4.4.1 Brincando com a característica

Nesta secção vamos apresentar alguns resultados importantes que se podem deduzir facilmente à custa de $car(A) + \text{nul}(A) = n$, onde A é uma matriz $m \times n$. Pressupõe-se que B é uma matriz tal que AB existe.

1. $car(AB) \leq car(A)$.
   Como vimos na secção anterior, $CS(AB) \subseteq CS(A)$, pelo que $\dim CS(AB) \leq \dim CS(A)$.

2. Se B é invertível então $car(A) = car(AB)$.

3. $N(B) \subseteq N(AB)$.
   Se $b \in N(B)$ então $Bb = 0$. Multiplicando ambos os lados, à esquerda, por A obtemos $ABb = 0$, pelo que $b \in N(AB)$.

4. $\text{nul}(B) \leq \text{nul}(AB)$.

5. $N(A^TA) = N(A)$.
   Resta mostrar que $N(A^TA) \subseteq N(A)$. Se $x \in N(A^TA)$ então $A^TAx = 0$. Multiplicando ambos os lados, à esquerda, por x^T obtemos $x^TA^TAx = 0$, pelo $(Ax)^TAx = 0$. Seja $(y_1, \ldots, y_n) = y = Ax$. De $y^Ty = 0$ obtemos $y_1^2 + y_2^2 + \ldots y_n^2 = 0$. A soma de reais não negativos é zero se e só se cada parcela é nula, pelo que cada $y_i^2 = 0$, e portanto $y_i = 0$. Ou seja, $y = 0$, donde segue que $Ax = 0$, ou seja, que $x \in N(A)$.

6. $\text{nul}(A^TA) = \text{nul}(A)$.

7. $car(A^TA) = car(A) = car(AA^T)$.
   De $car(A) + \text{nul}(A) = n = car(A^TA) + \text{nul}(A^TA)$ e $\text{nul}(A^TA) = \text{nul}(A)$ segue que $car(A^TA) = car(A)$. Da mesma forma, $car(A^T) = car(AA^T)$. Como $car(A) = car(A^T)$, obtemos $car(A) = car(AA^T)$.

8. Se $car(A) = n$ então A^TA é invertível.
   A^TA é uma matriz $n \times n$ com característica igual a n, pelo que é uma matriz não-singular, logo invertível.

4.4.2 Aplicação a sistemas impossíveis

Como motivação, suponha que se quer encontrar (caso exista) a recta r de $\mathbb{R}^2$ que incide nos pontos $(-2, -5), (0, -1), (1, 1)$. Sendo a recta não vertical, terá uma equação da forma $y = mx + c$, com $m, c \in \mathbb{R}$. Como r incide nos pontos indicados, então necessariamente

$$-5 = m \cdot (-2) + c,\ -1 = m \cdot 0 + c,\ 1 = m \cdot 1 + c.$$

A formulação matricial deste sistema de equações lineares (nas incógnitas m e c) é

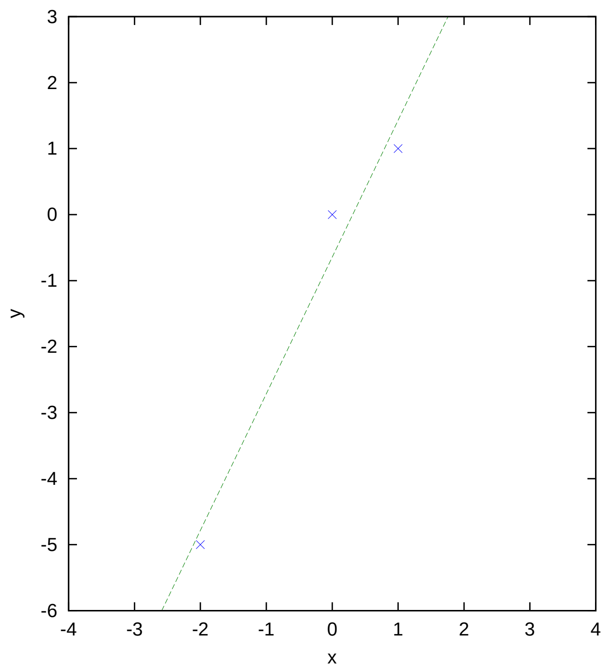
$$\begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \\ 1 \end{bmatrix}.$$

O sistema é possível determinado, pelo que a existência da recta e a sua unicidade está garantida. A única solução é $(m, c) = (2, 1)$ e portanto a recta tem equação $y = 2x - 1$.

