser obtida, sem perda de generalidade, tomando o eixo x na direção de $\vec{A}$ e colocando o vetor $\vec{B}$ no plano xy fazendo um ângulo θ com o eixo x. Nesse caso podemos usar apenas duas dimensões e os vetores são dados por $\vec{A}=(A,0)$ e $\vec{B}=(B\cos\theta,B\sin\theta)$. O produto escalar será $\vec{A}\cdot\vec{B}=AB\cos\theta$. Dois vetores não nulos para os quais $\vec{A}\cdot\vec{B}=0$ são perpendiculares entre si. Note que $\cos\theta=\dfrac{\vec{A}\cdot\vec{B}}{\sqrt{\vec{A}\cdot\vec{A}}\sqrt{\vec{B}\cdot\vec{B}}}$ é um número que varia entre $+1$ e -1.

Desigualdade de Schwartz no produto escalar:

$\left(\lambda\vec{A}-\vec{B}\right)\cdot\left(\lambda\vec{A}-\vec{B}\right)\geq 0 \quad \forall\lambda$ pois é o produto escalar de um vetor por ele mesmo. Mas $\left(\lambda\vec{A}-\vec{B}\right)\cdot\left(\lambda\vec{A}-\vec{B}\right)=\lambda^2\vec{A}\cdot\vec{A}-2\lambda\vec{A}\cdot\vec{B}+\vec{B}\cdot\vec{B}\geq 0$. Aqui usamos as propriedades distibutiva $\vec{A}\cdot\left(\vec{B}+\vec{C}\right)=\vec{A}\cdot\vec{B}+\vec{A}\cdot\vec{C}$ e comutativa $\vec{A}\cdot\vec{B}=\vec{B}\cdot\vec{A}$ do produto escalar. Mas a equação $\left(\vec{A}\cdot\vec{A}\right)\lambda^2-2\left(\vec{A}\cdot\vec{B}\right)\lambda+\left(\vec{B}\cdot\vec{B}\right)\geq 0$ é uma parábola em λ do tipo $a\lambda^2+b\lambda+c$ com o coeficiente a positivo, pois $a=\left(\vec{A}\cdot\vec{A}\right)>0$ e a desigualdade $a\lambda^2+b\lambda+c>0$ só será válida se $b^2-4ac<0$, ou sjea não existem raízes reais. Nesse caso, $4\left(\vec{A}\cdot\vec{B}\right)^2-4\left(\vec{A}\cdot\vec{A}\right)\left(\vec{B}\cdot\vec{B}\right)\leq 0$ que implica em $\dfrac{\left(\vec{A}\cdot\vec{B}\right)^2}{\left(\vec{A}\cdot\vec{A}\right)\left(\vec{B}\cdot\vec{B}\right)}\leq 1$ e $-1\leq\dfrac{\vec{A}\cdot\vec{B}}{\sqrt{\vec{A}\cdot\vec{A}}\sqrt{\vec{B}\cdot\vec{B}}}\leq +1$.

Se $\vec{A}$ é um vetor complexo então precisamos não apenas transpor o vetor coluna mas conjugá-lo para garantir um número real positivo após a operação. Chamamos de adjunto do vetor $\vec{A}=\begin{pmatrix}A_1\\A_2\\\vdots\\A_n\end{pmatrix}$ o vetor linha $\vec{A}^\dagger=\vec{A}'^*=\begin{pmatrix}A_1^* & A_2^* & \cdots & A_n^*\end{pmatrix}$, ou seja, é necessário transpor e conjugar. Agora sim, o produto $\vec{A}^\dagger\vec{A}=\begin{pmatrix}A_1^* & A_2^* & \cdots & A_n^*\end{pmatrix}\begin{pmatrix}A_1\\A_2\\\vdots\\A_n\end{pmatrix}=\sum_i A_i^*A_i=\sum_i\left|A_i\right|^2$ será um número real positivo e pode ser usado na definição da norma, ou módulo, de um vetor complexo.

Funções de Matrizes quadradas:

A nossa álgebra de matrizes nos permiet apenas realizar as operações adição, multiplicação por um escalar e multiplicação de matrizes. A adição só pode ser realizada com matrizes de mesma dimensão e a multiplicação exige condições sobre as dimensão das duas matrizes. Se trabalharmos apenas com matrizes quadradas tanto as condições da adição quanto as da multiplicação serão automaticamente satisfeitas. Sabendo a operação multiplicação podemos definir a operação potenciação de matrizes: $A^n=A\times A\times\cdots\times A$, sem problemas com matrizes quadradas.

Faz sentido falar de uma função de matrizes? Por exemplo: $\exp(A)$, ou $\cos(A)$? Sabemos calcular essas funções para um argumento real apenas. No entanto sabemos calcular a função $f(A)=A^n$ e isso é o suficiente para definirmos uma função de matrizes desde que essa função possa ser expandida em série de Taylor-McLaurin, ou seja, seja de classe C^∞ e com raio de convergência infinito. Dentre essas funções estão a exponencial, o seno, o coseno, etc. Assim, considere a série de Taylor da função $f(x)=\sum\limits_{n=0}^{\infty}\dfrac{f^{(n)}(0)}{n!}x^n$.

Note que $\dfrac{f^{(n)}(0)}{n!}$ são apenas escalares (números). Assim definimos a função de matrizes através da série:

$$f(A)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}A^n$$

Sabemos realizar essa operação? Sim, sabemos, pois sabemos calcular A^n, sabemos multiplicá-la por um escalar $\frac{f^{(n)}(0)}{n!}$, e sabemos somar as matrizes:

$$f(A)=f(0)I+f'(0)A+\frac{f''(0)}{2!}A^2+\frac{f^{(3)}(0)}{3!}A^3+\cdots+\frac{f^{(n)}(0)}{n!}A^n.$$

1. Se D é uma matriz diagonal na forma:

$$D=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$$

Então $f(D)=\begin{pmatrix}f(\lambda_1) & 0 & \cdots & 0\\ 0 & f(\lambda_2) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & f(\lambda_n)\end{pmatrix}$.

Para tanto basta notar que $D^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{pmatrix}$, portanto

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} D^n = \begin{pmatrix} \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}\lambda_1^n & 0 & \cdots & 0 \\ 0 & \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}\lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}\lambda_n^n \end{pmatrix}$$ logo

$f(D) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}$. Uma forma de mostrar que $D^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^n \end{pmatrix}$ é usando indução com a seguinte notação: $D_{ij} = \lambda_i \delta_{ij}$. Daí mostramos que $\left(D^2\right)_{ij} = \sum_k D_{ik} D_{kj} = \sum_k \lambda_i \lambda_k \delta_{ik} \delta_{kj} = \lambda_i^2 \delta_{ij}$, e que $\left(D^3\right)_{ij} = \sum_k D_{ik} D_{kj}^2 = \sum_k \lambda_i \lambda_k^2 \delta_{ik} \delta_{kj} = \lambda_i^3 \delta_{ij}$. Supomos verdadeiro para n, isto é, que $\left(D^n\right)_{ij} = \lambda_i^n \delta_{ij}$ e mostramos que é verdadeiro para $n+1$, ou seja $\left(D^{n+1}\right)_{ij} = \sum_k D_{ik} D_{kj}^n = \sum_k \lambda_i \lambda_k^n \delta_{ik} \delta_{kj} = \lambda_i^{n+1} \delta_{ij}$. Assim está provado que $\left(D^n\right)_{ij} = \lambda_i^n \delta_{ij}$ é sempre verdadeiro.

Matrizes Periódica, idempotente e nihilpotente.

Uma matriz quadrada A é **PERIÓDICA** com período $p = n-1,\ n \geq 2$, se $A^n = A$, onde $A^n = A A \cdots A$ é a matriz Amultiplicada por si própria n vezes.

Uma matriz periódica com período 1, i.e., $A^2 = A$, é chamada **IDEMPOTENTE**.

Uma matriz quadrada Aé **NIHILPOTENTE** de índice p se p é o menor inteiro tal que $A^p = O$.

a. Se A é periódica com período p então $A^{n+p+1}=A^n$, $n>0$.

b. Se A é idempotente então $A^n=A$, $n>1$.

c. Se A é nihilpotente de índice p então $A^{p+n}=O$, $n>0$.

Determinantes:

Determinantes são uma função que associa um número real à matrizes quadradas de números reais na forma $\det(A): m\times m \to \mathbb{R}$. A melhor forma de definir determinante é através dos tensores de Levi-Civita e por isso vamos começar com definições e estabelecer as propriedades desses tensores.

Definições.

A função sinal (sgn) é definida por $\operatorname{sgn}(x)=\begin{cases}1 & se \quad x>0\\ 0 & se \quad x=0\\ -1 & se \quad x<0\end{cases}$.

O produtório é representado por $\prod_{j=1}^{n}\lambda_j=\lambda_1\,\lambda_2\,\lambda_3\cdots\lambda_n$

Propriedade da função sinal: $\operatorname{sgn}(xy)=\operatorname{sgn}(x)\operatorname{sgn}(y)$ ou, generalizando, $\operatorname{sgn}\left(\prod_i x_i\right)=\prod_i \operatorname{sgn}(x_i)$. A operação produtório entra e sai do argumento da função sinal.

O Tensor de Levi-Civita é definido por: $\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)=\operatorname{sgn}\left[\prod_{1\le j<k\le n}(\lambda_k-\lambda_j)\right]$, e $\varepsilon(\lambda_1)=1$ quando houver apenas uma variável. Como $\varepsilon=\operatorname{sgn}(z)$ então ε só pode assumnir os valores $+1,\ 0\ e\ -1$.

Esclarecendo a notação. Note que:

$$\prod_{1\le j<k\le n}(\lambda_k-\lambda_j)=(\lambda_n-\lambda_{n-1})(\lambda_n-\lambda_{n-2})\cdots(\lambda_n-\lambda_1)(\lambda_{n-1}-\lambda_{n-2})\cdots(\lambda_{n-1}-\lambda_1)\cdots(\lambda_2-\lambda_1)$$

mantendo sempre $j<k$. Não podemos usar $j=k$ porque $\lambda_k-\lambda_j=0$ e o produtório inteiro seria nulo.

O número de termos no produtório $\prod_{1\le j<k\le n}(\lambda_k-\lambda_j)$ é $\dfrac{n(n-1)}{2}$. Para ver isso começamos de:

$$k=n \quad j<k\to \quad (\lambda_n-\lambda_{n-1})(\lambda_n-\lambda_{n-2})\cdots(\lambda_n-\lambda_1) \quad \to (n-1)\ termos$$
$$k=n-1 \quad j<k\to (\lambda_{n-1}-\lambda_{n-2})(\lambda_{n-1}-\lambda_{n-3})\cdots(\lambda_{n-1}-\lambda_1)\to (n-2)\ termos$$
$$k=2 \quad j<k\to \quad (\lambda_2-\lambda_1)\to \quad 1 \quad termo$$

Então devemos somar a PA $S = 1 + 2 + \cdots + (n-1) = \frac{(n-1)(n-1+1)}{2} = \frac{n(n-1)}{2}$.

Por definição se houver apenas um λ então $\varepsilon(\lambda) = 1$.

Já se existirem dois então $\varepsilon(\lambda_1, \lambda_2) = \operatorname{sgn}(\lambda_2 - \lambda_1)$.

Para $n = 3$ temos que $\varepsilon(\lambda_1, \lambda_2, \lambda_3) = \operatorname{sgn}\left[(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)\right]$.

Para $n = 4$ temos:

$$\varepsilon(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \operatorname{sgn}\left[(\lambda_4 - \lambda_3)(\lambda_4 - \lambda_2)(\lambda_4 - \lambda_1)(\lambda_3 - \lambda_2)(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)\right].$$

Exemplos de cálculos de $\varepsilon' s$:

1. $\varepsilon(4,5,1,2) = \operatorname{sgn}\left[\underset{+}{(2-1)}\underset{-}{(2-5)}\underset{-}{(2-4)}\underset{-}{(1-5)}\underset{-}{(1-4)}\underset{+}{(5-4)}\right] = (-1)^4 = +1$
2. $\varepsilon(2,3,1,2) = \operatorname{sgn}\left[\underset{+}{(2-1)}\underset{-}{(2-3)}\underset{0}{(2-2)}\underset{-}{(1-3)}\underset{-}{(1-2)}\underset{+}{(3-2)}\right] = 0$
3. $\varepsilon(1,2,3,4) = \operatorname{sgn}\left[\underset{+}{(4-3)}\underset{+}{(4-2)}\underset{+}{(4-1)}\underset{+}{(3-2)}\underset{+}{(3-1)}\underset{+}{(2-1)}\right] = +1$
4. $\varepsilon(4,3,2,1) = \operatorname{sgn}\left[\underset{-}{(1-2)}\underset{-}{(1-3)}\underset{-}{(1-4)}\underset{-}{(2-3)}\underset{-}{(2-4)}\underset{-}{(3-4)}\right] = (-1)^6 = +1$
5. $\varepsilon(3,2,1) = \operatorname{sgn}\left[\underset{-}{(1-2)}\underset{-}{(1-3)}\underset{-}{(2-3)}\right] = (-1)^3 = -1$

Note no caso 2 que se dois valores quaisquer são iguais o tensor se anula porque dois valores, o primeiro e o último, foram iguais (2).

Inversões: existe uma inversão sempre que $\lambda_j > \lambda_k$ e $j < k$. Cada inversão traz um sinal negativo. Se o conjunto $(\lambda_1, \lambda_2, \cdots, \lambda_n)$ com $\lambda_i \neq \lambda_j \quad \forall i \neq j$ possui p inversões então $\varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) = (-1)^p$. Logo ε será $+1$ se p for par e ε será -1 se p for ímpar. Foi a contando o número de inversões que calculamos

$$\varepsilon(4,5,1,2) = \operatorname{sgn}\left[\underset{+}{(2-1)}\underset{-}{(2-5)}\underset{-}{(2-4)}\underset{-}{(1-5)}\underset{-}{(1-4)}\underset{+}{(5-4)}\right] = (-1)^4 = +1$$

Tem 4 inversões. Não é necessário escrever todo o produtório para contar o número de inversões. Tome o exemplo da seqüência:

$$\begin{array}{ccccc} 2 & 5 & 4 & 1 & 3 \\ \downarrow & \downarrow & \downarrow & & \\ 1 & 3 & 2 & & \end{array}$$

Tem um total de $1+3+2=6$ inversões. As inversões do 2 foram o 1, o único número à direita menor do que 2. As inversões do 5 foram 4, 1 e 3, e do 4 foram 1 e 3. Com 6 inversões sabemos que $\varepsilon(2,5,4,1,3)=+1$. Contando as inversões podemos calcular o ε rapidamente. Por exemplo: $\varepsilon_{1762354}=+1$ pois a seqüência $1\underset{5}{7}\underset{4}{6}2\underset{1}{3}54$ tem $5+4+1=10$ inversões. Já $\varepsilon_{25314}=-1$ pois a seqüência $\underset{1}{2}\underset{3}{5}\underset{1}{3}14$ tem $1+3+1=5$ inversões.

