Exame de Qualificação – Análise Aplicada

$2.^o$ semestre de 2025 – 25/08/2025

Nome:_____

RA:_____

1.^{a} Questão. (Espaços métricos)

(a) Mostre que se $C \subseteq (X, d)$ é compacto, então C é um conjunto limitado e fechado.

(b) Sejam (X, d) um espaço métrico qualquer, Y um conjunto denso em X e $M \subseteq X$. Prove ou dê um contra-exemplo: se $Y \subseteq M$, então M é denso em X.

(c) Sejam d_A e d_B duas métricas definidas em X. Se (X, d_A) é completo e d_A e d_B são (uniformemente) equivalentes, então, (X, d_B) é completo?

2.^{a} Questão. (Espaços de Banach)

(a) Mostre que o espaço vetorial c das sequências convergentes de números complexos, munido da norma do supremo, é um espaço de Banach.

(b) Sejam $(X, ||\cdot||)$ e $(Y, ||\cdot||_0)$ espaços normados. Definimos $||\cdot||_s : X \times Y \to \mathbb{R}_{\geq 0}$, dada por $||(x, y)||_s = \max\{||x||, ||y||_0\}$. Mostre que, se $(X, ||\cdot||)$ e $(Y, ||\cdot||_0)$ forem espaços de Banach, então $(X \times Y, ||\cdot||_s)$ será espaço de Banach.

3.^{a} Questão. (Espaços de Hilbert)

(a) Provar que $L^2[a,b]$ é um espaço de Hilbert sobre $\mathbb{C}$ com o produto interno

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}dt.$$

(b) Sejam X espaço de Hilbert sobre o corpo K, Y um subespaço completo de X e $x \in X$. Defina a projeção ortogonal de x em Y.

(c) Sejam X espaço de Hilbert, $M = \{e_i\}_{i=1}^n \subset X$ subconjunto ortogonal e $Y = span(M)$. Provar que para todo $x \in X$

$y = \sum_{i=1}^{n} \alpha_i e_i$ é projeção ortogonal de x em Y se e somente se $\alpha_i = \dfrac{\langle x, e_i\rangle}{\langle e_i, e_i\rangle}$.

4.^{a} Questão. (Teorema do ponto fixo de Banach)

(a) Em análise, uma condição suficiente para a convergência de uma iteração

$$x_{n+1} = g(x_n),$$

é que g seja continuamente diferenciável e que

$$|g'(x)| \leq \alpha < 1.$$

Verifique isso por meio do uso do teorema do ponto fixo de Banach.

(b) Seja f uma função real duas vezes diferenciável e com a segunda derivada contínua em um intervalo $[a, b]$ e seja $\hat{x}$ uma raiz simples de f em (a, b).

Mostre que o método de Newton definido por

$$x_{n+1} = g(x_n), \quad g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)},$$

é uma contração em alguma vizinhança de $\hat{x}$.

Nota:

EXAME - ANÁLISE NUMÉRICA - 2s/2025

Nome:	**RA:**

29 de Agosto de 2025

Escolha 4 Questões daquelas indicadas abaixo (10 (dez) pontos no total)

Questões (pontuação)	*Pontos*
Questão 1 (2.5)	
Questão 2 (2.5)	
Questão 3 (2.5)	
Questão 4 (2.5)	
Questão 5 (2.5)	
Questão 6 (2.5)	
Questão 7 (2.5)	
T o t a l (10.0)	

Boa Prova

Questão 1. **(2.5 Pontos)**

Considere o problema de valor inicial (PVI) formado pela equação diferencial parcial (EDP) de evolução, sendo $u = u(x,t)$,

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}u^2\right) = 0, \qquad x \in \Re, \quad t > 0, \tag{1}$$

com a condição inicial,

$$u(x,t_0) = \eta(x). \tag{2}$$

A EDP não linear (1) é conhecida como a *equação de Burgers invíscida*. Esta equação é fundamental no estudo da dinâmica de fluidos e turbulência, pois pode levar à formação de descontinuidades ou choques na solução.

A equação de *Burgers inviscida* é uma equação hiperbólica, dentro da classe de *leis de conservação hiperbólica*, que admite soluções fracas entrópicas representadas por ondas de choque ou de rarefação em função da imposição da condição inicial.

Considerando o PVI dado por (1)-(2), pede-se com as devidas justficativas:

(a) Dê dois exemplos de dados iniciais para os quais tem-se a formação de uma rarefação para tempos longos

(b) Dê dois exemplos de dados iniciais para os quais tem-se a formação de um choque para tempos longos.

(c) Dê um exemplo de dado inicial para o qual tem-se a presença de choque com rarefação.

(d) Indique ao menos dois métodos numéricos que sejam adequados para a simulação numérica do PVI dado por (1)-(2), indicando em cada caso a condição de estabilidade CFL.

Questão 2. **(2.5 Pontos)**

Seja o problema de Poisson em duas dimensões, sendo $u = u(x, y)$,

$$\nabla^2 u = f(x, y), \quad (x, y) \in \Omega = (0, 1) \times (0, 1), \tag{3}$$

sujeita a uma condição de contorno de Dirichlet,

$$u(x, y) = g(x, y), \quad (x, y) \in \partial\Omega. \tag{4}$$

Considerando uma malha computacional igualmente espaçada ($\Delta x = \Delta y = h$) e as aproximações de segunda ordem para o Laplaciano,

$$\nabla^2 u = u_{xx} + u_{yy},$$

pede–se discutir e demonstrar para o PVC (3)-(4)

(a) Erro de truncamento local.