No entanto, se considerarmos como dados os pontos $(-2, -5), (0, 0), (1, 1)$, facilmente chegaríamos à conclusão que não existe uma recta incidente nos três pontos. Para tal, basta mostrar que o sistema de equações dado pelo problema (tal como fizemos no caso anterior) é impossível. Obtemos a relação

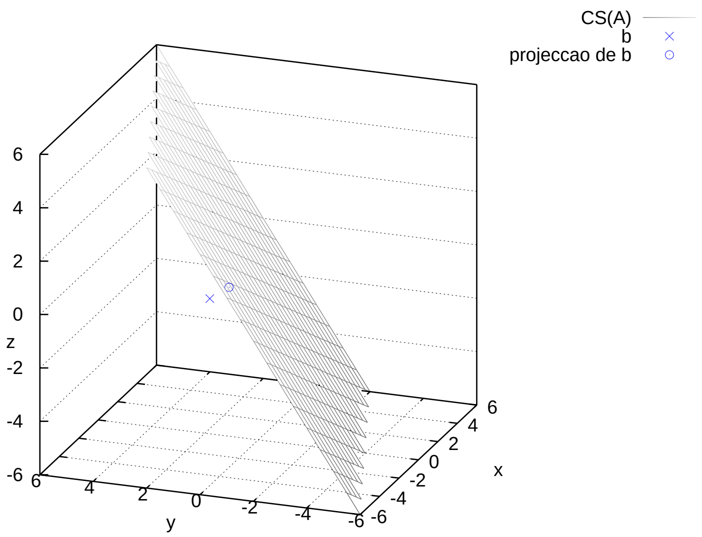
$$b \notin CS(A),$$

onde $A = \begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ e $b = \begin{bmatrix} -5 \\ 0 \\ 1 \end{bmatrix}$. Suponha que os pontos dados correspondem a leituras de uma certa experiência, pontos esses que, teoricamente, deveriam ser colineares. Ou seja, em algum momento houve um desvio da leitura em relação ao que se esperaria. Desconhece-se qual ou quais os pontos que sofreram incorrecções. Uma solução seria a de negligenciar um dos pontos e considerar os outros dois como correctos. É imediato concluir que este raciocínio pode levar a conclusões erróneas. Por exemplo, vamos pressupor que é o primeiro dado que está incorrecto (o ponto $(-2, -5)$). A rectas que passa pelos pontos $(0, 0), (1, 1)$ tem como equação $y = x$. Ora se o erro esteve efectivamente na leitura do ponto $(0, 0)$ (que deveria ser $(0, -1)$) então o resultado correcto está bastante distante do que obtivémos. O utilizador desconhece qual (ou quais, podendo haver leituras incorrectas em todos os pontos) dos dados sofreu erros. Geometricamente, a primeira estratégia corresponde a eliminar um dos pontos e traçar a recta que incide nos outros dois. Uma outra que, intuitivamente, parece a mais indicada, será a de, de alguma forma e com mais ou menos engenho, traçar uma recta que se tente *aproximar o mais possível* de todos os pontos, ainda que não incida em nenhum deles!

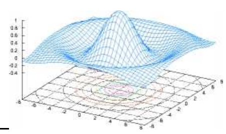


Vamos, de seguida, usar todo o engenho que dispomos para encontrar a recta que se aproxima o mais possível dos pontos $(-2,-5),(0,0),(1,1)$.

Sabendo que $b \notin CS(A)$, precisamos de encontrar $b' \in CS(A)$ por forma a que b' seja o ponto de $CS(A)$ mais próximo de b. Ou seja, pretendemos encontrar $b' \in CS(A)$ tal que $d(b,b') = \min_{c \in CS(A)} d(c,b)$, onde $d(u,v) = \|u-v\|$. O ponto b' é o de $CS(A)$ que minimiza a distância a b. Este ponto b' é único e é tal que $b-b'$ é ortogonal a todos os elementos de $CS(A)$. A b' chamamos *projecção ortogonal* de b sobre (ou ao longo) de $CS(A)$, e denota-se por $proj_{CS(A)} b$.