Arranjos de $(\lambda_1,\lambda_2,\cdots,\lambda_n)$:

Percebemos então que ε só não é nulo se todos os $\lambda's$ forem diferentes. Escolhido um conjunto qualquer $(\lambda_1,\lambda_2,\cdots,\lambda_n)$ de $\lambda's$ diferentes entre si de quantas formas podemos ordenar esse conjunto? Trata-se do problema de colocar n objetos distintos em n caixas em que a ordem dos objetos importa. Assim existem:

$$\frac{n}{1}\frac{(n-1)}{2}\frac{(n-2)}{3}\cdots\frac{1}{n}=n(n-1)(n-2)\cdots 1=n!$$

formas de reordenar esse conjunto.

Exemplos:

1. $n=2 \quad \to (1,2)\, e\, (2,1)=2\ formas$, e $2!=2$.
2. $n=3$ existem $3!=6\ formas$ mostradas abaixo:

Total 6 arranjos $(1\ \underset{+}{2}\ 3);(1\ \underset{-}{3}\ 2);(2\ \underset{-}{1}\ 3);(2\ \underset{+}{3}\ 1);(3\ \underset{+}{1}\ 2);(3\ \underset{-}{2}\ 1)$

Propriedades do tensor de Levi-Civita:

1. $\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)=0 \quad se \quad \lambda_i=\lambda_j \quad \forall i\neq j$. Corolário: se $\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)\neq 0$ então $\lambda_i\neq\lambda_j \quad \forall i\neq j$.
2. $\varepsilon(1,2,\cdots,n)=+1$.
3. Se $\lambda_1<\lambda_2<\cdots<\lambda_n$, isto é, $\lambda_i<\lambda_j \quad \forall i<j$ então $\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)=+1$. Nesse caso $(\lambda_k-\lambda_i)_{k>i}>0$ logo $\left[\prod_{i<k}(\lambda_k-\lambda_i)\right]>0$ e $\operatorname{sgn}\left[\prod_{i<k}(\lambda_k-\lambda_i)\right]=+1$.
4. $\varepsilon(1,\lambda_2,\lambda_3,\cdots,\lambda_n)=\varepsilon(\lambda_2,\lambda_3,\cdots,\lambda_n)$

5. Se $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, isto é, $\lambda_i > \lambda_j \quad \forall i < j$ então $\varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) = (-1)^{\frac{n(n-1)}{2}}$. Note que nesse caso todos os termos $(\lambda_k - \lambda_i)_{k>i} < 0$ e $\operatorname{sgn}(\lambda_k - \lambda_i)_{k>i} = -1$, como existem $\frac{n(n-1)}{2}$ termos, então $\varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) = (-1)^{\frac{n(n-1)}{2}}$.
6. $\varepsilon(\lambda_1 + p, \lambda_2 + p, \cdots, \lambda_n + p) = \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Basta notar que $\operatorname{sgn}(\lambda_k + p - \lambda_j - p) = \operatorname{sgn}(\lambda_k - \lambda_j)$.
7. $\varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_i, \cdots, \lambda_j, \cdots, \lambda_n) = -\varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_j, \cdots, \lambda_i, \cdots, \lambda_n)$. Aqui estamos supondo $i < j$. Ou seja, a transposição de dois índices troca o sinal de ε.

Prova: como só vamos trocar i por j basta analisar os termos contendo esses dois índices, pois todo restante não muda. No produtório vamos separar os termos de acordo com a ordem $k > j > i$, $j > l > i$, $j > i > m$ e $m < i < j$. Os termos que contém apenas i e j são:

$$\left[\prod_{k>j>i} (\lambda_k - \lambda_j)(\lambda_k - \lambda_i)\right] \times \left[\prod_{j>l>i} (\lambda_j - \lambda_l)(\lambda_l - \lambda_i)\right] \times$$
$$\times \left[\prod_{j>i>m} (\lambda_j - \lambda_m)(\lambda_i - \lambda_m)\right] \times (\lambda_j - \lambda_i)$$

Agora $(\lambda_k - \lambda_j)(\lambda_k - \lambda_i) = (\lambda_k - \lambda_i)(\lambda_k - \lambda_j)$ portanto trocar i por j nada muda nesse termo. Da mesma forma para o termo $(\lambda_j - \lambda_m)(\lambda_i - \lambda_m) = (\lambda_i - \lambda_m)(\lambda_j - \lambda_m)$. Já no termo:

$$(\lambda_j - \lambda_l)(\lambda_l - \lambda_i) = (-1)(\lambda_l - \lambda_j)(-1)(\lambda_i - \lambda_l) = (-1)^2 (\lambda_l - \lambda_j)(\lambda_i - \lambda_l)$$

O sinal foi trocado duas vezes. Portanto, trocar trocar i por j leva a $(\lambda_i - \lambda_l)(\lambda_l - \lambda_j) = (\lambda_j - \lambda_l)(\lambda_l - \lambda_i)$.

O único termo que sobrou foi $(\lambda_j - \lambda_i) = (-1)(\lambda_i - \lambda_j)$, mostrando que qualquer permuta nos índices troca o sinal de ε.

8. Se $(\mu_1 \quad \mu_2 \quad \cdots \quad \mu_n)$ é obtido $(\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n)$ através de p transposições então $\varepsilon(\mu_1 \quad \mu_2 \quad \cdots \quad \mu_n) = (-1)^p \varepsilon(\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n)$, pois cada transposição troca o sinal uma vez. Corolário: Se $(\mu_1 \quad \mu_2 \quad \cdots \quad \mu_n)$ é obtido através de p transposições de $(1 \quad 2 \quad \cdots \quad n)$ então $\varepsilon(\mu_1 \quad \mu_2 \quad \cdots \quad \mu_n) = (-1)^p$.

9. Reordenamento simultâneo: sejam $\begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}$, $\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix}$ e $\begin{pmatrix}k_1 & k_2 & \cdots & k_n\end{pmatrix}$ reordenamentos de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$, logo não existem índices iguais, i.e. $\nu_i \neq \nu_j \quad i \neq j$ é sempre verdade. Então:
$$\varepsilon\begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}\varepsilon\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix} = \varepsilon\begin{pmatrix}\mu_{k_1} & \mu_{k_2} & \cdots & \mu_{k_n}\end{pmatrix}\varepsilon\begin{pmatrix}\lambda_{k_1} & \lambda_{k_2} & \cdots & \lambda_{k_n}\end{pmatrix}.$$

Para provar basta notar que
$$\varepsilon\begin{pmatrix}\lambda_{k_1} & \lambda_{k_2} & \cdots & \lambda_{k_n}\end{pmatrix} = (-1)^p\,\varepsilon\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix}$$
$$\varepsilon\begin{pmatrix}\mu_{k_1} & \mu_{k_2} & \cdots & \mu_{k_n}\end{pmatrix} = (-1)^p\,\varepsilon\begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}$$

Pois $\begin{pmatrix}\lambda_{k_1} & \lambda_{k_2} & \cdots & \lambda_{k_n}\end{pmatrix}$ é um reordenamento de $\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix}$ e $\begin{pmatrix}\mu_{k_1} & \mu_{k_2} & \cdots & \mu_{k_n}\end{pmatrix}$ é o mesmo reordenamento de $\begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}$ logo as permutações são idênticas em ambos os casos. Como $(-1)^{2p} = \left[(-1)^2\right]^p = (+1)^p = +1$ a igualdade está provada.

10. $\varepsilon(1,\lambda_2,\lambda_3,\cdots,\lambda_n) = \varepsilon(\lambda_2,\lambda_3,\cdots,\lambda_n)$ onde $(\lambda_2,\lambda_3,\cdots,\lambda_n)$ é um reordenamento de $(2,3,\cdots,n)$. Todos os termos com $\operatorname{sgn}(\lambda_{i>1}-1) = +1$ e desaparecem do produtório.

Para provar as propriedades do determinante vamos precisar ainda dos teoremas de preenchimento.

Teoremas de preenchimento:

1. Sejam $\begin{pmatrix}\nu_1 & \nu_2 & \cdots & \nu_n\end{pmatrix}$ todos os reordenamentos possíveis de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$, portanto $\nu_i \neq \nu_j \quad i \neq j$ é sempre verdade, e $\begin{pmatrix}k_1 & k_2 & \cdots & k_n\end{pmatrix}$ um reordenamento fixo de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$, então $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ varre todos os reordenamentos possíveis de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$.

Antes de demonstrar o teorema vamos exemplificá-lo com $n=3$, e $\begin{pmatrix}1 & 2 & 3\end{pmatrix}$ e escolher $\begin{pmatrix}k_1 & k_2 & k_3\end{pmatrix} = \begin{pmatrix}2 & 3 & 1\end{pmatrix}$. Agora vamos reordenar $\begin{pmatrix}\nu_1 & \nu_2 & \nu_3\end{pmatrix}$ e ver o que obtemos com $\begin{pmatrix}\nu_2 & \nu_3 & \nu_1\end{pmatrix}$:

$$\begin{pmatrix}\nu_1 & \nu_2 & \nu_3\end{pmatrix}=\begin{bmatrix}\begin{pmatrix}1 & 2 & 3\end{pmatrix}\\\begin{pmatrix}1 & 3 & 2\end{pmatrix}\\\begin{pmatrix}2 & 1 & 3\end{pmatrix}\\\begin{pmatrix}2 & 3 & 1\end{pmatrix}\\\begin{pmatrix}3 & 1 & 2\end{pmatrix}\\\begin{pmatrix}3 & 2 & 1\end{pmatrix}\end{bmatrix}\Rightarrow\begin{pmatrix}\nu_2 & \nu_3 & \nu_1\end{pmatrix}=\begin{bmatrix}\begin{pmatrix}2 & 3 & 1\end{pmatrix}\\\begin{pmatrix}3 & 2 & 1\end{pmatrix}\\\begin{pmatrix}1 & 3 & 2\end{pmatrix}\\\begin{pmatrix}3 & 1 & 2\end{pmatrix}\\\begin{pmatrix}1 & 2 & 3\end{pmatrix}\\\begin{pmatrix}2 & 1 & 3\end{pmatrix}\end{bmatrix}$$

Note que com $\begin{pmatrix}\nu_2 & \nu_3 & \nu_1\end{pmatrix}$ obtivemos de volta os 6 arranjos apenas em uma ordem diferente. Ou seja $\begin{pmatrix}\nu_2 & \nu_3 & \nu_1\end{pmatrix}$ é tão bom quanto $\begin{pmatrix}\nu_1 & \nu_2 & \nu_3\end{pmatrix}$ para varrer todas as possibilidades de arranjos.

Para $\begin{pmatrix}\nu_1 & \nu_2 & \cdots & \nu_n\end{pmatrix}$ exaurir todas as possibilidades de reordenamento ele precisa de $n!$ conjuntos de n elementos, diferentes entre si [os conjuntos]. Se $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ contém $n!$ conjuntos diferentes, obrigatoriamente contém todos os possíveis reordenamentos de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$. Para cada $\begin{pmatrix}\nu_1 & \nu_2 & \cdots & \nu_n\end{pmatrix}$ corresponde um $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$, logo o número de conjuntos de $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ é igual ao de $\begin{pmatrix}\nu_1 & \nu_2 & \cdots & \nu_n\end{pmatrix}$. Na realidade $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ é um reordenamento de $\begin{pmatrix}\nu_1 & \nu_2 & \cdots & \nu_n\end{pmatrix}$. Como o número de conjuntos são iguais entre si e igual a $n!$ então basta provar que não existem $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ repetidos.

Vamos provar por absurdo:

Suponha que os dois conjuntos $\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}\neq\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_n\end{pmatrix}$, diferentes entre si, levam a dois conjuntos $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ repetidos, i.e.:

$$\begin{pmatrix}\alpha_{k_1} & \alpha_{k_2} & \cdots & \alpha_{k_n}\end{pmatrix}=\begin{pmatrix}\beta_{k_1} & \beta_{k_2} & \cdots & \beta_{k_n}\end{pmatrix}$$

Mas isso implica que $\alpha_{k_i}=\beta_{k_i} \quad \forall i$ o que significa que $\alpha_j=\beta_j \quad \forall j$, fazendo $j=k_i$, e, portanto que $\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}=\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_n\end{pmatrix}$ em contradição com a hipótese de que $\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}\neq\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_n\end{pmatrix}$. Logo não há repetição nos conjuntos $\begin{pmatrix}\nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n}\end{pmatrix}$ e eles varrem todos os possíveis reordenamentos de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$.

2. Sejam $\begin{pmatrix}\nu_1 & \nu_2 & \cdots & \nu_n\end{pmatrix}$ todos os reordenamentos possíveis de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$ e $\begin{pmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{pmatrix}$ um reordenamento fixo de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$, então o reordenamento $\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix}$ tal que $\nu_{\lambda_1}=\mu_1, \quad \nu_{\lambda_2}=\mu_2, \quad \cdots \quad ,\nu_{\lambda_n}=\mu_n$ também varre todos os reordenamentos possíveis de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$.

Antes de demonstrar o teorema vamos aplicá-lo para entender a regra de associação entre os diversos índices. Vamos escolher $(\mu_1 \quad \mu_2 \quad \mu_3) = (3 \quad 1 \quad 2)$ e a regra para encontrar o λ é dada por $\nu_{\lambda_1} = 3, \quad \nu_{\lambda_2} = 1, \nu_{\lambda_3} = 2$. Então se $(\nu_1 \quad \nu_2 \quad \nu_3) = (2 \quad 1 \quad 3)$, $\lambda_1 = 3$ porque $3 = \nu_3$, $\lambda_2 = 2$ porque $1 = \nu_2$ e $\lambda_3 = 1$ porque $2 = \nu_1$. Trata-se, portanto, de encontrar a função inversa $\nu(\lambda_i) = \mu_i \to \lambda_i = \nu^{-1}(\mu_i)$.

No exemplo dado, $\nu_{\lambda_1} = 3$, $\nu_{\lambda_2} = 1$ e $\nu_{\lambda_3} = 2$, teremos:

$$(\nu_1 \quad \nu_2 \quad \nu_3) = \begin{bmatrix} (1 \quad 2 \quad 3) \\ (1 \quad 3 \quad 2) \\ (2 \quad 1 \quad 3) \\ (2 \quad 3 \quad 1) \\ (3 \quad 1 \quad 2) \\ (3 \quad 2 \quad 1) \end{bmatrix} \Rightarrow (\lambda_1 \quad \lambda_2 \quad \lambda_3) = \begin{bmatrix} (3 \quad 1 \quad 2) \\ (2 \quad 1 \quad 3) \\ (3 \quad 2 \quad 1) \\ (2 \quad 3 \quad 1) \\ (1 \quad 2 \quad 3) \\ (1 \quad 3 \quad 2) \end{bmatrix}$$

Note que $(\lambda_1 \quad \lambda_2 \quad \lambda_3)$ varreu todas as possibilidades de reordenamento de $(1 \quad 2 \quad 3)$ também.

A prova é semelhante e basta provar que não há repetição, ou seja, se $\left\{\begin{matrix} (\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n) \\ (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n) \end{matrix}\right\} \Rightarrow (\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_n)$ então $\alpha_{\lambda_i} = \mu_i \quad e \quad \beta_{\lambda_i} = \mu_i \quad \Rightarrow \quad \alpha_{\lambda_i} = \beta_{\lambda_i} \quad \Rightarrow \alpha_j = \beta_j$ logo $(\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n) = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_n)$.

com $(\nu_2 \quad \nu_3 \quad \nu_1)$ obtivemos de volta os 6 arranjos apenas em uma ordem diferente. Ou seja $(\nu_2 \quad \nu_3 \quad \nu_1)$ é tão bom quanto $(\nu_1 \quad \nu_2 \quad \nu_3)$ para varrer todas as possibilidades de arranjos.

