23/05/2021
FGV EMAp

Teste 2
Álgebra Linear Numérica

Rener Oliveira

Questão 1

a) Fazendo $A = \begin{bmatrix} 1 & -3 \\ 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$ e $y = \begin{bmatrix} 1 \\ 1.2 \\ 2.1 \\ 1.5 \\ 2.1 \\ 2.5 \\ 2 \end{bmatrix}$ queremos encontrar $\alpha = \begin{bmatrix} a \\ b \end{bmatrix}$ tal que $A\alpha$ seja a melhor aproximação de y no espaço coluna de A. Tal vetor de coeficientes pode ser encontrado via resolução do sistema linear $A^T A\alpha = A^T y$.

Usando o Scilab, obtemos $a \approx 1.771$ e $b = 0.2$.

```
--> x = [-3;-2;-1;0;1;2;3];
--> A = [ones(7,1) x];
--> y = [1;1.2;2.1;1.5;2.1;2.5;2];
--> alpha = Gaussian_Elimination_4(A'*A,A'*y)
alpha  =
1.7714286
0.2

```

b) Com o modelo $y(x) = a + bx$ do item anterior, realizando os cálculos no Scilab obtemos:

$$y(4) = a + 4b \approx 2.571$$
$$y(10) = a + 10b \approx 3.771$$

c) No item (a) para a aproximação por função afim, usamos uma matriz com uma coluna de *uns* e outra coluna com os elementos x da tabela. Para o modelo quadrático deste exercício, adicionaremos uma coluna x^2 para aproximar uma parábola. Dessa forma temos uma matriz modificada:

$$A_2 = \begin{bmatrix} 1 & -3 & 9 \\ 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}$$

e mantemos y como mesmo vetor do item (a).

Agora o que estamos buscando é um vetor $\alpha_2 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, pois a multiplicação $A_2\alpha_2$ estará representando $a + bx + cx^2$ com x vetor coluna dos dados da tabela.

Agora os três coeficientes serão encontrados pela solução do sistema $A_2^T A_2 \alpha_2 = A_2^T y$. Pelo Scilab encontra-se:

$$\begin{aligned} a &\approx 1.943 \\ b &= 0.2 \\ c &\approx -0.043 \end{aligned}$$

```
--> A2 = [ones(7,1) x x^2];
--> alpha2 = Gaussian_Elimination_4(A2'*A2,A2'*y)
alpha2  =
1.9428571
0.2
-0.0428571

```

d) Usando os coeficientes encontrados na função $y(x) = a + bx + cx^2$, e o Scilab para realização dos cálculos,temos:

$$\begin{aligned} y(4) &= a + 4b + 16c \approx 2.052 \\ y(10) &= a + 10b + 100c \approx -0.115 \end{aligned}$$

```
--> a=alpha2(1);b=alpha2(2);c=alpha2(3);
--> a+4*b+16*c
ans  =
2.0571429

--> a+10*b+100*c
ans  =
-0.3428571

```

e) Para determinar qual modelo melhor se ajustou aos dados vamos calcular o erro quadrático de cada um deles. O erro do modelo linear do item (a) será $e_1 = ||y - A\alpha||$ e do modelo quadrático do item (c) será $e_2 = ||y - A_2\alpha_2||$. Utilizando o Scilab para os cálculos, encontramos:

$$e_1 \approx 0.821$$
$$e_2 \approx 0.721$$

Veja que $e_2 < e_1$, portanto, o modelo quadrático se ajustou melhor aos dados.

```
--> e1 = norm(y-A*alpha)
e1  =
0.8211490

--> e2 = norm(y-A2*alpha2)
e2  =
0.7211103