Apresentamos, de seguida, uma forma fácil de cálculo dessa projecção, quando as colunas de A são linearmente independentes. Neste caso, A^TA é invertível e a projecção de b sobre $CS(A)$ é dada por
$$b' = A(A^TA)^{-1}A^Tb.$$



Octave

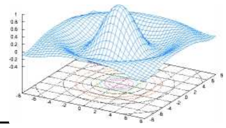
Aconselhamos que experimente o código seguinte no *octave* e que manipule o gráfico por forma a clarificar que $b \notin CS(A)$:

```
x=[-2;0;1]; y=[-5; 0; 1];b=y;
A=[x [1;1;1]]
alfa=-6:0.5:6;beta=alfa;
[AL,BE]=meshgrid (alfa,beta);
Z=0.5*(-AL+3*BE);
mesh(AL,BE,Z)
hold on
```

```
plot3([-5],[0],[1],'x')
projb=A*inv(A'*A)*A'*b;
plot3([projb(1,1)],[projb(2,1)],[projb(3,1)],'o')
axis ([-6, 6,-6 , 6, -6,6], "square")
legend('CS(A)','b','projeccao de b' );
xlabel ('x');ylabel ('y');zlabel ('z');
```

Calculou-se `projb` a projecção de b sobre $CS(A)$.

Pretendemos agora encontrar x por forma a que $Ax = b'$, ou seja, x por forma a que a distância de Ax a b seja a menor possível. Repare que se $Ax = b$ é impossível, então essa distância será, seguramente, não nula. A equação $Ax = b'$ é sempre possível, já que $b' = A(A^T A)^{-1} A^T b \in CS(A)$; ou seja, b' escreve-se como Aw, para algum w (bastando tomar $w = (A^T A)^{-1} A^T b$). No entanto, o sistema pode ser indeterminado, e nesse caso poderá interessar, de entre todas as soluções possíveis, a que tem norma mínima. O que acabámos por expôr, de uma forma leve e ingénua, denomina-se o *método dos mínimos quadrados*, e a x solução de $Ax = b'$ de norma minimal, denomina-se a solução no sentido dos mínimos quadrados de norma minimal.



Octave

Vamos agora mostrar como encontrámos a recta que melhor se ajustava aos 3 pontos apresentados no início desta secção.

```
x=[-2;0;1]; y=[-5; 0; 1];b=y;
A=[x [1;1;1]]
xx=-6:0.5:6; solminq=inv(A'*A)*A'*b
yy=solminq(1,1)*xx+solminq(2,1);
plot(x,y,'x',xx,yy); xlabel ('x');ylabel ('y');
```

Para mudar as escalas basta fazer `set(gca,"XLim",[-4 4]); set(gca,"YLim",[-6 6])`, ou em alternativa `axis ([-4, 4,-6 , 3], "square");`. Para facilitar a leitura dos pontos, digite `grid on`.

Uma forma alternativa de encontrar x solução de $Ax = proj_{CS(A)} b$ seria `solminq=A\b` em vez de `solminq=inv(A'*A)*A'*b`.

Ao invés de procurarmos a recta que melhor se adequa aos dados disponíveis, podemos procurar o polinómio de segundo, terceiro, etc, graus. Se os dados apresentados forem pontos de $\mathbb{R}^3$, podemos procurar o plano que minimiza as somas das distâncias dos pontos a esse plano. E assim por diante, desde que as funções que definem a curva ou superfície sejam lineares nos parâmetros. Por exemplo, $ax^2 + bx + c = 0$ não é uma equação linear em x mas é-o em a e b.

No endereço `http://www.nd.edu/~powers/ame.332/leastsquare/leastsquare.pdf` pode encontrar informações adicionais sobre o método dos mínimos quadrados e algumas aplicações.

Em `enacit1.epfl.ch/cours_matlab/graphiques.html` pode encontrar alguma descrição das capacidades gráficas do *Gnu-octave* recorrendo ao *GnuPlot*.