Para $(\nu_1 \quad \nu_2 \quad \cdots \quad \nu_n)$ exaurir todas as possibilidades de reordenamento ele precisa de $n!$ conjuntos de n elementos, diferentes entre si [os conjuntos]. Se $(\nu_{k_1} \quad \nu_{k_2} \quad \cdots \quad \nu_{k_n})$ contém $n!$ conjuntos diferentes, obrigatoriamente contém todos os possíveis reordenamentos de $(1 \quad 2 \quad \cdots \quad n)$. Para cada $(\nu_1 \quad \nu_2 \quad \cdots \quad \nu_n)$ corresponde um $(\nu_{k_1} \quad \nu_{k_2} \quad \cdots \quad \nu_{k_n})$, logo o número de conjuntos de $(\nu_{k_1} \quad \nu_{k_2} \quad \cdots \quad \nu_{k_n})$ é igual ao de $(\nu_1 \quad \nu_2 \quad \cdots \quad \nu_n)$. Na realidade $(\nu_{k_1} \quad \nu_{k_2} \quad \cdots \quad \nu_{k_n})$ é um reordenamento de $\begin{pmatrix} \nu_1 & \nu_2 & \cdots & \nu_n \end{pmatrix}$. Como o número de conjuntos são iguais entre si e igual a $n!$ então basta provar que não existem $\begin{pmatrix} \nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n} \end{pmatrix}$ repetidos.

Vamos provar por absurdo: Suponha que os dois conjuntos $\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix} \neq \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}$, diferentes entre si, levam a dois conjuntos $\begin{pmatrix} \nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n} \end{pmatrix}$ repetidos, i.e.:

$$\begin{pmatrix} \alpha_{k_1} & \alpha_{k_2} & \cdots & \alpha_{k_n} \end{pmatrix} = \begin{pmatrix} \beta_{k_1} & \beta_{k_2} & \cdots & \beta_{k_n} \end{pmatrix}$$

Mas isso implica que $\alpha_{k_i} = \beta_{k_i} \quad \forall i$ o que significa que $\alpha_j = \beta_j \quad \forall j$, fazendo $j = k_i$, e, portanto que $\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix} = \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}$ em contradição com a hipótese de que $\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix} \neq \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix}$. Logo não há repetição nos conjuntos $\begin{pmatrix} \nu_{k_1} & \nu_{k_2} & \cdots & \nu_{k_n} \end{pmatrix}$ e eles varrem todos os possíveis reordenamentos de $\begin{pmatrix} 1 & 2 & \cdots & n \end{pmatrix}$.

DEFINIÇÃO DE DETERMINANTE:

O determinante é uma função que associa um número à matrizes quadradas, ou seja, é uma função de matrizes, o domínio é o conjunto de matrizes quadradas, e o contradomínio são os reais, definida pela regra:

$$\det(A) = \sum_{\lambda_1=1}^{n} \sum_{\lambda_2=1}^{n} \cdots \sum_{\lambda_n=1}^{n} \varepsilon\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{n\lambda_n}.$$

Também denotado por:

$$\det(A) = \sum_{\lambda_1, \lambda_2, \cdots, \lambda_n} \varepsilon\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{n\lambda_n}$$

Ou ainda na forma mais abreviada:

$$\det(A) = \sum_{\vec{\lambda}} \varepsilon\left(\vec{\lambda}\right) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{n\lambda_n} \text{ com } \vec{\lambda} = \left(\lambda_1, \lambda_2, \cdots, \lambda_n\right)$$

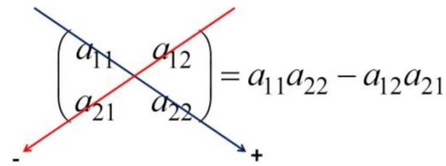
Note que a somatória é feita sobre o segundo índice dos elementos da matriz, que correspondem às colunas. Também se percebe que a propriedade (1) das $\varepsilon's$ nos garante que podemos trabalhar apenas com reordenamentos de $\begin{pmatrix} 1 & 2 & \cdots & n \end{pmatrix}$, evitando assim índices repetidos.

Exemplos: encontrar a forma explícita dos determinantes de matrizes 1×1, 2×2 e 3×3.

1. $\det A_{1\times1} = \sum_{\lambda_1=1}^{1} \varepsilon\left(\lambda_1\right) A_{1\lambda_1} = A_{11}$.

2. $\det A_{2\times 2} = \sum_{\lambda_1=1}^{2}\sum_{\lambda_2=1}^{2} \varepsilon(\lambda_1,\lambda_2) A_{1\lambda_1} A_{2\lambda_2} = \varepsilon(1,2) A_{11}A_{22} + \varepsilon(2,1) A_{12}A_{21}$, ou seja, $\det A_{2\times 2} = A_{11}A_{22} - A_{12}A_{21}$.

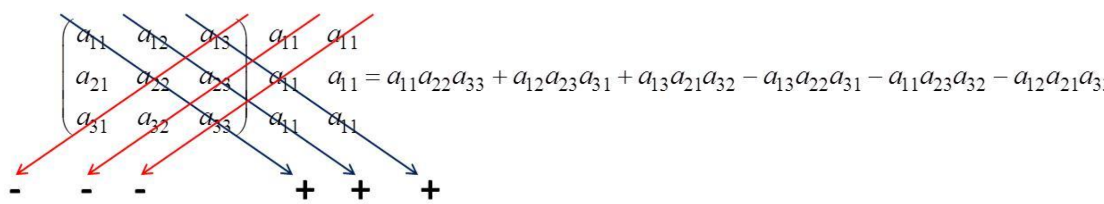
Trata-se da famosa regra:



3. $\det A_{3\times 3} = \sum_{\lambda_1=1}^{3}\sum_{\lambda_2=1}^{3}\sum_{\lambda_3=1}^{3} \varepsilon(\lambda_1,\lambda_2,\lambda_3) A_{1\lambda_1} A_{2\lambda_2} A_{3\lambda_3}$

$$\det A_{3\times 3} = \varepsilon(1,2,3) A_{11}A_{22}A_{33} + \varepsilon(2,3,1) A_{12}A_{23}A_{31} + \varepsilon(3,1,2) A_{13}A_{21}A_{32} +$$
$$+\varepsilon(3,2,1) A_{13}A_{22}A_{31} + \varepsilon(1,3,2) A_{11}A_{23}A_{32} + +\varepsilon(2,1,3) A_{12}A_{21}A_{33}$$

Agora note que $\varepsilon(1,2,3) = \varepsilon(2,3,1) = \varepsilon(3,1,2) = +1$ e que $\varepsilon(3,2,1) = \varepsilon(1,3,2) = \varepsilon(2,1,3) = -1$ e mais ainda que varremos todos os reordenamentos possíveis [devem existir 2 começando com 1, dois começando com 2 e 2 começando com 3]. Trata-se também de outra regra famosa:



O determinante de uma matriz 4×4 terá $4! = 24$ termos e o processo de cálculo vai se tornando tedioso. O EXCEL já tem uma função embutida que faz esse cálculo. Na mão o melhor é calcular usando a regra dos menores que provaremos adiante.

Propriedades dos determinantes:

1. $\det(A) = \sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n) A_{\mu_1\lambda_1} A_{\mu_2\lambda_2} \cdots A_{\mu_n\lambda_n}$

Onde $(\mu_1,\mu_2,\cdots,\mu_n)$ é um reordenamento de $(1 \quad 2 \quad \cdots \quad n)$. Note que a diferença com a definição inicial do determinante $\det(A) = \sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{n\lambda_n}$ foi o reordenamento de $(1 \quad 2 \quad \cdots \quad n)$.

Prova: por definição temos que:

$$\det(A)=\sum_{\nu_1,\nu_2,\cdots,\nu_n}\varepsilon(1,2,\cdots,n)\varepsilon(\nu_1,\nu_2,\cdots,\nu_n)\,A_{1\nu_1}A_{2\nu_2}\cdots A_{n\nu_n}$$

Agora aplicamos dois reordenamentos iguais e simultâneos:

$$(1,2,\cdots,n)\to(\mu_1,\mu_2,\cdots,\mu_n)$$
$$(\nu_1,\nu_2,\cdots,\nu_n)\to(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

e usamos a propriedade (7):

$\varepsilon(1,2,\cdots,n)\varepsilon(\nu_1,\nu_2,\cdots,\nu_n)=\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

Para completar, precisamos usar o teorema de preenchimento para garantir que os $\lambda's$ varrerão todos o reordenamentos possíveis de $\begin{pmatrix}1 & 2 & \cdots & n\end{pmatrix}$ como faziam os $\nu's$ iniciais. Isso garante que a somatória incluirá todos os termos.

2. $\det(A)=\sum_{\mu_1,\mu_2,\cdots,\mu_n}\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)A_{\mu_1\lambda_1}A_{\mu_2\lambda_2}\cdots A_{\mu_n\lambda_n}$. Note que agora a somatória é feita sobre o primeiro índice, ou seja, sobre as linhas.

Prova: novamente começamos de:

$$\det(A)=\sum_{\nu_1,\nu_2,\cdots,\nu_n}\varepsilon(1,2,\cdots,n)\varepsilon(\nu_1,\nu_2,\cdots,\nu_n)\,A_{1\nu_1}A_{2\nu_2}\cdots A_{n\nu_n}$$

O procedimento de reordenação agora será o seguinte – reordenar cada $(\nu_1,\nu_2,\cdots,\nu_n)$ até chegar a um mesmo $(\lambda_1,\lambda_2,\cdots,\lambda_n)$fixo, o que requer que $\nu_{\mu_j}=\lambda_j$. Se aplicamos o mesmo reordenamento para $(1,2,\cdots,n)$ obteremos $(\mu_1,\mu_2,\cdots,\mu_n)$, ou seja, se ν_j virou λ_j então j virou μ_j. Como os reordenamentos são, novamente, simultâneos, vale a propriedade (7) $\varepsilon(1,2,\cdots,n)\varepsilon(\nu_1,\nu_2,\cdots,\nu_n)=\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)$. A diferença agora é que os $\lambda's$ estão fixos e os $\mu's$ variando nos reordenamentos de $(1,2,\cdots,n)$. O teorema 2 do preenchimento garante que essa variação preencherá todos os $n!$ reordenamentos. Assim a propriedade está provada.

3. $\det(\bar{A})=\det(A)$.

$$\det(\bar{A})=\sum_{\mu_1,\mu_2,\cdots,\mu_n}\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)\bar{A}_{\mu_1\lambda_1}\bar{A}_{\mu_2\lambda_2}\cdots\bar{A}_{\mu_n\lambda_n}=$$
$$=\sum_{\mu_1,\mu_2,\cdots,\mu_n}\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)A_{\lambda_1\mu_1}A_{\lambda_2\mu_2}\cdots A_{\lambda_n\mu_n}=\det(A)$$

Usando a propriedade (2). O determinante de uma matriz é igual ao da sua transposta.

4. Se A' é obtida de A multiplicando a iésima linha por α então $\det A' = \alpha \det A$.

A transformação é dada por $A'_{ij} = \alpha A_{ij} \quad e \quad A'_{k\neq ij} = A_{ij}$. Nesse caso

$$\det A' = \sum_{\vec{\lambda}} \varepsilon\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix} A'_{1\lambda_1} A'_{2\lambda_2} \cdots A'_{i\lambda_i} \cdots A'_{n\lambda_n} =$$
$$= \sum_{\vec{\lambda}} \varepsilon\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix} A_{1\lambda_1} A_{2\lambda_2} \cdots \alpha A_{i\lambda_i} \cdots A_{n\lambda_n} =$$
$$= \alpha \sum_{\vec{\lambda}} \varepsilon\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix} A_{1\lambda_1} A_{2\lambda_2} \cdots A_{i\lambda_i} \cdots A_{n\lambda_n} = \alpha \det A$$

4.a. $\det(\alpha A) = \alpha^n \det(A)$

$$\det(\alpha A) = \sum_{\vec{\lambda}} \varepsilon\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_n\end{pmatrix} \alpha A_{1\lambda_1} \alpha A_{2\lambda_2} \cdots \alpha A_{i\lambda_i} \cdots \alpha A_{n\lambda_n} = \alpha^n \det A$$

5. Se A' é obtida de A multiplicando a jésima coluna por β então $\det A' = \beta \det A$.

Nesse caso notamos que $\bar{A}'$ é obtida de $\bar{A}$ multiplicando a jésima linha por β, e como $\det \bar{A}' = \det A'$ então $\det A' = \beta \det A$.

6. Trocar duas linhas, ou duas colunas, de uma matriz troca o sinal do determinante. Para mostra a troca de colunas basta usar o fato de que $\det \bar{A} = \det A$.

$$D = \det(A) = \sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{n\lambda_n}$$

$$\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_i,\cdots,\lambda_j,\cdots,\lambda_n) = -\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_j,\cdots,\lambda_i,\cdots,\lambda_n)$$

$$D' = \sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_i,\cdots,\lambda_j,\cdots,\lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{j\lambda_i} \cdots A_{i\lambda_j} \cdots A_{n\lambda_n} =$$
$$= -\sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_j,\cdots,\lambda_i,\cdots,\lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{j\lambda_i} \cdots A_{i\lambda_j} \cdots A_{n\lambda_n}$$

Chamando $\lambda_j = \lambda'_i$ e $\lambda_i = \lambda'_j$

$$D' = -\sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon(\lambda_1,\lambda_2,\cdots,\lambda'_i,\cdots,\lambda'_j,\cdots,\lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{j\lambda'_j} \cdots A_{i\lambda'_i} \cdots A_{n\lambda_n} = -\det A$$

7. Se uma matriz tem uma linha (ou uma coluna) nula então $\det A = 0$. Multiplica a linha por zero e usa propriedade (4).
8. Se dua linhas (ou colunas) de A são iguais então $\det A = 0$. Se duas linhas i e j são iguais, trocá-las não muda a matriz, i.e., $A' = A$. Por outro lado, pela propriedade (6) então $\det A' = -\det A \quad \to \det A = -\det A \quad \to 2\det A = 0 \quad \to \det A = 0$.

9. E duas linhas (ou duas colunas) de A são proporcionais então $\det A = 0$. Nesse caso $A_{ij} = kA'_{ij}$ onde a matriz A' possui duas linhas iguais, logo $\det A = k \det A' = k \times 0 = 0$.

10. Considere duas matrizes A_1 e A_2 que só diferem na iésima linha, i.e.,

$$A_1 = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \quad \text{e} \quad A'_1 = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A'_{i1} & A'_{i2} & \cdots & A'_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}, \quad \text{então}$$

$$\det A = \det \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} + A'_{i1} & A_{i2} + A'_{i2} & \cdots & A_{in} + A'_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \det A + \det A' .$$

$$\det A = \sum_{\vec{\lambda}} \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots \left(A_{i\lambda_i} + A'_{i\lambda_i} \right) \cdots A_{n\lambda_n} =$$
$$= \sum_{\vec{\lambda}} \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A_{i\lambda_i} \cdots A_{n\lambda_n} + \sum_{\vec{\lambda}} \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) A_{1\lambda_1} A_{2\lambda_2} \cdots A'_{i\lambda_i} \cdots A_{n\lambda_n} =$$
$$= \det A_1 + \det A_2$$

A mesma propriedade vale para duas colunas.

