UNIVERSIDADE FEDERAL FLUMINENSE
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ANÁLISE

Primeira Avaliação – 2019.1 (Matemática para Economia III e Álgebra Linear)
Professor: Bruno Santiago

Exercício 1 (1 pt)**.** *Sejam* $u = (0, 0, 2)$*,* $v = (1, 1, 0)$*,* $w = (-1, 1, 0)$ *e* $x = (4, 8, 16)$ *vetores em* $\mathbb{R}^3$*. Determine números reais* α*,* β *e* γ *tais que*

$$\alpha u + \beta v + \gamma w = x.$$

Solução. Se $\alpha u + \beta v + \gamma w = x$ então

$$\begin{aligned} \beta - \gamma &= 4 \\ \beta + \gamma &= 8 \\ 2\alpha &= 16. \end{aligned}$$

Somando membro a membro as duas primeiras equações deduzimos que $2\beta = 12$ e portanto $\beta = 6$. Com isso, da primeira equação concluímos que $\gamma = 2$. Finalmente, a terceira equação fornece $\alpha = 8$. $\square$

Exercício 2 (1.5 pt)**.** *Seja* E *um espaço vetorial com produto interno e seja* $\|x\| = \sqrt{\langle x, x\rangle}$ *a norma proveniente desse produto interno. Dados* $x, y \in E$ *prove que vale a fórmula a seguir, que* ***generaliza a lei dos cossenos ordinária da geometria euclideana****,*

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2.$$

Usando a lei dos cossenos, prove que $x, y \in E$ *são ortogonais (i.e.* $\langle x, y\rangle = 0$*) se, e somente se,*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Demonstração. Pela definição da norma vemos que $\|x + y\|^2 = \langle x + y, x + y\rangle$. Pela linearidade do produto interno, podemos deduzir

$$\begin{aligned} \langle x + y, x + y\rangle &= \langle x, x + y\rangle + \langle y, x + y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle. \end{aligned} \tag{1}$$

Pela simetria do produto interno $\langle x, y\rangle = \langle y, x\rangle$, segue então que

$$\langle x + y, x + y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle.$$

Como $\langle x + y, x + y\rangle = \|x + y\|^2$, $\langle x, x\rangle = \|x\|^2$ e $\langle y, y\rangle = \|y\|^2$, concluímos que

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2,$$

como desejado. Agora, uma vez que

$$\|x\|^2 + 2\langle x, y\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2 \iff 2\langle x, y\rangle = 0,$$

a lei dos cossenos implica que $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ se, e somente se $\langle x, y\rangle = 0$, ou seja, se e somente se x e y são ortogonais. $\square$

Exercício 3 (2 pt)**.** *Suponha que $x \in \mathbb{R}^n$ represente o vetor de contagem de palavras de um documento, dentro de um dicionário de n palavras, ou seja x_i é a quantidade de aparições da i-ésima palavra no dicionário. Vamos supor que todas as palavras do documento estão no dicionário. Considere $\beta = \langle \mathbb{I}, x \rangle$, onde $\mathbb{I} = (1, ..., 1) \in \mathbb{R}^n$. Seja $y \in \mathbb{R}^n$ o vetor que dá o histograma de contagem de palavras, ou seja, y_i é a fração do total de palavras no documento que são iguais a i-ésima palavra no dicionário. Expresse y em função de x e β.*

Solução. Como x_i é a quantidade de aparições da i-ésima palavra no dicionário,

$$\beta = \langle \mathbb{I}, x \rangle = \sum_{\ell=1}^{n} x_i$$

fornece o total de palavras usadas no documento. Logo, a fração do total de palavras no documento que são iguais a i-ésima palavra no dicionário é dada por