Exemplo 4.4.6. O exemplo que de seguida apresentamos baseia-se no descrito em [3, pag.58]

Suponha que se está a estudar a cinética de uma reacção enzimática que converte um substrato S num produto P, e que essa reacção segue a equação de Michaelis-Menten,

$$r = \frac{k_2[E]_0[S]}{K_m + [S]},$$

onde

1. $[E]_0$ indica concentração enzimática original adicionada para iniciar a reacção, em gramas de E por litro,

2. r é o número de gramas de S convertido por litro por minuto (ou seja, a velocidade da reacção),

3. k_2 é o número de gramas de S convertido por minuto por grama de E.

Depois de se efectuar uma série de experiências, obtiveram-se os dados apresentados na tabela seguinte, referentes à taxa de conversão de gramas de S por litro por minuto:

$[S]$ g s/l	$[E]_0 = 0.005$ g$_E$/l	$[E]_0 = 0.01$ g$_E$/l
1.0	0.055	0.108
2.0	0.099	0.196
5.0	0.193	0.383
7.5	0.244	0.488
10.0	0.280	0.569
15.0	0.333	0.665
20.0	0.365	0.733
30.0	0.407	0.815

Re-escrevendo a equação de Michaelis-Menten como

$$\frac{[E]_0}{r} = \frac{K_m}{k_2}\frac{1}{[S]} + \frac{1}{k_2},$$

obtemos um modelo linear

$$y = b_1 x + b_0$$

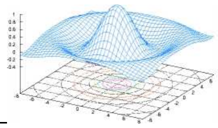
com

$$y = \frac{[E]_0}{r},\ x = \frac{1}{[S]},\ b_0 = \frac{1}{k_2},\ b_1 = \frac{K_m}{k_2}.$$

Denotemos os dados x e y por x_i e y_i, com $i = 1, \ldots, 8$. Este sistema de equações lineares tem a representação matricial

$$A \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} = y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_8 \end{bmatrix}$$

com $A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_8 & 1 \end{bmatrix}$. A única solução de $A^T A \begin{bmatrix} b_1 \\ b_0 \end{bmatrix} = y$ indica-nos a solução no sentido dos mínimos quadrados da equação matricial, e daqui obtemos os valores de k_2 e de K_m.



Octave

Vamos definir `S`, `r1` e `r2` como os vectores correspondentes às colunas da tabela:

```
> S=[1;2;5;7.5;10;15;20;30];
> r1=[0.055;0.099;0.193;0.244;0.280;0.333;0.365;0.407];
> r2=[0.108;0.196;0.383;0.488;0.569;0.665;0.733;0.815];
> x=1./S;
> y1=0.005./r1;
> y2=0.01./r2;
```

Definimos também os quocientes respeitantes a y. A notação `a./b` indica que se faz a divisão elemento a elemento do vector `b`. Finalmente, definimos a matriz do sistema. Usou-se `ones(8)` para se obter a matriz 8×8 com 1's nas entradas, e depois seleccionou-se uma das colunas.

```
 > A=[x ones(8)(:,1)]
A =

   1.000000   1.000000
   0.500000   1.000000
   0.200000   1.000000
   0.133333   1.000000
   0.100000   1.000000
   0.066667   1.000000
   0.050000   1.000000
   0.033333   1.000000

> solucao1=inv(A'*A)*A'*y1
solucao1 =

   0.0813480
```

```
   0.0096492

> solucao2=inv(A'*A)*A'*y2
solucao2 =

   0.0831512
   0.0094384
```

Recorde que se poderia ter usado `solucao1=A\y1`. Daqui obtemos valores para k_2, K_m para cada uma das experiências. Vamos denotá-los, respectivamente, por `k21 Km1 k22 Km2`.

```
 > k21=1./solucao1(2,1)
k21 =  103.64
> k22=1./solucao2(2,1)
k22 =  105.95
> Km1=solucao1(1,1)*k21
Km1 =  8.4306
> Km2=solucao2(1,1)*k22
Km2 =  8.8098
> s=0:0.1:35;
> R1=(k21.*0.005*s)./(Km1+s);
> R2=(k22.*0.01*s)./(Km2+s);
> plot(s,R1,S,r1,'o',s,R2,S,r2,'o')
> grid on; legend('E0=0.005','' ,'E0=0.01','' );
```