11. Somar à uma linha um múltiplo de outra linha não altera o determinante. Basta usar as propriedades acima: $\det A = \det A_1 + \det A_2$ mas A_2 possui duas linhas proporcionais logo $\det A_2 = 0$. O mesmo vale para colunas.

12. $\det(AB) = \det A \times \det B$. Seja $C = AB$ então $C_{ij} = \sum_k A_{ik} B_{kj}$ e nnn

$$\det C = \sum_{\vec{\lambda}} \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) C_{1\lambda_1} C_{2\lambda_2} \cdots C_{n\lambda_n} =$$

$$= \sum_{\vec{\lambda}} \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) \left(\sum_{\mu_1} A_{1\mu_1} B_{\mu_1\lambda_1} \right) \left(\sum_{\mu_2} A_{2\mu_2} B_{\mu_2\lambda_2} \right) \cdots \left(\sum_{\mu_n} A_{n\mu_n} B_{\mu_n\lambda_n} \right) = \text{n}$$

$$= \sum_{\vec{\mu}} A_{1\mu_1} A_{2\mu_2} \cdots A_{n\mu_n} \sum_{\vec{\lambda}} \varepsilon(\lambda_1, \lambda_2, \cdots, \lambda_n) B_{\mu_1\lambda_1} B_{\mu_2\lambda_2} \cdots B_{\mu_n\lambda_n} =$$

Agora $\varepsilon^2(\mu_1,\mu_2,\cdots,\mu_n)=\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)=+1$ desde que $(\mu_1,\mu_2,\cdots,\mu_n)$ seja um reordenamento de $(1,2,\cdots,n)$. Logo podemos incluí-lo na somatória acima obtendo:

$$\begin{aligned}\det C &= \sum_{\vec{\mu}}\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)A_{1\mu_1}A_{2\mu_2}\cdots A_{n\mu_n}\times\\ &\times\sum_{\vec{\lambda}}\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)B_{\mu_1\lambda_1}B_{\mu_2\lambda_2}\cdots B_{\mu_n\lambda_n}=\\ &=\sum_{\vec{\mu}}\varepsilon(\mu_1,\mu_2,\cdots,\mu_n)A_{1\mu_1}A_{2\mu_2}\cdots A_{n\mu_n}\times\det B=\det A\times\det B\end{aligned}$$

13. $\det(\bar{A}\bar{B})=\det(\bar{B}\bar{A})=\det(A\bar{B})=\det(\bar{B}A)=\det(\bar{A}B)=\det(B\bar{A})=\det A\times\det B$ pois $\det\bar{M}=\det M$ e $\det(MN)=\det M\times\det N$.

14. Se a matriz A é diagonal então $\det(A)=\prod_{i=1}^{n}A_{ii}=A_{11}A_{22}\cdots A_{nn}$.

Se A é diagonal então $A_{ij}=A_i\delta_{ij}$. Nesse caso o determinante é dado por:

$$\det(A)=\sum_{\lambda_1,\lambda_2,\cdots,\lambda_n}\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)A_1\delta_{1\lambda_1}A_2\delta_{2\lambda_2}\cdots A_n\delta_{n\lambda_n}$$

Mas os deltas de Kronecker dentro da somatória garantem que que apenas os termos com $\lambda_1=1,\lambda_2=2,\cdots,\lambda_n=n$ sobrevivem, logo a somatória será $\det(A)=A_1A_2\cdots A_n$, ou seja, a multiplicação dos elementos da diagonal.

15. Se a matriz A é triangular superior, ou inferior, então $\det(A)=\prod_{i=1}^{n}A_{ii}=A_{11}A_{22}\cdots A_{nn}$.

Trata-se de uma propriedade intuitiva porque exceto pela diagonal todos os outros termos serão multiplicados por algum elemento nulo do triângulo inferior. Basta mostrar para a triangular superior, pois a inferior sai com $\det\bar{A}=\det A$. Se A é matriz trinagular superiro então $A_{ij}=0 \quad se \quad i>j$, o que impõe restrições nos $\lambda's$.

$$\det(A)=\sum_{\lambda_1,\lambda_2,\cdots,\lambda_n}\varepsilon(\lambda_1,\lambda_2,\cdots,\lambda_n)A_{1\lambda_1}A_{2\lambda_2}\cdots A_{n\lambda_n}$$

Note que para $A_{1\lambda_1}\neq 0 \quad 0<\lambda_1<2 \quad \rightarrow \lambda_1=1$. Se $\lambda_1=1$ então $\lambda_i>1\,\forall i>1$. O próximo termo é o λ_2, que pela propriedade da matriz deve ser menor do que 3, mas acima de 1, logo $\lambda_2=2$. Continuando o argumento percebemos que a única solução não nula é para $\lambda_1=1,\lambda_2=2,\cdots,\lambda_n=n$ e $\det(A)=A_{11}A_{22}\cdots A_{nn}$.

Essa propriedade (15) é a base do método numérico de triangularização da matriz, usando o fato de que operações elementares não alteram o determinante, mais eficiente para cálculo de determinantes.

Definição do COFATOR.

Vamos obter a matriz Δ_{ij} à partir da matriz $A_{n\times n}$ extraindo a iésima linha e a jésima coluna. A matriz cofator ij de A é definida como: $cof\,(A)_{ij} = (-1)^{i+j}\det\left(\Delta_{ij}\right)$.

Exemplo, seja $A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}$.

$$\Delta_{11} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad \det\Delta_{11} = 1 \quad cofA_{11} = 1$$

$$\Delta_{12} = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix} \quad \det\Delta_{12} = 4 \quad cofA_{12} = -4$$

$$\Delta_{13} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} \quad \det\Delta_{13} = 1 \quad cofA_{13} = 1$$

$$\Delta_{21} = \begin{pmatrix} 0 & 2 \\ 1 & 2 \end{pmatrix} \quad \det\Delta_{21} = -2 \quad cofA_{21} = 2$$

$$\Delta_{22} = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix} \quad \det\Delta_{22} = -2 \quad cofA_{22} = -2$$

$$\Delta_{23} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \quad \det\Delta_{23} = 1 \quad cofA_{23} = -1$$

$$\Delta_{31} = \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} \quad \det\Delta_{31} = -2 \quad cofA_{31} = -2$$

$$\Delta_{32} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \quad \det\Delta_{32} = -5 \quad cofA_{32} = 5$$

$$\Delta_{33} = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \quad \det\Delta_{33} = 1 \quad cofA_{33} = 1$$

$$cof\left(A\right) = \begin{pmatrix} 1 & -4 & 1 \\ 2 & -2 & -1 \\ -2 & 5 & 1 \end{pmatrix}$$

Propriedades do cofator:

1. $\det A = \det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \det\Delta_{11} = \det\begin{pmatrix} A_{22} & A_{23} & \cdots & A_{2n} \\ A_{32} & A_{33} & \cdots & A_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n2} & A_{n3} & \cdots & A_{nn} \end{pmatrix}$, ou seja, iniciamos com uma matriz $n \times n$ e caímos em uma matriz $\left(n-1\right)\times\left(n-1\right)$. Note que os elementos da primeira linha podem ser escritos como $A_{1\lambda_1} = \delta_{1\lambda_1}$, anulando todos os elementos com $\lambda_1 \neq 1$.

$$\begin{aligned}
\det(A) &= \sum_{\lambda_1,\lambda_2,\cdots,\lambda_n} \varepsilon\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)\delta_{1\lambda_1}A_{2\lambda_2}\cdots A_{n\lambda_n} = \\
&= \sum_{\left(\lambda_2,\cdots,\lambda_n\right)>1} \varepsilon\left(1,\lambda_2,\cdots,\lambda_n\right)A_{2\lambda_2}\cdots A_{n\lambda_n} = \\
&= \sum_{\left(\lambda_2,\cdots,\lambda_n\right)>1} \varepsilon\left(\lambda_2,\cdots,\lambda_n\right)A_{2\lambda_2}\cdots A_{n\lambda_n} = \det\Delta_{11}
\end{aligned}$$

Expansão nos cofatores:

$$\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = \det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} + \det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & A_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} + \cdots +$$

$$+\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

$$\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{in} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} = A_{i1}\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} + A_{i2}\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} + \cdots +$$

$$+A_{in}\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

Portanto $\det A = \sum_k A_{ik} D_{ik}$ com $D_{ik} = \det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2k} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nk} & \cdots & A_{nn} \end{pmatrix}$, com a linha i substituida pro zeros exceto na coluna k em que é substituída por 1. Agora vamos trocando as linhas até levar a linha i para a primeira linha na ordem. Não podemos simplesmente trocar a linha i pela linha 1 porque isso desordena o resto da matriz. Vamos fazer o seguinte: primeiro trocamos a linha i com a linha $i-1$, depois com a linha $i-2$ e assim sucessivamente até chegar a primeira linha. Nesse processo trocamos de linhas $(i-1)$ vezes, portanto:

$$\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2k} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nk} & \cdots & A_{nn} \end{pmatrix} = (-1)^{i-1}\det\begin{pmatrix} 0 & 0 & \cdots & 1 & \cdots & 0 \\ A_{11} & A_{12} & \cdots & A_{1k} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2k} & \ddots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ \\ A_{n1} & A_{n2} & \cdots & A_{nk} & \cdots & A_{nn} \end{pmatrix}$$

Agora vamos trocando as colunas da mesma forma até levar a coluna k para a primeira coluna, então:

$$\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2k} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nk} & \cdots & A_{nn} \end{pmatrix} = (-1)^{i-1}(-1)^{k-1}\det\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ A_{1k} & A_{11} & A_{12} & \cdots & A_{1n} \\ A_{2k} & A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ & & & \ddots & \vdots \\ A_{nk} & A_{n1} & A_{n2}\cdots & \cdots & A_{nn} \end{pmatrix}$$

Ou seja

$$D_{ik} = (-1)^{i+k}\det\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ A_{1k} & A_{11} & A_{12} & \cdots & A_{1n} \\ A_{2k} & A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ & & & \ddots & \vdots \\ A_{nk} & A_{n1} & A_{n2}\cdots & \cdots & A_{nn} \end{pmatrix} = cof(A)_{ik}$$

Pois extraimos a linha i e a coluna k.

Então, acabamos de provar que $\det A = \sum_k A_{ik} cof(A)_{ik} = \sum_k A_{kj} cof(A)_{kj}$

Nesse caso temos dois métodos de decomposição no cofatores para calcular um determinante – escolhe uma linha e usa a primeira somatória, ou escolhe uma coluna e usa a segunda. A idéia é procurar linhas ou colunas com muitos zeros para facilitar os cálculos. Para fazer determinantes 4×4 ou 5×5 na mão, dependendo da quantidade de zeros na matriz, esse método pode ser até rápido. Se precisar de uma resposta analítica em lugar de numérica também deve-se usar esse método. Para calcular valores numéricos de qualquer determinante não é o mais adequado. Entretanto, o método é valioso demais para provar o teorema a seguir:

Teoremas:

$\sum_k A_{ik} cof\left(A\right)_{jk} = \delta_{ij} \det A$ ou $\sum_k A_{kj} cof\left(A\right)_{ki} = \delta_{ij} \det A$

Prova: vamos criar uma matriz A' à partir da A trocando a linha j pela i, i.e.,

$$A'_{pq} = \begin{cases} A_{pq} & se \quad p \neq j \\ A_{iq} & se \quad p = j \end{cases}$$

Note que se $i \neq j$ ficamos com duas linhas iguais. Agora $\det A' = \sum_k A'_{jk} \, cof\left(A'\right)_{jk} = \sum_k A_{ik} \, cof\left(A\right)_{jk} = 0 \quad i \neq j$ por conta das duas linhas iguais. Se escolhemos a linha j para a expansão dos cofatores então $cof\left(A'\right) = cof\left(A\right)$. Se, por outro lado, $i = j$, a troca não muda nada e $\sum_k A_{ik} \, cof\left(A\right)_{ik} = \det A$. Nesse caso podemos então escrever que

$$\sum_k A_{ik} \, cof\left(A\right)_{jk} = \sum_k A_{kj} \, cof\left(A\right)_{ki} = \delta_{ij} \det A$$

Note agora que $\sum_k A_{ik} \, cof\left(A\right)_{jk} = \sum_k A_{ik} \, \overline{cof\left(A\right)}_{kj} = \sum_k A_{ik} \, A^{\dagger}_{kj}$ onde matriz adjunta de A, dada por $A^{\dagger}_{kj} = \overline{cof\left(A\right)}_{kj}$, é a transposta da matriz cofator. Note o que temos aqui:

$$AA^{\dagger} = \left(\det A\right) I \text{ ou } A\left[\frac{1}{\det A} A^{\dagger}\right] = I$$

Por outro lado também vale:

$$A^{\dagger} A = \left(\det A\right) I \text{ ou } \left[\frac{1}{\det A} A^{\dagger}\right] A = I$$

Logo acabamos de encontrar a matriz inversa de A ($AA^{-1} = A^{-1}A = I$): $A^{-1} = \frac{1}{\det A} A^{\dagger}$. A inversa é a transposta do cofator dividida pelo determinante de A. Note que essa fórmula só pode ser utilizada se $\det A \neq 0$.

Assim podemos encontrar facilmente a matriz inversa de uma matriz 2×2 não singular:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \rightarrow cof\left(A\right) = \begin{bmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{bmatrix} \rightarrow A^{\dagger} = \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$$

Ou seja, para obter a matriz adjunta simplesmente trocamos os dois elementos da diagonal e os sinais dos dois elementos fora da diagonal. A matriz inversa é dada por:

$$A^{-1} = \frac{\begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}}{(A_{11}A_{22} - A_{12}A_{21})}$$

Formalização de Matriz Inversa

1. Definição de matrizes SINGULAR e NÃO-SINGULAR. A matriz **A** é **SINGULAR** se $\det(A) = 0$, e é **NÃO SINGULAR** se $\det(A) \neq 0$.

1.1. Divisores da matriz NULA: Se $AB = O$ então $A = O$ ou $B = O$ ou A e B, as duas, são matrizes singulares. Claro que se $A = O$ ou $B = O$ então $AB = O$. Mas estamos interssados no caso em que $A \neq O$, $B \neq O$ mas $AB = O$. Aplicando o determinante temos que $\det(AB) = \det A \times \det B = 0$. Pareceria então que exigir que $\det A = 0$ ou $\det B = 0$ seria suficiente, mas essa caso é mais forte, os dois devem ser nulos. Suponha que $\det A \neq 0$ e $\det B = 0$. Mas se $\det A \neq 0$ então A admite inversa, portanto $A^{-1}AB = A^{-1}O \quad \to B = O$ em contradição com $B \neq O$.

2. Definição de matriz inversa. Se existir uma matriz X tal que $AX = I$ e que $XA = I$ então X é a matriz inversa de A, denotada por $X = A^{-1}$.