$$\frac{x_i}{\beta}.$$

Portanto, $y = \frac{1}{\beta}x$. □

Exercício 4 (2.5 pt)**.** *Nesse exercício, como vimos em aula, vamos considerar que vetores $\mathbb{R}^d$ representem séries temporais associadas a investimentos no mercado financeiro. Nesse caso, a j-ésima entrada dá o retorno (percentual) de um investimento no mês j após o início. Assim, a média das entradas dá o retorno do investimento e o desvio padrão dá o risco. Seja $x \in \mathbb{R}^d$ um investimento com retorno $\mu(x)$ e risco $\sigma(x)$. Vamos considerar $m \in \mathbb{R}^d$ dado por $m = \lambda\mathbb{I}$, com $\lambda \in (0, 1)$, que representa um investimento com retorno fixo ao longo do tempo. Vamos criar um portfólio simples que consiste em investir uma fração $\xi \in (0, 1)$ do capital em x e a outra $(1 - \xi)$ em m. Nesse caso, a série temporal de retornos é o vetor z dado por*

$$z = \xi x + (1 - \xi)m.$$

(a) *Deduza fórmulas para o retorno $\mu(z)$ e o risco $\sigma(z)$ do investimento z, em função de $\mu(x), \sigma(x), \xi$ e λ.*

Solução. Pela linearidade da média,

$$\mu(\xi x + (1 - \xi)m) = \xi\mu(x) + (1 - \xi)\mu(m).$$

Como $\mu(m) = \lambda$, deduzimos que

$$\mu(z) = \xi\mu(x) + (1 - \xi)\lambda.$$

Para deduzir uma fórmula para o risco de z, começamos observando que $\sigma(m) = 0$[1] Lembrando a fórmula do desvio padrão da soma,

$$\sigma(a + b) = \sqrt{\sigma(a)^2 + 2\rho(a, b)\sigma(a)\sigma(b) + \sigma(b)^2},$$

onde $\rho(a, b)$ é o coeficiente de correlação. Como $\sigma(m) = 0$ concluímos que

$$\sigma(z) = \sqrt{\sigma(\xi x)^2}$$

Como $\xi > 0$, sabemos que $\sigma(\xi x) = \xi\sigma(x)$. Portanto, $\sigma(z) = \xi\sigma(x)$. □

[1] Isso tem uma interpretação interessante: como o retorno é constante, o risco é zero.

(b) *Supondo que o investimento x tem risco positivo, prove que se o risco $\sigma(z)$ é 20% menor do que $\sigma(x)$ e o retorno $\mu(z)$ é 20% maior do que $\mu(x)$ então o retorno de x é metade do retorno de m, ou seja, $\mu(x) = 0.5\lambda$.*

Demonstração. Se $\sigma(z)$ é 20% menor do que $\sigma(x)$ então $\sigma(z) = 0.8\sigma(x)$. Como $\sigma(z) = \xi\sigma(x)$ deduzimos que

$$\xi\sigma(x) = 0.8\sigma(x) \quad \Longrightarrow \quad \xi = 0.8.$$

Como $\mu(z)$ é 20% maior do que $\mu(x)$, deduzimos que $\mu(z) = 1.2\mu(x)$. Pelo item (a), segue que

$$0.8\mu(x) + 0.2\lambda = 1.2\mu(x) \quad \Longrightarrow \quad 0.2\lambda = 0.4\mu(x),$$

e portanto $\mu(x) = 0.5\lambda$. □

Exercício 5 (3 pt)**.** *Suponha que o vetor $x \in \mathbb{R}^{100}$ represente a distribuição de idades em uma determinada população, de modo que x_i seja o número de pessoas com idade $i-1$. Suponha, por simplicidade que $x \neq 0$ e que ninguém tem idade superior a 99 nessa população. Determine o vetor $y \in \mathbb{R}^{100}$ tal que a idade média da população seja dada pela fórmula*

$$\textit{Idade média} = \frac{\langle y, x\rangle}{\langle x, \mathbb{I}\rangle},$$

onde $\langle .,. \rangle$ é o produto interno euclideano e $\mathbb{I} = (1, ..., 1) \in \mathbb{R}^{100}$ é o vetor de uns.