```

Questão 2

a) Tendo a fatoração QR da matriz A, se R é invertível a solução por mínimos quadrados do sistema $Ax = b$ é dada por $x = R^{-1}Q^Tb$. Para demonstrar tal fato, sabemos que tal solução satisfaz $A^TAx = A^Tb$. Substituindo A por QR e usando $Q^{-1} = Q^T$ mais o fato de R

invertível $\Rightarrow R^T$ invertível, temos:

$$\begin{aligned} A^TAx = A^Tb &\Rightarrow (QR)^T(QR)x = (QR)^Tb \\ &\Rightarrow R^TQ^TQRx = R^TQ^Tb \\ &\Rightarrow R^T(Q^{-1}Q)Rx = R^TQ^Tb \\ &\Rightarrow R^TIRx = R^TQ^Tb \\ &\Rightarrow Rx = Q^Tb \qquad (1) \\ &\Rightarrow x = R^{-1}Q^Tb \qquad (2) \end{aligned}$$

Computacionalmente ou mesmo para cálculos manuais, pode ser preferível encontrar x resolvendo o sistema linear (1) de inverter R e usar (2). Sendo $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ o vetor de respostas obtido através do número da matrícula. A solução procurada é obtida resolvendo o sistema linear em (1), que no caso será:

$$Rx = Q^Tb$$

$$\begin{bmatrix} 5 & 0.4 \\ 0 & 2.417 \end{bmatrix} x = \begin{bmatrix} -0.8 & 0.6 & 0 \\ 0.546 & 0.728 & -0.414 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

A solução a partir daqui é individual, mas pode ser resolvida via `Gaussian_Elimination_4(R,Q'*b)` após definir R, Q e b no console, ou substituindo b no sistema acima e prosseguir os cálculos manualmente.

b) Seja $x = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$ a primeira coluna de B, vamos transformá-la em $\pm \begin{bmatrix} ||x|| \\ 0 \\ 0 \end{bmatrix}$ via transformação ortogonal H_1 à ser descoberta pelo Método de Householder.

Utilizaremos o Refletor de Householder $H_1 = I - 2uu^T$ com $u = \dfrac{x - H_1x}{||x - H_1x||}$. Como $||x|| = \sqrt{4^2 + 3^2 + 0^2} = \sqrt{16+9} = \sqrt{25} = 5$, podemos escolher H_1x como $[5\ 0\ 0]^T$ ou $[-5\ 0\ 0]^T$; Apesar de ambos darem fatorações QR válidas no final, vamos escolher H_1x tal que $||x - H_1x||$ seja máximo. No caso, é fácil ver que tal H_1x será $[5\ 0\ 0]^T$ nesse critério, bastando olhar para a primeira coordenada de x e ver que subtraindo com 5 teremos -9 que é maior em módulo do que -4-(-5)=1.

Encontrando o vetor u:

$$
\begin{aligned}
u = \frac{x - H_1x}{||x - H_1x||} &= \frac{[-9\ 3\ 0]^T}{||[-9\ 3\ 0]^T||} \\
&= \frac{[-9\ 3\ 0]^T}{\sqrt{9^2 + 3^2}} = \frac{[-9\ 3\ 0]^T}{\sqrt{90}} = \\
&= \frac{1}{3\sqrt{10}}[-9\ 3\ 0]^T \\
&= \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}
\end{aligned}
$$

Encontrando a matriz H_1:

$$
\begin{aligned}
H_1 &= I - 2uu^T \\
&= I - 2 \cdot \frac{1}{\sqrt{10}} \cdot \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -3 & 1 & 0 \end{bmatrix} \\
&= I - \frac{1}{5} \begin{bmatrix} 9 & -3 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} 9 & -3 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} -4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}
\end{aligned}
$$

Para encontrar R_1 multiplicamos H_1 por B.

$$
\begin{aligned}
R_1 &= H_1 B \\
&= \frac{1}{5}\begin{bmatrix} -4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}\begin{bmatrix} -4 & 2 \\ 3 & 2 \\ 0 & -1 \end{bmatrix} \\
&= \frac{1}{5}\begin{bmatrix} 4^2+3^2 & -4\cdot 2+3\cdot 2 \\ -3\cdot 4+4\cdot 3 & 3\cdot 2+4\cdot 2 \\ 5\cdot 0 & -5 \end{bmatrix} \\
&= \frac{1}{5}\begin{bmatrix} 25 & -2 \\ 0 & 14 \\ 0 & -5 \end{bmatrix} = \begin{bmatrix} 5 & -2/5 \\ 0 & 14/5 \\ 0 & -1 \end{bmatrix}
\end{aligned}
$$

Consolidando a resposta, temos $H_1 = \frac{1}{5}\begin{bmatrix} -4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ e $R_1 = \begin{bmatrix} 5 & -2/5 \\ 0 & 14/5 \\ 0 & -1 \end{bmatrix}$.