2.1. Se a inversa existe ela é única. Supor que existem duas diferentes, ou seja, $\exists Y \neq X \;/\; AY = I$. Mas $X = XI = X(AY) = (XA)Y = IY = Y$ ou seja $X = Y$ em contradição com $X \neq Y$. Logo a inversa, se existir, é única.

2.2. Se A admite inversa então A é não singular. $AA^{-1} = I \quad \to \det(AA^{-1}) = \det I = 1 \quad \to \det(A)\det(A^{-1}) = 1$, logo $\det A \neq 0$ e $\det A^{-1} \neq 0$. Mais ainda $\det A^{-1} = \dfrac{1}{\det A}$.

2.3. Vale também o converso, se A é não singular então A admite inversa. Nesse caso $\dfrac{1}{\det A} A^{\dagger}$ existe e $A^{-1} = \dfrac{1}{\det A} A^{\dagger}$.

Então podemos afirmar A admite inversa □ A é não singular.

2.4. Se A é singular, A não admite inversa, pois $\det(A)\det(A^{-1}) = 1$ logo $0 \times \det(A^{-1}) = 1 \quad \Rightarrow 0 = 1$ gera uma contradição.

2.5. A é a inversa de A^{-1}. Pois $AA^{-1} = I \quad e \quad A^{-1}A = I$.

2.6. Se A é não singular, $\bar{A}$ também é não singular, pois $\det \bar{A} = \det A \neq 0$.

2.7. Se A e B são não singulares então AB é não singular, admite inversa e $(AB)^{-1} = B^{-1}A^{-1}$ é a inversa de AB.

$$\left(B^{-1}A^{-1}\right)(AB) = B^{-1}\left(A^{-1}A\right)B = B^{-1}IB = B^{-1}B = I$$

2.8. Se A é não singular então $\left(\bar{A}\right)^{-1} = \overline{A^{-1}}$.

Partindo de $AA^{-1} = I$ e aplicando a transposta $\overline{AA^{-1}} = \bar{I} = I \quad \Rightarrow \overline{A^{-1}}\bar{A} = I$ logo $\overline{A^{-1}} = \left(\bar{A}\right)^{-1}$.

OPERAÇÕES ELEMENTARES:

São três as operações elementares sobre matrizes quadradas.

1. Multiplicar uma linha, ou uma coluna, por um escalar $\alpha \neq 0$. O determinante é multiplicado por α.
2. Permutar duas linhas, ou duas colunas. O determinante é multiplicado por -1.
3. Adicionar à uma linha, ou uma coluna, um múltiplo de outra linha, ou coluna. Nesse caso o determinante é preservado.

Vamos denotar um operador por $\hat{O}$ que aplicado em uma matriz A a modifica para a matriz A', ou seja $A' = \hat{O}A$. Para diferenciar operações nas linhas das colunas vamos chamar o operador nas linhas de $\hat{O}_L$ e o operador nas colunas de $\hat{O}_C$.

Teorema: $\hat{O}_L(AB) = \left(\hat{O}_L A\right)B$ e $\hat{O}_C(AB) = A\left(\hat{O}_C B\right)$.

O operador $\hat{O}_L$ só atua nas linhas, ou seja, apenas no primeiro índice do produto $(AB)_{ij} = \sum_k A_{ik}B_{kj}$, portanto $\hat{O}_L(AB)_{ij} = \sum_k \left(\hat{O}_L A_{ik}\right)B_{kj}$ que significa $\hat{O}_L(AB) = \left(\hat{O}_L A\right)B$. Já o operador $\hat{O}_C$ só atua no índice j e $\hat{O}_C(AB)_{ij} = \sum_k A_{ik}\left(\hat{O}_C B_{kj}\right)$.

Matrizes elementares.

Uma matriz elementar é obtida aplicando operadores linha e/ou coluna sobre a matriz identidade. Denotamos essas matrizes por E e E_L ou E_C se a operação for apenas sobre linhas ou colunas.

Exemplo: A matriz $E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ troca as linhas 1 e 2. Veja $E_1 A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} A_{21} & A_{22} & A_{23} \\ A_{11} & A_{12} & A_{13} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$. A matriz $E_2 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ multiplica a linha 1 por α. A matriz $E_3 = \begin{pmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ soma à linha 1, $\alpha \times$ linha 2. Note:

$$E_3 A = \begin{pmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} A_{11} + \alpha A_{21} & A_{12} + \alpha A_{22} & A_{13} + \alpha A_{23} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

Já a matriz $E_4 = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ soma á coluna 1, $\alpha \times$ coluna 2, ou seja, opera nas colunas, em lugar das linhas.

Teorema.

1. $\det E \neq 0$, pois selecionamos as operações elementares com essa propriedade. Note que $\det I = 1$ e que $\det E = k \neq 0$ para a operação multiplicar uma linha por uma constante não nula, $\det E = -1$ para a operação trocar duas linhas e $\det E = 1$ para a operação adicionar á uma linha um múltiplo de outra linha.
2. Realizar uma operação sobre as linhas é equivalente à pré-multiplicar a matriz pela matriz elementar da operação desejada. Já realizar a operação sobre colunas é equivalente a pós-multiplicar pela matriz elementar.

$A = IA \quad \rightarrow O_L IA = (O_L I)A = E_L A$ por outro lado $A = AI \quad \rightarrow O_C AI = A(O_L I) = AE_C$

3. Como o determinante das matrizes elementares não é nulo elas admitem inversa. A inversa de E também é uma matriz elementar. Se $E = \hat{O} I \quad \rightarrow I = \hat{O}^{-1} E \quad \rightarrow IE^{-1} = \hat{O}^{-1} E E^{-1} \quad \rightarrow E^{-1} = \hat{O}^{-1} I$, ou seja E^{-1} também é obtida pela operação elementar reversa aplicada na matriz identidade.

Matrizes equivalentes $A \sim B$

Definição: $A \sim B$, A é equivalente à B, se B pode ser obtido de A através de uma cadeia de operações elementares.

Teoremas:

1. $A \sim A$, pois $A = IA$ e I é uma matriz elementar.
2. Se $A \sim B$ então $B \sim A$, pois se $A = E_1 E_2 \cdots E_n B \quad \rightarrow B = E_n^{-1} \cdots E_2^{-1} E_1^{-1} A$.
3. Se $A \sim B$ e $B \sim C$ então $A \sim C$. Note que $A = E_1 E_2 \cdots E_n B$ e $B = E_m'^{-1} \cdots E_2'^{-1} E_1'^{-1} C$ logo $A = E_1 E_2 \cdots E_n E_m'^{-1} \cdots E_2'^{-1} E_1'^{-1} C$.
4. Todas as matrizes quadrada não singulares podem ser expressas como um produto de matrizes elementares.

Se A é não singular então A admite inversa. Isso significa que um conjunto de operações elementares leva A até I, ou seja, $E_1 E_2 \cdots E_n A = I$, logo $A = E_n^{-1} \cdots E_2^{-1} E_1^{-1} I = E_n^{-1} \cdots E_2^{-1} E_1^{-1}$. Isso também significa que $A^{-1} = E_1 E_2 \cdots E_n$.

Método para achar matriz inversa através de operações elementares.

Coloque as duas matrizes, A e I lado a lado. Agora, usando operações elementares de linhas vá transformando A na matriz identidade. Toda operação realizada sobre A também deve ser realizada sobre I. Quando A se tornar a identidade a identidade se tornou A^{-1}.

Exemplo 1: achar inversa de $A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$. Usando a matriz adjunta sabemos que $A^{-1} = \frac{1}{(10-12)} \begin{pmatrix} 5 & -3 \\ -4 & 2 \end{pmatrix} = \begin{pmatrix} -5/2 & 3/2 \\ 2 & -1 \end{pmatrix}$. Agora vamos usar apenas as operações elementares. Colocar as duas lado à lado:

$$\left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right) \div 2 \rightarrow \left(\begin{array}{cc|cc} 1 & 3/2 & 1/2 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right) \begin{matrix} \times -4 \\ + linha1 \end{matrix} \rightarrow \left(\begin{array}{cc|cc} 1 & 3/2 & 1/2 & 0 \\ 0 & -1 & -2 & 1 \end{array}\right) \begin{matrix} + linha2 \\ \times 3/2 \end{matrix} \rightarrow$$

$$\rightarrow \left(\begin{array}{cc|cc} 1 & 0 & -5/2 & 3/2 \\ 0 & -1 & -2 & 1 \end{array}\right) \times -1 \rightarrow \left(\begin{array}{cc|cc} 1 & 0 & -5/2 & 3/2 \\ 0 & 1 & 2 & -1 \end{array}\right)$$

Então $A^{-1} = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} -5/2 & 3/2 \\ 2 & -1 \end{pmatrix}$.

Exemplo 2: achar inversa de $A=\begin{pmatrix} 2 & 4 & 5 \\ 1 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}$. Faremos mais de uma operação elementar por vez.

$$\left(\begin{array}{ccc|ccc} 2 & 4 & 5 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right)\begin{array}{l} \div 2 \\ \\ \div 2 \end{array} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 5/2 & 1/2 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 1 & 1 & 3/2 & 0 & 0 & 1/2 \end{array}\right)\begin{array}{l} \\ -linha1 \rightarrow \\ -linha1 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 5/2 & 1/2 & 0 & 0 \\ 0 & 1 & -1/2 & -1/2 & 1 & 0 \\ 0 & -1 & -1 & -1/2 & 0 & 1/2 \end{array}\right)\begin{array}{l} -2\times linha2 \\ \\ +linha2 \end{array} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 7/2 & 3/2 & -2 & 0 \\ 0 & 1 & -1/2 & -1/2 & 1 & 0 \\ 0 & 0 & -3/2 & -1 & 1 & 1/2 \end{array}\right)\begin{array}{l} \\ \\ \times -2/3 \end{array} \rightarrow$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 7/2 & 3/2 & -2 & 0 \\ 0 & 1 & -1/2 & -1/2 & 1 & 0 \\ 0 & 0 & 1 & 2/3 & -2/3 & -1/3 \end{array}\right)\begin{array}{l} -7/2\times linha3 \\ +1/2\times linha3 \rightarrow \\ \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5/6 & 2/6 & 7/6 \\ 0 & 1 & 0 & -1/6 & 4/6 & -1/6 \\ 0 & 0 & 1 & 4/6 & -4/6 & -2/6 \end{array}\right)$$

Ou seja $A^{-1}=\begin{pmatrix} 2 & 4 & 5 \\ 1 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}^{-1}=\frac{1}{6}\begin{pmatrix} -5 & 2 & 7 \\ -1 & 4 & -1 \\ 4 & -4 & -2 \end{pmatrix}$.

Matrizes na forma escada reduzida por linha:

Uma matriz na forma escada reduzida obedece aos seguintes critérios:

1. O primeiro elemento não nulo de uma linha é 1.
2. Se A_{ij} é o primeiro elemento não nulo da linha i então $A_{ij}=0 \quad \forall j \neq i$, ou seja, todos os elementos da coluna j, exceto A_{ij}, são nulos.
3. Todas as linhas nulas estão abaixo das linhas não nulas.
4. Se as linhas $1,2,\cdots,p$ são não nulas e k_i é o primeiro elemento não nulo da linha i então $k_1<k_2<\cdots<k_p$.

Teorema: Toda matriz $A_{m\times n}$ pode ser colocada na forma escada através de operações elementares.

A demonstração do teorema é a descrição do procedimento para atingir o objetivo – colocar a matriz na forma escada reduzida por linhas. O procedimento é o seguinte:

1. Suponha que já se colocou as $i-1$ linhas com primeiro elemento não nulo iguais à 1 e todo o resto da coluna desse elemento nula. Tome agora a linha i :
2. Se ela é nula nada precisa ser feito.
3. Se o primeiro elemento não nulo é o A_{ij} divida a linha por A_{ij}. Assim tornamos o primeiro elemento não nulo igual a 1. Falta anular todos os elementos da coluna j .
4. Some todos os elementos da linha $k \neq i$ com a linha i multiplicada por $\left(-A_{kj}\right)$. Assim anulamos todos os elementos da linha j exceto o $A_{ij} = 1$.
5. Repita o procedimento para as próximas linhas. Note que os zeros abaixo e acima do elemento igual a 1 não desfazem os 1 e zeros já obtidos.
6. Permute as linhas até colocar a matriz na forma escada.

Exemplo: colocar a matriz $M = \begin{pmatrix} 0 & 2 & 1 & 4 \\ 3 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$ na forma escada reduzida por linhas.

$$\begin{pmatrix} 0 & 2 & 1 & 4 \\ 3 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{matrix} \times 1/2 \\ \times 1/3 \\ \times 1/3 \end{matrix} \rightarrow \begin{pmatrix} 0 & 1 & 1/2 & 2 \\ 1 & 0 & 5/3 & 2/3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} -2 \times linha3 \\ -2/3 \times linha3 \\ \ \end{matrix} \rightarrow \begin{pmatrix} 0 & 1 & 1/2 & 0 \\ 1 & 0 & 5/3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{matrix} trocar\ linha1 \\ com\ linha2 \end{matrix} \rightarrow \begin{pmatrix} 1 & 0 & 5/3 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Unicidade da forma escada reduzida por linhas.

Qualquer operação elementar de linhas em uma matriz na forma escada reduzida por linhas ou a tira da condição de escada ou nada modifica, deixando a matriz invariante.

1. Multiplicar uma linha por $\alpha \neq 1$. Se a linha for nula nada muda. Se a linha não for nula contém o primeiro elemento não nulo igual a 1, que se torna igual a α após a operação e a matriz não está mais na forma escada.
2. Trocar duas linhas não nula altera a ordem escada $k_1 < k_2 < \cdots < k_p$.
3. Somar duas linhas não nulas colocaria 1 onde deveria ser zero

Isso significa que se $A_{esc} \square B_{esc}$ então $A_{esc} = B_{esc}$. Matematicamente podemos provar a unicidade da forma usual, supondo que existe $A'_{esc} \neq A_{esc}$. Nesse caso então $A_{esc} = E_1 E_2 \cdots E_n A$ e $A'_{esc} = E'_1 E'_2 \cdots E'_m A$, logo $A_{esc} \square A$ e $A'_{esc} \square A$ então $A_{esc} \square A'_{esc}$ o que implica que $A_{esc} = A'_{esc}$. Cuidado para manter apenas operações de linhas, porque operações elementares de colunas podem gerar duas matrizes escadas. Exemplo:

$$\begin{pmatrix} 1 & 0 & 0 & 5/3 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 5/3 & 0 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Nesse caso trocamos as colunas 3 e 4 e as duas matrizes, embora diferentes, estão na forma escada.

POSTO e NULIDADE de uma Matriz: ao terminar de colocar a matriz na forma escada reduzida por linhas o número r de linhas não nulas será o posto [rank em inglês] da matriz M. Ou seja, posto de A, $p(A)$, é o número de linhas não nulas de A_{esc} □ A. A nulidade de $A_{m\times n}$ é $N(A) = n - p$, o número de colunas subtraído do posto. Note que no caso de um sistema de equações lineares o número de colunas é igual ao número de incógnitas.

Antes de utilizar o conceito poderoso de posto de uma matriz na resolução de sistema de equações lineares vamos estabelecer alguns teoremas.

Teoremas do posto de uma matriz $A_{m\times n}$.