Solução. Existem x_1 pessoas recém nascidas (0 anos de idade), x_2 pessoas com 1 ano de idade e assim por diante. Logo, o número total de pessoas na população é

$$N = \sum_{i=1}^{100} x_i.$$

Agora, lembre o produto interno de x pelo vetor de uns fornece a soma das entradas de x. Ou seja,

$$\langle x, \mathbb{I}\rangle = \sum_{i=1}^{100} x_i.$$

Portanto $N = \langle x, \mathbb{I}\rangle$.

Por outro lado, para cada $i = 1, 2, ..., 99, 100$ existem x_i pessoas com idade $i-1$. Então, deduzimos que a idade total da população é

$$T = \sum_{i=1}^{100} (i-1)x_i.$$

Lembrando da definição do produto interno euclideano e considerando o vetor $y = (0, 1, 2, 3, ..., 98, 99) \in \mathbb{R}^{100}$ temos que

$$\langle y, x\rangle = \sum_{i=1}^{100} (i-1)x_i.$$

Portanto $T = \langle y, x\rangle$.

Finalmente, a idade média da população é o quociente entre a idade total da população o número de pessoas nela. Ou seja,

$$\text{Idade média} = \frac{T}{N}.$$

Como $T = \langle y, x\rangle$ e $N = \langle x, \mathbb{I}\rangle$, concluímos que $y = (0, 1, 2, 3, ..., 98, 99) \in \mathbb{R}^{100}$ é o vetor procurado. $\square$

Desafio 1 (1 pt)**.** *A nota de cada estudante em um turma de Álgebra Linear é dada por um vetor $x \in \mathbb{R}^{10}$, onde $x_1, ..., x_8$ são as notas das 8 listas de exercícios para entregar, cada uma valendo nota em uma escala 0-10, x_9 é nota da primeira avaliação (em escala 0-120) e x_{10} é a nota da avaliação final (numa escala de 0-160). A nota final, numa escala de 0-100, é baseada 25% nas listas de exercícios, 35% na primeira avaliação e 40% na segunda avaliação. Determine o vetor $y \in \mathbb{R}^{10}$ tal que a nota final do estudante seja dada pelo produto interno $\langle x, y\rangle$.*

Solução. Como as notas das 8 listas somam 80 pontos e devem corresponder a no máximo 25 pontos na nota final, para fazer a média ponderada que dará a nota final devemos somar o total de pontos na lista e dividir por $\frac{25}{80}$. Similarmente, como a P1 vai corresponder a no máximo 35 pontos na nota final, e vale no máximo 120 pontos, devemos multiplicar a nota na P1 por $\frac{35}{120}$. Finalmente, o mesmo raciocínio permite concluir que o coeficiente a multiplicar pela nota na P2 para formar a nota final é $\frac{40}{160}$. Assim se consideramos o vetor

$$y = (\frac{25}{80}, ..., \frac{25}{80}, \frac{35}{120}, \frac{40}{160}) \in \mathbb{R}^{10},$$

vemos que a nota final será

$$\langle x, y\rangle = \sum_{i=1}^{8} \frac{25}{80} x_i + \frac{35}{120} x_9 + \frac{40}{160} x_{10}.$$

Portanto, $y = (\frac{25}{80}, ..., \frac{25}{80}, \frac{35}{120}, \frac{40}{160}) \in \mathbb{R}^{10}$ é o vetor procurado. $\square$

Desafio 2 (1 pt)**.** *Suponha que $x, \hat{x} \in \mathbb{R}^d$ sejam, respectivamente, a temperatura medida a cada hora em d intervalos e a previsão para esta medição, fornecida pela meteorologia. Vamos supor que a previsão acerte em cheio 87% das medições, e que os erros cometidos são de pelo menos 1ºC, e no máximo 8ºC. Mostre que a distância* rms *entre x e $\hat{x}$ é limitada independentemente da quantidade de medições, enquanto que a distância euclideana entre x e $\hat{x}$ aumenta arbitrariamente quando o número d de medições cresce.*