Questão 3

Tome a fórmula $x = \dfrac{30-\sqrt{896}}{2}$ e multiplique espertamente por $\dfrac{30+\sqrt{896}}{30+\sqrt{896}}$, assim, teremos:

$$
\begin{aligned}
x &= \frac{(30-\sqrt{896})(30+\sqrt{896})}{2(30+\sqrt{896})} \\
&= \frac{30^2-896}{60+2\sqrt{896}} \\
&= \frac{900-896}{30+2\sqrt{896}} = \frac{4}{60+2\sqrt{896}} \\
&= \frac{2}{30+\sqrt{896}}
\end{aligned}
$$

Utilizando 29.9 como aproximação da raiz nessa nova fórmula, temos $x = \dfrac{2}{30+29.9} = \dfrac{2}{59.9} \approx 0.0333890$

O erro relativo à aproximação do enunciado que usa encontra $x \approx 0.0333705$ será:

$$\frac{0.0333890 - 0.0333705}{0.0333705} \approx 0.0005544$$

ou ainda, 0.05544% que é bem menor que 49%.
Uma nova proposta de algoritmo de aproximação é usar então a fórmula alternativa $\dfrac{2}{30+\sqrt{896}}$.
Porque tivemos tamanha redução? Seja $r = \sqrt{896}$ a raiz exata e Δr uma perturbação de forma que $r + \Delta r$ seja uma aproximação da raiz de 896.
Usando $x = \dfrac{30-\sqrt{896}}{2}$, o erro relativo será:

$$\begin{aligned}
\frac{|(30-r-\Delta r)/2-(30-r)/2|}{(30-r)/2} &= \frac{|30-r-\Delta r-30+r|}{30-r} \\
&= \frac{|-\Delta r|}{30-r} \\
&= \frac{|\Delta r|}{30-r}
\end{aligned}$$

Como 30 é numericamente muito próximo de r (na ordem de 0.066741) é necessário uma perturbação Δr bem menor, para que o erro relativo seja pequeno. O mal condicionamento numérico se dá por essa proximidade de 30 e r no denominador do erro, o que produz erros grandes se o numerador não for pequeno suficiente. A aproximação de 29.9 por exemplo, não e uma aproximação ruim, mas equivale à um Δr da ordem de 0.033 que corresponde aos 49% vistos no enunciado.
A modificação proposta $x = \dfrac{2}{30+\sqrt{896}}$, aceita perturbações Δr mais gros-

seiras do que a fórmula anterior, como vemos no cálculo a seguir:

$$
\begin{aligned}
&\frac{|2/(30+r+\Delta r)-2/(30+r)|}{2/(30+r)}\\
&=\frac{|1/(30+r+\Delta r)-1/(30+r)|}{1/(30+r)}\\
&=\left|\frac{30+r}{30+r+\Delta r}-1\right|\\
&=\left|\frac{30+r-30-r-\Delta r}{30+r+\Delta r}\right|\\
&=\frac{|\Delta r|}{30+r+\Delta r}
\end{aligned}
$$

Esta última expressão mostra que a nova fórmula é muito mais estável numericamente, pois o denominador é grande suficiente (na ordem de 60) para aceitar perturbações bem grosseiras sobre r e ainda assim manter o erro relativo pequeno. Apenas como exemplo se aproximarmos $\sqrt{896}$ por 28, com $\Delta r \approx 1.9332591$ teremos um erro relativo de 0.0312489 ou aproximadamente 3.1%. Achou muito? Usando tal Δr com o primeiro método chega-se a um erro da ordem de 28.966630 ou ainda 2896.63% !!!