1. Se $p(A) = 0$ então $A = O$. Claro, pois todas as linhas são nulas.
2. $p(A) \leq m$. Trivial pois A_{esc} também é $m \times n$ e o número de linhas não nulas não pode ser maior do que o número de linhas.
3. $p(A) \leq n$. Aqui complica um pouco por estamos usando colunas em lugar de linhas. Isso implica que $N(A) = n - p \geq 0$. Na forma escada $k_1 < k_2 < \cdots < k_p$ onde $p =$ posto = número de linhas não nulas. Por outro lado k_i é a coluna do iésimo elemento diferente de zero da linha i. Logo $k_i \leq n$ e $k_p \leq n$. As condições da forma escada requerem que $k_1 \geq 1$, $k_2 > k_1 \geq 2$ e, continuando, $k_p \geq p$. Então concluímos que $p \leq k_p \leq n$, ou seja, $p \leq n$.

Definição equivalente de Posto.

Posto de uma matriz A é o máximo p para o qual uma sub-matriz quadrada de A na forma $A_{k\times k}$ é não singular, ou seja, tem determinante não nulo. Fica claro a equivalência com a definição anterior para matrizes na forma escada. As operações elementares jamais anulam o determinante de uma matriz não singular, portanto podemos permutar as colunas da matriz na forma escada para colocar os 1´s em colunas adjacentes crescentes. Todas as linhas após p serão nulas, então o máximo determinante não nulo tem dimensão $p \times p$. Qualquer sub-matriz $A_{l\times l}$ terá linhas nulas com determinantes nulos. Procurar sub-matriz não singular pode ser feito diretamente na matriz A sem necessidade de reduzi-la à forma escada. O fato de que $\det E \neq 0$ sempre garante que as operações elementares sobre A jamais anularão qualquer determinante que já não fosse nulo. A vantagem dessa definição é que facilita a demonstração de teoremas. Do ponto de vista de operacionalidade pode ser muito fácil descobrir a sub-matriz não singular mas fica complicado operacionalizar o procedimento de sair procurando uma a uma as submatrizes. Melhor começar pelas matrizes de dimensão mais alta possível.

Exemplos:

Achar $p(A)$ onde $A_{3\times4} = \begin{pmatrix} 0 & 2 & 1 & 4 \\ 3 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$. Como $A_{3\times4}$ o posto será no máximo 3. Fica fácil de ver que a sub-matriz $\left(\begin{array}{c|ccc} 0 & 2 & 1 & 4 \\ 3 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{array}\right)$ marcada é triangular superior com determinante diferente de zero, logo $p(A)=3$.

Achar $p(A)$ onde $A_{4\times5} = \begin{pmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. Como a última linha é nula sabemos que o posto será no máximo 3. Esolhendo a sub-matriz $\begin{pmatrix} \boxed{1} & 2 & \boxed{0} & 3 & \boxed{0} \\ \boxed{0} & 0 & \boxed{1} & 2 & \boxed{0} \\ \boxed{0} & 0 & \boxed{0} & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ obtemos uma sub-matriz identidade com determinante não nulo, logo $p(A)=3$.

Teoremas. Com essa definição fica fácil mostrar que:

1. $0 \le p(A_{m\times n}) \le \min\{m,n\}$ pois uma sub-matriz quadrada só pode ser, no máximo, de ordem $\min\{m,n\}$.
2. $p(\overline{A}) = p(A)$. Sabemos que $\det \overline{A} = \det A$. Portanto se extraímos uma sub-matriz $A_{p\times p}$ de A com determinante não nulo também podemos extrair a $\overline{A_{p\times p}}$ da matriz $\overline{A}$. Se A tiver uma linha nula a $\overline{A}$ terá uma correspondente coluna nula.
3. Se A' é sub-matriz de A então $p(A') \le p(A)$.
4. $p(EA) = p(AE) = p(A)$, pois $\det(EA) = \det E \times \det A$. Operações elementares, tanto nas linhas quanto nas colunas, não alteram o posto de uma matriz.
5. $p(\lambda A) = p(A)$, $\lambda =$constante.
6. $p(AB) = \min\{p(A), p(B)\}$.

Se A e B são conformes para multiplicação então $A_{m\times n}$ e $B_{n\times k}$ e o produto $A_{m\times n}B_{n\times k}=C_{m\times k}$. Note que as operações elementares de linha só se aplicam à matriz A do produto. Através de operações elementares podemos levar a matriz A para a forma escada com $m-p(A)$ linhas nulas. O produto então também terá, pelo menos, $m-p(A)$ linhas nulas. sobre colunas só se aplica à matriz B. Isso implica que $p(AB)\leq p(A)$. Por outro lado $p(AB)=p\left(\overline{AB}\right)=p\left(\overline{B}\,\overline{A}\right)\leq p\left(\overline{B}\right)=p(B)$, então $p(AB)=\min\left\{p(A),p(B)\right\}$.

7. O posto de uma matriz não muda se ele é pré/pós-multiplicada por uma matriz quadrada não singular.

Pré-multiplicação: $C_{m\times n}=X_{m\times m}A_{m\times n}$. Como X é não singular então $p(X)=m$ e $p(A)\leq m$, logo $p(XA)\leq p(A)$, pela teorema 6. Por outro lado $p(A)=p\left(X^{-1}XA\right)\leq p(XA)$, que nos leva a $p(XA)\leq p(A)\leq p(XA)$ com a única solução possível $p(XA)=p(A)$. Basta transpor para concluir o mesmo para a pós-multiplicação $p(AX)=p(A)$.

Dois sistemas de equações lineares são equivalentes se eles possuem as mesmas soluções e as mesmas nulidades, ou variáveis livres.

8. $p\left(\overline{A}A\right)=p(A)$. Note que $\overline{A}A$ é uma matriz simétrica quadrada.

Considere o sistema $A_{m\times n}\vec{X}_{n\times 1}=\vec{0}_{m\times 1}$ e o sistema $\overline{A}_{n\times m}\,A_{m\times n}\vec{X}_{n\times 1}=\vec{0}_{n\times 1}$. Nesse caso $\overline{\vec{X}_{1\times n}}\,\overline{A}_{n\times m}\,A_{m\times n}\vec{X}_{n\times 1}=\vec{0}_{1\times 1}=0$. Isso pode ser re-escrito como $\left(\overline{A\vec{X}}\right)\left(A\vec{X}\right)=0$. Mas $\left(A\vec{X}\right)=\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$ e $\left(\overline{A\vec{X}}\right)=\begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix}$ e $\left(\overline{A\vec{X}}\right)\left(A\vec{X}\right)=v_1^2+v_2^2+\cdots+v_m^2=0$ só será verdadeiro se $v_i=0\quad\forall i$. Logo $A\vec{X}=\vec{0}$ é a única solução para $\overline{A}A\vec{X}=\vec{0}$, ou seja, $\overline{A}A\vec{X}=\vec{0}$ e $A\vec{X}=\vec{0}$ são equivalentes. Igualando as nulidades $n-p(A)=n-p\left(\overline{A}A\right)$ concluímos que $p\left(\overline{A}A\right)=p(A)$.

Sistema de Equações Lineares:

1. Uma equação linear nas incógnitas $x_1,x_2,\cdots,x_n$ é escrito como: $a_1x_1+a_2x_2+\cdots+a_nx_n=b$. Se $b=0$ a equação é chamada de equação linear HOMOGÊNEA.

2. Um sistema de m equações lineares nas incógnitas $x_1, x_2, \cdots, x_n$ é escrito como:

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{matrix}$$

ou, na forma matricial:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Se $b_1 = b_2 = \cdots = b_m = 0$, ou seja, $\vec{b} = \vec{0}$, ou ainda $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, o sistema é chamado de HOMOGÊNEO.

3. Equações redundantes. Suponha as seguintes equações: $\begin{matrix} 1. & x + y = 5 \\ 2. & 2x + y = 3 \\ 3. & 3x + 2y = 8 \end{matrix}$. A terceira equação foi obtida simplesmente somando as duas primeiras. Não contém qualquer informação extra. É uma equação redundante. Correta, mas desnecessária.
4. Solução. Qualquer conjunto $(\xi_1, \xi_2, \cdots, \xi_n)$ que satisfaça a todas equações do sistema é uma solução. Em termos das soluções os sistemas são classificados em CONSISTENTES ou INCONSISTENTES. Os sistemas inconsistentes não admitem qualquer solução, ou seja, o conjunto solução é vazio. Os sistemas consistentes admitem pelo menos uma solução.

Exemplo: $\begin{matrix} 1. & x_1 + x_2 = 1 \\ 2. & x_1 + x_2 = 0 \end{matrix}$ nos leva a $1 = 0$, uma inconsistência. Não existe solução para esse sistema.

O sistemas consistentes podem ter uma única solução ou infinitas soluções.

Por exemplo, a equação dois do sistema $\begin{matrix} 1. & x_1 + x_2 = 1 \\ 2. & x_1 - x_2 = 0 \end{matrix}$ implica que $x_2 = x_1$ que substituída na equação um implica que $x_1 = 1/2$ e o conjunto $(\xi_1, \xi_2) = (1/2, 1/2)$ é a única solução do sistema. Por outro lado a equação $x_1 + x_2 = 1$ admite infinitas soluções. Escolhe um valor qualquer para x_1 e $x_2 = 1 - x_1$.

SOLUÇÃO TRIVIAL: um sistema homogêneo

Um sistema homogêneo do tipo:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Admite pelo menos uma solução, a solução trivial $(0,0,\cdots,0)$. Todo sistema homogêneo é consistente.

No sistema:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Chamamos $A_{m\times n} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ de matriz dos coeficientes do sistema. Já $\vec{x}$ é o vetor das incógnitas $\vec{x}_{n\times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e $\vec{b}_{m\times 1} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$. Chamamos de matriz ampliada a matriz: $\left(A\middle|\vec{b}\right)_{m\times(n+1)} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$. O sistema pode então ser escrito na forma matricial como: $A\vec{x} = \vec{b}$.

Se o sistema for consistente então $A\vec{x} = \vec{b}$ admite pelo menos uma solução. Vamos chamar $\vec{x}_o$ uma solução particular do sistema, então $A\vec{x}_o = \vec{b}$ é verdadeira. O sistema homogêneo associado $A\vec{x} = \vec{0}$ é sempre consistente, logo também admite pelo menos uma solução $\vec{x}_H$ tal que $A\vec{x}_H = \vec{0}$. Nesse caso então $\vec{x}_o + \vec{x}_H$ também é uma solução de $A\vec{x} = \vec{b}$, pois $A\left(\vec{x}_o + \vec{x}_H\right) = A\vec{x}_o + A\vec{x}_H = \vec{b} + \vec{0} = \vec{b}$. Se a única solução do sistema homogêneo for a trivial então $\vec{x}_H = \vec{0}$ e $\vec{x}_o + \vec{x}_H = \vec{x}_o$.

Resolução de sistemas de equações lineares:

Vamos voltar ao sistema
$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

As soluções não mudam por qualquer operação elementar entre linhas realizadas de ambos os lados da igualdade. Considere agora a matriz dos coeficientes A e a matriz ampliada $\left(A\middle|\vec{b}\right)$. Podemos então mostra os seguintes teoremas:

1. Se $p(A) < p\left(A\middle|\vec{b}\right)$ o sistema é inconsistente e não admite qualquer solução.

Nesse caso ao reduzir o sistema à forma escada obeteríamos um sistema do tipo: $0x_1' + 0x_2' + \cdots + 0x_n' = b_k' \neq 0$ que gera a inconsistência $0 = b_k' \neq 0$.

2. Se $p(A) = p\left(A\middle|\vec{b}\right) = n$ o sistema é consistente e admite apenas uma solução. Note que a nulidade desse sistema é zero. Se $p(A) = n$ é porque $\left(k_1, k_2, \cdots, k_p\right) = (1, 2, \cdots, n)$ o que significa que as matrizes A e $\left(A\middle|\vec{b}\right)$ reduzidas na forma escada tem que ser do tipo:

$$A_{esc} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ e } \left(A\middle|\vec{b}\right) = \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & c_1 \\ 0 & 1 & 0 & 0 & c_2 \\ 0 & 0 & 1 & 0 & \vdots \\ 0 & 0 & 0 & 1 & c_4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$
o que levaria a algumas equações redundantes e a solução única $x_i = c_i \quad i \in [1, n]$.

3. Se $p(A) = p\left(A\middle|\vec{b}\right) < n$ o sistema é consistente e admite infinitas soluções. Neste caso a nulidade $n - p$ é o número de variáveis livres que podem passar para o lado direito da equação. Sejam $k_1, k_2, \cdots, k_p$ as colunas dos primeiros termos não nulos de cada linha. Nesse caso sabemos que $A_{esc\,1k_1} = 1,\ A_{esc\,2k_2} = 1$, ..., $A_{esc\,pk_p} = 1$ e que esses termos são os únicos não nulos nas colunas $k_1, k_2, \cdots, k_p$. Passamos para o lado direito as $n - p$ incógnitas restantes para obter um sistema de equações do tipo:

$$
\begin{aligned}
x_{k_1} &= c_1 - \sum_j \gamma_{1j} x_j \qquad j \notin \{k_1, k_2, \cdots, k_p\} \\
x_{k_2} &= c_2 - \sum_j \gamma_{2j} x_j \\
\vdots \quad & \quad \vdots \\
x_{k_p} &= c_p - \sum_j \gamma_{pj} x_j
\end{aligned}
$$

$$
\begin{pmatrix} x_{k_1} \\ x_{k_2} \\ \vdots \\ x_{k_p} \end{pmatrix} = \begin{pmatrix} -\gamma_{1j_1} & -\gamma_{1j_2} & \cdots & -\gamma_{1j_{n-p}} \\ -\gamma_{2j_1} & -\gamma_{2j_2} & \cdots & -\gamma_{2j_{n-p}} \\ \vdots & \vdots & \ddots & \vdots \\ -\gamma_{pj_1} & -\gamma_{pj_2} & \cdots & -\gamma_{pj_{n-p}} \end{pmatrix} \begin{pmatrix} x_{j_1} \\ x_{j_2} \\ \vdots \\ x_{j_{n-p}} \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{pmatrix}
$$

$$
\vec{X}_{dep} = C\vec{X}_{ind} + \vec{d}
$$

Dados $x_j \ j \notin \{k_1, k_2, \cdots, k_p\}$ os valores de $x_i \ i \in \{k_1, k_2, \cdots, k_p\}$ são únicos, mas como os $x_j{}'s$ podem assumir quaisquer valores existem infinitas soluções.

Suponha um sistema consistente com n variáveis, posto p e nulidade $n-p$. Note então que o sistema possui $n-p$ variáveis livres, independentes, e p variáveis dependentes. Sempre podemos escolher as $p's$ primeiras como as dependentes e as $n-p$ últimas como independentes. Agora vamos reescrever o sistema

$$
\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}
$$

Exemplo: suponha que após a redução à forma escada o sistema ficou na forma:

$$\begin{pmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$$ no qual a matriz ampliada é dada por $\left(\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 0 & 1 \\ 0 & 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$. Vemos que estamos no caso em que $p(A) = p\left(A\middle|\vec{b}\right) = 3 < 5$ e a nulidade é $n - p = 5 - 3 = 2$. Logo existem duas variáveis livres. Nesse problema, seguindo a nossa notação, $k_1 = 1$, $k_2 = 3$ e $k_3 = 5$, assim escolhemos $x_1, x_3\ e\ x_5$ para ficarem do lado esquerdo das equações:

$$\begin{aligned} x_1 + 2x_2 + 3x_4 &= 1 \quad \to \quad x_1 = 1 - 2x_2 - 3x_4 \\ x_3 + 2x_2 &= 2 \quad \to \quad x_3 = 2 - 2x_2 \\ x_5 &= 3 \quad \to \quad x_5 = 3 \end{aligned}$$

As duas variáveis livres são $x_2\ e\ x_4$, x_5 está determinada mas $x_1\ e\ x_3$ podem assumir infinitos valores dependendo de $x_2\ e\ x_4$.