Demonstração. Por hipótese temos que para 13% dos índices $i = 1, ..., d$ temos $1 \leq |x_i - \hat{x}_i| \leq 8$, correspondendo às medições nas quais um erro foi cometido. Além disso, existe um subconjunto de índices $\mathcal{K} \subset \{1, ..., d\}$ com $0.87d$ elementos, tal que

$$j \in \mathcal{K} \implies x_j = \hat{x}_j,$$

correspondendo aos acertos em cheio da previsão do tempo, os quais formam, como dito no enunciado, 87% das medições. Assim, ao considerarmos a distância euclideana precisamos apenas considerar os índices não pertencentes a $\mathcal{K}$. Por isso, ela pode ser estimada inferiormente por

$$\|x - \hat{x}\| = \sqrt{\sum_{j \notin \mathcal{K}} |x_i - \hat{x}_i|^2} \geq \sqrt{0.13d}.$$

Por outro lado, o desvio rms entre x e $\hat{x}$ pode ser estimado por

$$\mathrm{rms}(x - \hat{x}) = \sqrt{\frac{\sum_{i=1}^{d} |x_i - \hat{x}_i|^2}{d}} \leq \sqrt{\frac{8^2 d}{d}} = 8.$$

Como $\lim_{d\to\infty} \sqrt{0.13d} = +\infty$, vemos que o desvio rms entre x e $\hat{x}$ é limitada independentemente da quantidade de medições, enquanto que a distância euclideana entre x e $\hat{x}$ aumenta arbitrariamente quando o número d de medições cresce, como queria demonstrar. □

Desafio 3 (2 pt). *Suponha que $x \in \mathbb{R}^d$ represente uma série temporal ou um sinal discreto. A quantidade*

$$\mathcal{D}(x) \stackrel{\text{def}}{=} \sum_{\ell=1}^{d-1} (x_\ell - x_{\ell+1})^2 = (x_1 - x_2)^2 + ... + (x_{d-1} - x_d)^2,$$

que dá a soma dos quadrados das diferenças entre entradas adjacentes, é chamada a energia de Dirichlet *do sinal x. Ela fornece uma medida da oscilação absoluta do sinal.*

(a) Descreva $\mathcal{D}(x)$ em notação vetorial (usando coordenadas, produto interno, etc...)

Solução. Dado $x \in \mathbb{R}^d$ formamos dois vetores $x_{[1,d-1]} \in \mathbb{R}^{d-1}$ e $x_{[2,d]} \in \mathbb{R}^{d-1}$ pondo

$$x_{[1,d-1]} \stackrel{\text{def}}{=} (x_1, x_2, ..., x_{d-1}) \quad \text{e} \quad x_{[2,d]} \stackrel{\text{def}}{=} (x_2, x_3, ..., x_d).$$

Vemos então que a energia de Dirichlet é dada pelo quadrado da norma euclideana do vetor diferença $x_{[1,d-1]} - x_{[2,d]} \in \mathbb{R}^{d-1}$. Com efeito,

$$\|x_{[1,d-1]}, -x_{[2,d]}\|^2 = \sum_{\ell=1}^{d-1} (x_\ell - x_{\ell+1})^2.$$

Isso descreve a energia de Dirichelet em notação vetorial. □

(b) Qual o menor valor possível da energia de Dirichlet? Que vetores assumem esse menor valor de $\mathcal{D}(x)$?

Solução. Como a energia de Dirichelet é o comprimento euclideano ao quadrado de um vetor, seu menor valor possível é 0, e ele só é atingindo quando esse vetor é o vetor nulo. Portanto, o menor valor ocorre quando $x_{[1,d-1]}, -x_{[2,d]}$ é o vetor nulo de $\mathbb{R}^{d-1}$. Logo, temos $x_1 = x_2$, $x_2 = x_3$, assim por diante até $x_{d-1} = x_d$. Portanto, a energia de Dirichelet de um vetor é zero se, somente se esse vetor possui todas as entradas iguais. □

(c) *Encontre $x \in \mathbb{R}^d$ com todas as entradas menores ou iguais do que 1 em valor absoluto e tal que $\mathcal{D}(x)$ seja o mair possível. Expresse esse maior valor possível em função de d.*