4. O sitema homogêneo $A\vec{x} = \vec{0}$ com n incógnitas admite uma solução NÃO TRIVIAL se, e somente se, $p(A) < n$. Prova: se $p(A) = n$ então $p\left(A\middle|\vec{0}\right) = n$ pois o acréscimo de uma coluna de zeros não altera o posto da matriz, então o sistema admite apenas uma solução. Sabemos que a solução trivial $\vec{x} = \vec{0}$ é uma solução, logo é a única solução. Se $p(A) < n$ então o sistema admite infinitas soluções. No caso do sistema acima teríamos:

$$\begin{aligned} x_1 + 2x_2 + 3x_4 &= 0 \quad \to \quad x_1 = -2x_2 - 3x_4 \\ x_3 + 2x_2 &= 0 \quad \to \quad x_3 = -2x_2 \\ x_5 &= 0 \quad \to \quad x_5 = 0 \end{aligned}$$

5. Se o sistema homgêneo $A\vec{x} = \vec{0}$ com $A_{n\times n}$ possui solução NÃO TRIVIAL, então A é singular. Se o sistema admite solução não trivial então $p(A) < n$, o que significa que ao reduzir a matriz para a forma escada via operações elementares, que jamais anulam o determinante, terminaremos com uma ou mais linhas nulas. O determinante de uma matriz com linha nula é ZERO. Logo $\det A = 0$ e A é singular.
6. Se $p(A) = p$ então A contém uma sub-matriz quadrada $p \times p$ com determinante não nulo. Novamente podemos usar o fato de que as operações elementares, via linhas ou colunas, jamais anulam o determinante. Então podemos transformar a matriz original, via operações elementares de linhas e ou colunas na forma:

$$\left(\begin{array}{cccc|ccc}
1 & 0 & 0 & & & & \\
0 & 1 & 0 & & & & \\
0 & 0 & 1 & & & & \\
 & & & 1 & & & \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & & 0 & 0
\end{array}\right)$$

Nesse caso o determinante da sub-matriz $p \times p$ das p primeiras linhas e colunas é obviamente não nulo, vale 1, pois todos os elementos da diagonal valem 1. Além do resultado acima ainda podemos afirmar mais, que todos os determinantes de sub-matrizes $k \times k, \quad k > p$ serão nulos pois terão pelo menos uma linha nula.

Um último ponto em relação às operações elementares é que qualquer matriz de posto p pode ser reduzida, através de operações elementares de linha e ou coluna, à forma:

$$\left(\begin{array}{c|c}
I_{p\times p} & O_{p\times(n-p)} \\
\hline
O_{(m-p)\times p} & O_{(m-p)\times(n-p)}
\end{array}\right)$$

Ou seja, a matriz identidade no canto superior esquerdo e matriz nulas nos outros 3 quadrantes. É fácil perceber como proceder para isso. Primeiro coloca-se a matriz na forma escada com operações elementares de linhas, depois com operações de colunas leva-se a matriz para a forma:

$$\left(\begin{array}{cccc|ccc}
1 & 0 & 0 & 0 & c_{11} & \cdots & c_{1(n-p)} \\
0 & 1 & 0 & 0 & c_{21} & \cdots & c_{2(n-p)} \\
0 & 0 & 1 & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & c_{p1} & \cdots & c_{p(n-p)} \\
\hline
0 & 0 & 0 & \cdots & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0
\end{array}\right)$$

Agora utilzamos os 1´s da diagonal para anular via subtração de colunas todos os termos $c_{ij} \neq 0$ da matriz $C_{p\times(n-p)}$ que sobrou.

Sistemas de equações lineares com n incógnitas e n equações. São sistemas que podem ser escritos na forma:

$A\vec{x} = \vec{b}$ com A sendo uma matriz $n \times n$. Se A é não singular, i.e., $\det A \neq 0$ então A admite inversa e podemos calcular o vetor $\vec{x}$ da única solução imediatamente através de $\vec{x} = A^{-1}\vec{b}$. Em particular se a equação é homogênea então só existe a solução trivial $\vec{x} = A^{-1}\vec{0} = \vec{0}$. Se A é singular significa que posto de A é menor do que n e teremos soluções diferentes da trivial.

Regra de Cramer:

Note que para matrizes dos coeficientes não singular obtivemos $\vec{x} = A^{-1}\vec{b}$. Por outro lado $A^{-1} = \frac{A^\dagger}{\det A} = \frac{1}{\det A}\begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{nn} \end{pmatrix}$ e o vetor $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ np é tal que cada componente é dada por $v_i = \sum_k b_k \Delta_{ki} = \det\begin{pmatrix} A_{11} & A_{12} & \cdots & b_1 & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & b_2 & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & b_n & \cdots & A_{nn} \end{pmatrix}$ de onde extraímos a regra de Cramer:

$$x_i = \frac{\det\begin{pmatrix} A_{11} & A_{12} & \cdots & b_1 & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & b_2 & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & b_n & \cdots & A_{nn} \end{pmatrix}}{\det\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1i} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2i} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{ni} & \cdots & A_{nn} \end{pmatrix}}$$

A regra é: troca-se a iésima coluna da matriz do numerador pelo vetor $\vec{b}$ e divide-se o detrminante dessa matriz pelo determinante da matriz dos coeficientes.

Exemplo:

$$\begin{matrix} 2x+3y=2 \\ 3x+2y=1 \end{matrix} \quad \text{então} \quad A=\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \quad \text{e} \quad \vec{b}=\begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{então} \quad x=\frac{\det\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}}{\det\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}}=\frac{4-3}{4-9}=-\frac{1}{5} \quad \text{e}$$

$$y=\frac{\det\begin{pmatrix} 2 & 2 \\ 3 & 1 \end{pmatrix}}{\det\begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}}=\frac{2-6}{4-9}=\frac{4}{5} \quad \text{então} \quad \begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} -{}^{1}\!/_{5} \\ {}^{4}\!/_{5} \end{pmatrix}$$ é a solução. Recalculando através da inversa:

$$A^{-1}=\frac{1}{(4-9)}\begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix}=\begin{pmatrix} -{}^{2}\!/_{5} & {}^{3}\!/_{5} \\ {}^{3}\!/_{5} & -{}^{2}\!/_{5} \end{pmatrix} \quad \text{então} \quad \begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} -{}^{2}\!/_{5} & {}^{3}\!/_{5} \\ {}^{3}\!/_{5} & -{}^{2}\!/_{5} \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}=\begin{pmatrix} -{}^{2}\!/_{5} \\ {}^{4}\!/_{5} \end{pmatrix}.$$

Combinações lineares:

Equações paramétricas.

Suponha que exista uma curva no plano x-y, ou no espaço $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. A equação do LUGAR GEOMÉTRICO da curva é uma relação do tipo, $R(x,y)=0$ entre as variáveis x e y. Exemplo: círculo centrado na origem possui a relação $x^2+y^2=R^2$, ou $x^2+y^2-R^2=0$, onde R é o raio do círculo. Vamos tentar extrair dessa relação o y em função de x. Mas note que $y=\pm\sqrt{R^2-x^2}$ não é uma função $y(x):\mathbb{R}\to\mathbb{R}$ porque admite dois valores para cada x. Uma função é uma regra de associação entre todos os elementos de um conjunto chamado de domínio com elementos de outro conjunto chamado contra-domínio e não admite dúvida sobre a qual elemento do contra-domínio devemos associar cada elemento do domínio. Caso contrário seria impossível obedecer a ordem associe os elementos através da regra – o operador ficaria em dúvida sobre qual elemento associar.

A idéia da equação paramétrica [de parâmetro] é usar um parâmetro t em termos do qual as funções $x(t):\mathbb{R}\to\mathbb{R}$ e $y(t):\mathbb{R}\to\mathbb{R}$ existem sendo bem definidas. Podemos agora fazer um gráfico de y vs x da seguinte forma: construímos uma coluna com os valores de t, com ele calculamos $x(t)$ e $y(t)$ e no final fazemos o gráfico y vs x esquecendo o t. Criamos uma função $\left[x(t),\ y(t)\right]:\mathbb{R}\to\mathbb{R}^2$. Para obter a relação do lugar geométrico é necessário eliminar o parâmetro. Vejamos o exemplo do círculo:

$x(t)=R\cos t$ e $y(t)=R\sin t$. Agora elevando as duas ao quadrado e somando obtemos $x^2+y^2=R^2\left[\cos^2 t+\sin^2 t\right]=R^2$, eliminando o parâmetro t e ficando com a equação do lugar gemétrico dos pontos $x^2+y^2=R^2$.

Outro exemplo interessante é o da elipse: $x(t)=a\cos t$ e $y(t)=b\sin t$. Agora somamos a equação $\frac{x^2}{a^2}=\cos^2 t$ com $\frac{y^2}{a^2}=\sin^2 t$ para obter $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$.

Com um parâmetro podemos descrever uma curva, com dois uma superfície, com três um volume e assim por diante. Note que se mantenho R fixo nas equações $x(t)=R\cos t$ e $y(t)=R\sin t$ consigo gerar um círculo. Mas se ganhamos a liberdade de variar R de podemos preencher toda a área de um círculo, de um anel e até todo o plano x-y.

Equação paramétrica da reta.

Tome um vetor $\vec{V}$ e o multiplique por t. Variando t de $-\infty,+\infty$ podemos gerar todos os pontos ao longo da reta na direção do vetor $\vec{V}$. Note que estamos usando as equações paramétricas: $\left[x(t),y(t)\right]=t\vec{V}=t\left[V_1,V_2\right]$, ou seja, $x(t)=tV_1$ e $y(t)=tV_2$. Eliminando o parâmetro obtemos $\frac{y}{x}=\frac{V_2}{V_1}$ que é uma reta que passa no origem. Já a equação $\vec{X}=\vec{X}_o+t\vec{V}$, $x(t)=x_o+tV_1$ e $y(t)=y_o+tV_2$ é uma reta que passa no ponto (x_o,y_o) com a equação do lugar geométrico dada por: $\frac{y-y_o}{x-x_o}=\frac{V_2}{V_1}$.

1. Mostre que se o sistema homogêneo de n equações e n incógnitas:

$$\mathbf{A_{11}\, t_1 + A_{12}\, t_2 + \ldots + A_{1n}\, t_n = 0}$$

$$\mathbf{A_{21}\, t_1 + A_{22}\, t_2 + \ldots + A_{2n}\, t_n = 0}$$

$$\vdots$$

$$\mathbf{A_{21}\, t_1 + A_{22}\, t_2 + \ldots + A_{2n}\, t_n = 0}$$

admite solução diferente da trivial, i.e, $(t_1,t_2,\cdots,t_n)\neq(0,0,\cdots,0)$ então det(**A**) = 0.

2. Prove a regra de Cramer $x_i = \dfrac{\begin{bmatrix} A_{11} & \cdots & B_1 & \cdots A_{1n} \\ A_{21} & \cdots & B_2 & \cdots A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & \cdots & B_n & \cdots A_{nn} \end{bmatrix}}{\begin{bmatrix} A_{11} & \cdots & A_{1i} & \cdots A_{1n} \\ A_{21} & \cdots & A_{2i} & \cdots A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & \cdots & A_{ni} & \cdots A_{nn} \end{bmatrix}}$ para as soluções do sistema de n equações e n incógnitas:

$$\mathbf{A_{11}\, t_1 + A_{12}\, t_2 + \ldots + A_{1n}\, t_n = B_1}$$

$$\mathbf{A_{21}\, t_1 + A_{22}\, t_2 + \ldots + A_{2n}\, t_n = B_2}$$

$$\vdots$$

$$\mathbf{A_{21}\, t_1 + A_{22}\, t_2 + \ldots + A_{2n}\, t_n = B_n}$$

em que **A** é NÃO SINGULAR.

Espaços Vetoriais

Seja $\Im$ o conjunto de todos os números reais positivos. Defina a operação adição $\oplus$ pelo produto usual de dois números, i.e., $\vec{\mathbf{u}} \oplus \vec{\mathbf{v}} = \mathrm{u\,v}$. Seja $\alpha \in \Re$ um escalar e defina a operação multiplicação por um escalar através da regra $\alpha \otimes \vec{\mathbf{u}} = \mathrm{u}^{\alpha}$. Mostre que $\Im$ é um espaço vetorial. Qual a necessidade da restrição para valores positivos?

Se $\Im$ e $\aleph$ são espaços vetoriais mostre que $\Im \cap \aleph$ também é um espaço vetorial.

Considere o conjunto das matrizes $m \times n$ com as operações adição e multiplicação por escalar matriciais usuais. Mostre que esse conjunto é um espaço vetorial. Quem é o elemento nulo?

Considere o conjunto das matrizes 2×2 e as operações matriciais usuais. Encontre uma base para esse espaço. Qual a dimensão desse espaço? O conjunto de matrizes 2×2 simétricas é um subespaço vetorial? Encontre uma base para esse espaço. Qual a dimensão desse sub-espaço?

Sitemas superdeterminados.

Tome o caso em que $\sum_{j=1}^{n} a_{ij} x_j = b_i \quad i \in [1,\cdots,m] \quad m > n$. Nesse caso temos mais equações m do que incógnitas. Em termos matriciais temos que $A_{m\times n}\,\vec{x}_{1\times n} = \vec{b}_{1\times m}$. Mesmo que o sistema seja inconsistente, de modo que $A_{m\times n}\,\vec{x}_{1\times n} = \vec{b}_{1\times m}$ não admite solução, queremos achar os valores $\vec{x}^{*}_{1\times n}$ para os quais de a soma dos desvios, i.e., se $A_{m\times n}\,\vec{x}^{*}_{1\times n} = \vec{b}^{*}_{1\times m}$ então $\left\|\vec{b}^{*}_{1\times m} - \vec{b}_{1\times m}\right\|^2 = \min$. A norma de um vetor é dada por $\left\|\vec{V}\right\| = \sqrt{\sum_i V_i^2} = \sqrt{\overline{\vec{V}}\vec{V}}$. Então $\vec{b}^{*} - \vec{b} = A\vec{x}^{*} - \vec{b}$ logo $\overline{\vec{b}^{*} - \vec{b}} = \overline{A\vec{x}^{*} - \vec{b}} = \overline{A\vec{x}^{*}} - \overline{\vec{b}} = \overline{\vec{x}^{*}}\,\overline{A} - \overline{\vec{b}}$ logo $\left(\overline{\vec{b}^{*} - \vec{b}}\right)\left(\vec{b}^{*} - \vec{b}\right) = \left(\overline{\vec{x}^{*}}\,\overline{A} - \overline{\vec{b}}\right)\left(A\vec{x}^{*} - \vec{b}\right) = \overline{\vec{x}^{*}}\,\overline{A}A\vec{x}^{*} - \overline{\vec{x}^{*}}\,\overline{A}\vec{b} - \overline{\vec{b}}A\vec{x}^{*} + \overline{\vec{b}}\vec{b}$.

Agora note que $\overline{\vec{b}}_{1\times m} A_{m\times n} \vec{x}_{n\times 1} = \alpha_{1\times 1}$ é um número, um escalar, logo, $\overline{\alpha} = \alpha$ então $\overline{\vec{b}}A\vec{x} = \overline{\overline{\vec{b}}A\vec{x}} = \overline{\vec{x}}\,\overline{A}\,\overline{\overline{\vec{b}}} = \overline{\vec{x}}\,\overline{A}\vec{b}$, logo a nossa equação é dada por:

$$Min\left[\overline{\vec{x}^{*}}\,\overline{A}A\vec{x}^{*} - 2\overline{\vec{x}^{*}}\,\overline{A}\vec{b} + \overline{\vec{b}}\vec{b}\right]$$

Vamos escreve-la na forma mais adequada à diferenciação:

$$Min\left\{\sum_i\sum_j x_i^{*}\left(\overline{A}A\right)_{ij} x_j^{*} - 2\sum_i\sum_j x_i^{*}\overline{A}_{ij} b_j + b^2\right\}$$ onde $b^2 = \left\|\vec{b}\right\|^2$.

A solução é dada quando o gradiente da função de $\vec{x}$ é nulo, ou seja:

$$\frac{\partial}{\partial x_k^{*}}\left\{\sum_i\sum_j x_i^{*}\left(\overline{A}A\right)_{ij} x_j^{*} - 2\sum_i\sum_j x_i^{*}\overline{A}_{ij} b_j + b^2\right\} = 0 \quad \forall k \in [1,\cdots,n].$$

Mas nesse caso temos que:

$$\frac{\partial}{\partial x_k^*}\left\{\sum_i\sum_j x_i^*\left(\overline{A}A\right)_{ij}x_j^* - 2\sum_i\sum_j x_i^* A_{ij}b_j + b^2\right\} =$$

$$= \sum_i\sum_j x_i^*\left(\overline{A}A\right)_{ij}\frac{\partial}{\partial x_k^*}x_j^* + \sum_i\sum_j\left(\frac{\partial}{\partial x_k^*}x_i^*\right)\left(\overline{A}A\right)_{ij}x_j^* - 2\sum_i\sum_j\left(\frac{\partial}{\partial x_k^*}x_i^*\right)A_{ij}b_i =$$

$$= \sum_i\sum_j x_i^*\left(\overline{A}A\right)_{ij}\delta_{jk} + \sum_i\sum_j\delta_{ik}\left(\overline{A}A\right)_{ij}x_j^* - 2\sum_i\sum_j\delta_{ik}A_{ij}b_i =$$

$$= \sum_i x_i^*\left(\overline{A}A\right)_{ik} + \sum_j\left(\overline{A}A\right)_{kj}x_j^* - 2\sum_i A_{kj}b_j$$

Agora $\overline{A}\,A$ é simétrica, logo:

$$\sum_i x_i^*\left(\overline{A}A\right)_{ik} + \sum_j\left(\overline{A}A\right)_{kj}x_j^* - 2\sum_i \overline{A}_{kj}b_j = \sum_i x_i^*\left(\overline{A}A\right)_{ki} + \sum_j\left(\overline{A}A\right)_{kj}x_j^* - 2\sum_i \overline{A}_{kj}b_j =$$

$$= \sum_j\left(\overline{A}A\right)_{kj}x_j^* + \sum_j\left(\overline{A}A\right)_{kj}x_j^* - 2\sum_i \overline{A}_{kj}b_j = 2\sum_j\left(\overline{A}A\right)_{kj}x_j^* - 2\sum_i \overline{A}_{kj}b_j$$

Então note que $2\sum_j\left(\overline{A}A\right)_{kj}x_j^* - 2\sum_i \overline{A}_{kj}b_j = 0$ implica na equação $\sum_j\left(\overline{A}A\right)_{kj}x_j^* = \sum_i \overline{A}_{kj}b_j$ que pode ser escrita na forma matricial como $\left(\overline{A}\,A\right)\vec{x}^* = \overline{A}\vec{b}$. Então a $\vec{x}^* = \left(\overline{A}\,A\right)^{-1}\overline{A}\vec{b}$ embora não seja uma solução de $A\vec{x} = \vec{b}$ minimiza o desvio $\delta = \left\|A\vec{x}^* - \vec{b}\right\|^2$. Se $\delta = 0$ então $\vec{x}^* = \left(\overline{A}\,A\right)^{-1}\overline{A}\vec{b}$ é uma solução de $A\vec{x} = \vec{b}$.

Regressão linear:

Problema de autovalores e autovetores:

O problema de diagonalização de matrizes é resolvido através do problema de autovalor. Problema do auto-valor de uma matriz M quadrada: queremos saber se existem autovetores $\vec{u}_i$ e autovalores λ_i tais que $M\vec{u}_i = \lambda_i \vec{u}_i$.

A forma canônica com que esse problema é resolvido é escrever a equação na forma do sistema de equações lineares $(M-\lambda I)\vec{u} = \vec{0}$ que só admite solução diferente da trivial $\vec{u} = \vec{0}$ se $\det(M-\lambda I) = 0$. Note que isso leva a uma equação de grau n pois $\det\begin{pmatrix} M_{11}-\lambda & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22}-\lambda & \cdots & M_{2n} \\ \vdots & \vdots & & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn}-\lambda \end{pmatrix} = 0$ e não é possível se afirmar à priori, mesmo se M for uma matriz real, que as raízes, correspondentes aos autovalores, serão reais ou complexas.

Entretanto se M for real e simétrica podemos afirmar com certeza que os autovalores serão reais. Poderíamos relaxar afirmando que se M for Hermitiana os autovalores são reais.

Matrizes Hermitianas

Uma matriz M é Hermitiana se $M^{\dagger} = M$, ou seja, $M^{*\prime} = M$. Se M é real e Hermitiana então M é simplesmente simétrica, $M' = M$. Ou seja, matrizes reais simétricas são Hermitianas – um sub-conjunto das matrizes Hermitianas. Qual a propriedade interessante das matrizes Hermitianas?

1. Os autovalores são reais
2. Os autovetores são ortogonais

Prova: Temos que $M\,\vec{u}_i = \lambda_i \vec{u}_i$ e $M\,\vec{u}_j = \lambda_j \vec{u}_j$. Aplicando o operador adjunto na segunda equação $\left(M\,\vec{u}_j\right)^{\dagger} = \vec{u}_j^{\dagger} M^{\dagger} = \lambda_j^* \vec{u}_j^{\dagger}$ como $M^{\dagger} = M$ então $\vec{u}_j^{\dagger} M = \lambda_j^* \vec{u}_j^{\dagger}$. Agora multiplicamos a primeira equação por $\vec{u}_j^{\dagger}$ obtendo $\vec{u}_j^{\dagger} M\,\vec{u}_i = \lambda_i \vec{u}_j^{\dagger} \vec{u}_i$ e a última por $\vec{u}_i$ obtendo $\vec{u}_j^{\dagger} M\,\vec{u}_i = \lambda_j^* \vec{u}_j^{\dagger} \vec{u}_i$. Subtraindo uma da outra temos que $\left(\lambda_i - \lambda_j^*\right)\vec{u}_j^{\dagger}\vec{u}_i = 0$. Se $i = j$ então $\vec{u}_i^{\dagger}\vec{u}_i > 0$ e a igualdade só é satisfeita se $\lambda_i = \lambda_i^*$, o que significa que λ_i é real. Por outro lado se $i \neq j$, $\lambda_i \neq \lambda_j$ [caso não degenerado] é preciso que $\vec{u}_j \cdot \vec{u}_i = \vec{u}_j^{\dagger}\vec{u}_i = 0$, o que significa que os autovetores são ortogonais entre si. Se M é real e simétrica os autovalores e autovetores serão reais.

Agora, se além de real e simétrica M for definida positiva então os autovalores, além de reais, são obrigatoriamente positivos, pois $\vec{u}_i' M\,\vec{u}_i = \lambda_i \vec{u}_i \cdot \vec{u}_i > 0$ e como $\vec{u}_i \cdot \vec{u}_i > 0$ sempre, então $\lambda_i > 0$.

Existe uma arbitrariedade na escolha do autovetor uma vez que é possível multiplicar a equação $M\,\vec{u}_i = \lambda_i \vec{u}_i$ por uma constante de ambos os lados sem perda de validade. Já que se pode multiplicar por uma constante qualquer pode-se multiplicar pela constante que garanta que $\vec{u}_i \cdot \vec{u}_i = 1$ e $\vec{u}_i \cdot \vec{u}_j = \delta_{ij}$. Neste caso geramos uma base ortonormal – ortogonal normalizada à 1. Todos os autovetores são unitários, i.e., possuem norma igual a 1.

Diagonalização de Matrizes:

Vamos considerar que encontramos todos os autovalores da base ortonormal de autovetores de modo que $M\,\vec{u}_i = \lambda_i \vec{u}_i$ e $\vec{u}_i \cdot \vec{u}_j = \delta_{ij}$. O que acontece se realizamos a seguinte operação:

$\begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix} M \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix}$. Note que nossa notação para o símbolo de vetor é de uma matriz coluna logo $\vec{u}_j = \begin{pmatrix} u_{j1} \\ u_{j2} \\ \vdots \\ u_{jn} \end{pmatrix}$ e que $\begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix}$ e $\begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix}$ são matrizes $n \times n$. Na primeira os vetores estão alinhados na forma de linhas e na segunda na forma de colunas. Além disso note que $\begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix} = \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix}'$. Aplicando M a cada um dos autovetores normalizados teremos:

$\begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix} M \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix} = \begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix} \begin{pmatrix} \lambda_1 \vec{u}_1 & \lambda_2 \vec{u}_2 & \cdots & \lambda_n \vec{u}_n \end{pmatrix}$ e, finalmente, que $\begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix} M \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \vec{u}_1 \cdot \vec{u}_1 & \lambda_2 \vec{u}_1 \cdot \vec{u}_2 & \cdots & \lambda_n \vec{u}_1 \cdot \vec{u}_n \\ \lambda_1 \vec{u}_2 \cdot \vec{u}_1 & \lambda_2 \vec{u}_2 \cdot \vec{u}_2 & \cdots & \lambda_n \vec{u}_2 \cdot \vec{u}_n \\ \vdots & \vdots & & \vdots \\ \lambda_1 \vec{u}_n \cdot \vec{u}_1 & \lambda_2 \vec{u}_n \cdot \vec{u}_2 & \cdots & \lambda_n \vec{u}_n \cdot \vec{u}_n \end{pmatrix}$. Agora usamos o fato de que os vetores são ortonormais para chegar ao resultado $\begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix} M \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$. Ou seja a matriz M foi diagonalizada. A matriz que a diagonaliza é $S = \begin{pmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{pmatrix} = \begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} \\ \vdots & \vdots & & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{pmatrix}$ onde o primeiro índice se refere ao autovetor e o segundo às suas componentes. Note que $S' = \begin{pmatrix} \vec{u}_1 \\ \vec{u}_2 \\ \vdots \\ \vec{u}_n \end{pmatrix}$ e que $S'S = I$, ou seja, $S' = S^{-1}$. São chamadas matrizes ortogonais. Aplicando o determinante $\det(S'S) = \det(I)$, e usando o fato de que $\det(AB) = \det(A)\det(B)$, $\det(A') = \det(A)$ e $\det(I) = 1$ vemos que $\det^2(S) = 1$ que nos leva a $\det(S) = 1$. Na realidade $\det(S) = \pm 1$, sendo 1 em rotações e -1 em operações que envolvem reflexões, que não é o caso aqui.

Então a operação $S'M\,S = D$ transforma a matriz M em uma matriz diagonal. Para voltar da diagonal à M basta inverter S´com S: $S'M\,S = D \quad \rightarrow \quad SS'M\,SS' = S\,D\,S' \quad \rightarrow \quad M = S\,D\,S'$.

Transformação de similaridade:

Quando transformamos uma matriz segundo a forma: $M' = S^{-1}M\,S$ dizemos que fizemos uma transformação de similaridade. A transformação de similaridade preserva o determinante e o traço da matriz. A preservação do determinantes é demonstrada facilmente pois $\det(M') = \det(S^{-1}M\,S) = \det(S^{-1}S)\det(M) = \det(M)$. Já o traço é dado pela soma dos elementos da diagonal $tr(M) = \sum_i M_{ii}$ e tem as seguinte propriedades: $tr(A+B) = tr(A) + tr(B)$, $tr(cA) = c\,tr(A)$ e $tr(AB) = tr(BA)$. As duas primeiras são triviais. A última é demonstrada da seguinte forma: $(AB)_{ij} = \sum_k A_{ik}B_{kj}$ e $tr(AB) = \sum_i \sum_k A_{ik}B_{ki} = \sum_k \sum_i B_{ki}A_{ik} = tr(BA)$. Com isso a preservação do traço pode ser facilmente demonstrada $tr(S^{-1}M\,S) = tr(M\,SS^{-1}) = tr(M)$. Se a matriz M foi diagonalizada para D e o determinante e o traço são preservados então $\det(M) = \lambda_1\lambda_2\cdots\lambda_n$ e o $tr(M) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Diagonalização da Matriz inversa:

Podemos mostrar que a mesma matriz S que diagonaliza a matriz M também diagonaliza sua inversa. Sabemos que $MM^{-1} = I$ e $S'MS = D$. Inserindo $I = SS'$ dos dois lados de M: $SS'MSS'M^{-1} = I$, e usando o fato de que M se torna diagonal obtemos $SDS'M^{-1} = I$. Agora multiplicamos à direita e à esquerda pela inversa de S ou S´ para obter $S'SDS'M^{-1}S = S'IS$, que se transforma em $D(S'M^{-1}S) = I$, o que significa que $(S'M^{-1}S) = D^{-1}$. Entretanto se $D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ então $D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n} \end{pmatrix}$ e também é diagonal. Ou seja, a transformação S também diagonaliza a matriz inversa e, além disso, os autovalores da matriz inversa são os recíprocos da direta.

Em resumo se λ_i é um autovalor de M com autovetor $\vec{u}_i$ então $\frac{1}{\lambda_i}$ é o autovalor de M^{-1} correspondendo ao meso autovetor $\vec{u}_i$.

$$\det\begin{pmatrix} -x & 0 & 0 & \cdots & -a_o \\ 1 & -x & 0 & \cdots & -a_1 \\ 0 & 1 & -x & \cdots & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & (-a_{n-1}-x) \end{pmatrix} = -x\det\begin{pmatrix} -x & 0 & 0 & \cdots & -a_1 \\ 1 & -x & 0 & \cdots & -a_2 \\ 0 & 1 & -x & \cdots & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & (-a_{n-1}-x) \end{pmatrix} -$$

$$A_{ij} = -x\delta_{ij} - a_{n-1}\delta_{in}\delta_{jn} + \delta_{i(j+1)} - a_{i-1}\delta_{jn}$$