Solução. Como estamos supondo que as entradas possuem valor absoluto ≤ 1, elas oscilam entre -1 e 1. Logo, a diferença entre cada duas entradas pode ser no máximo $|-1-1| = 2$. O maior valor possível vai ser atingindo quando todas as parcelas que aprecem na definição da energia de Dirichlet forem as maiores possíveis. Assim, o maior valor é atingindo pelos vetores

$$v = (1, -1, 1, -1, ..., \pm 1) \text{ e } v^* = (-1, 1, -1, ..., \pm 1).$$

A última entrada de v depende se d é a par ou ímpar. Quando for par, será -1, e quando d for ímpar será 1. O mesmo se aplica a v^*. Vemos que

$$\mathcal{D}(v) = 4(d-1).$$

A energia de Dirichlet é uma medida não-estatística do grau de oscilação de um vetor $x \in \mathbb{R}^d$. Não à toa, seu máximo é atingindo no vetor que mais oscila. O

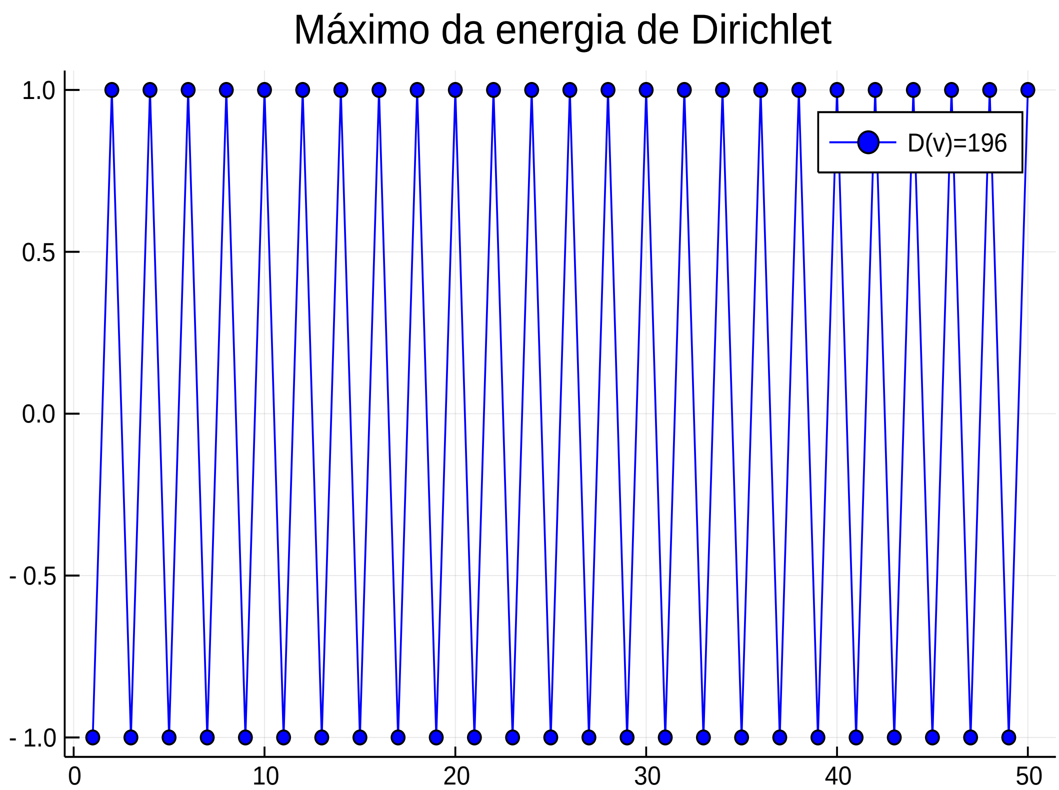

Figura 1: Figura produzida em Julia.

código usado para fazer a figura 1 foi

```
function v(n)
x=ones(n)
for j=1:n
x[j]=(-1)^ (j%2)x[j]
end
return x
end
plot(v(50),marker=:circle, color=:blue,title="Máximo da energia de Dirichlet",
label=["D(v)=196"])
```

□