(b) Erro Global.

(c) Convergência.

Dica: A matriz "A" do problema de álgebra linear associado a discretização da EDP (3) com as aproximações de segunda ordem é simétrica, negativa-definida, e na norma-2 tem-se que:

$$||A||_2 = \rho(A) = \max_{1 \leq p,q \leq m} \{|\lambda_{p,q}|\},$$

sendo, $\lambda_{p,q}$ os autovalores, dados por:

$$\lambda_{p,q} = \frac{2}{h^2}[(cos(p\pi h) - 1) + (cos(q\pi h) - 1)], \quad p, q = 1, 2, \cdots, m.$$

Questão 3. **(2.5 Pontos)**

Considere o problema de valor inicial (PVI) formado pela equação diferencial parcial (EDP) de evolução, sendo $u = u(x,t)$,

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial^2 u}{\partial x^2}, \qquad x \in \Re, \quad t > 0, \tag{5}$$

com a condição inicial,

$$u(x, t_0) = \eta(x), \qquad x \in \Re, \quad t_0 = 0, \tag{6}$$

sendo κ constante ($\kappa > 0$). Se $\kappa \equiv 0$ temos uma problema parabólico degenerado. Se $\kappa < 0$ temos uma problema mal-posto no sentido "clássico" de Hadamard.

Aplicando ao PVI (5)-(6) aproximações combinadas de diferenças finitas no tempo e no espaço (i.e., $U_j^n \approx u(x_j, t_n)$) pode-se obter o seguinte esquema geral, sendo $r \equiv \kappa \dfrac{k}{h^2}$ e supondo κ constante (por simplicidade):

$$-\alpha\, r\, U_{j-1}^{n+1} + (1+2\,\alpha\, r) U_j^{n+1} - \alpha\, r\, U_{j+1}^{n+1} = (1-\alpha)\, r\, U_{j-1}^n + [1-2(1-\alpha)\, r] U_j^n + (1-\alpha)\, r\, U_{j+1}^n, \tag{7}$$

$$t_n = n\,k, \quad n = 0, 1, 2, \ldots \qquad x_j = j\,h, \quad j = \ldots -2, -1, 0, 1, 2, \ldots \tag{8}$$

onde $\alpha \in [0, 1]$, $k = \Delta t$ denota o passo de tempo e $h = \Delta x$ denota o espaçamento da malha.

Pede–se, com as devidas justificativas:

(a) Prove consistência e estabilidade do método (7)-(8) para a aproximação do PVI (5)-(6).

(b) Enuncie o teorema de Lax-Richtmyer que estabelece a conexão entre consistência e estabilidade para convergência de métodos numéricos lineares de um passo para problemas de EDP de evolução lineares. Utilize esse resultado para demonstrar/provar a convergência do método (7)-(8) para a solução do PVI (5)-(6), considerando obrigatoriamente neste contexto a **Análise de Fourier** e a **Identidade de Parseval**.

(c) Considere o PVI (5)-(6), mas agora com o coeficiente de condução de calor de um meio material não constante, i.e., $\kappa = \kappa(x)$. Neste contexto, explique como um coeficiente variável de condução de calor $\kappa(x)$ pode afetar a *análise numérica* e, assim, a relação de estabilidade do método (7)-(8) para o PVI (5)-(6), levando em conta também estratégias de eficiência na implementação computacional, como por exemplo, recurso de memória, complexidade do algoritmo e tempo computacional.

Questão 4. **(2.5 Pontos)**

Considere o problema de valor de inicial (PVI) formado pela equação diferencial de primeira ordem,

$$u'(t) = f(u(t), t), \quad t > 0, \tag{9}$$

com condição inicial,

$$u(t_0) = u_0. \tag{10}$$

Um método linear de $r-$passos múltiplos para o PVI (9)-(10) tem a forma geral dada por,

$$\sum_{j=0}^{r} \alpha_j \, U^{n+j} = k \sum_{j=0}^{r} \beta_j \, f(U^{n+j}, t_{n+j}). \tag{11}$$

Pede-se, com as devidas justificativas,

(a) Escreva a forma geral dos polinômios característicos $\rho(\xi)$ e $\sigma(\xi)$, associados ao método linear de $r-$passos múltiplos (11) para o PVI (9)-(10).

(b) Escreva a definição de estabilidade-zero para métodos lineares de $r-$passos múltiplos.

(c) Deduza os métodos de Euler avançado, Euler recuado, e Leapfrog (ponto médio).

(d) Para um método linear de $r-$passos múltiplos da forma (11), a região de estabilidade absoluta é definida pela aplicação deste método na equação diferencial de primeira ordem (9), com $f(u(t), t) = \lambda u(t)$, que resulta em,

$$\sum_{j=0}^{r} \alpha_j \, U^{n+j} = k \sum_{j=0}^{r} \beta_j \, \lambda U^{n+j}, \tag{12}$$

e que pode ser reescrito como,

$$\sum_{j=0}^{r} (\alpha_j - z\beta_j) U^{n+j}, \quad z = k\lambda. \tag{13}$$

Escreva os polinômios-estabilidade da forma,

$$\pi(\xi, z) = \rho(\xi) - z\sigma(\xi),$$

e esboce a região de estabilidade absoluta (no plano complexo) para os métodos de Euler avançado, Euler recuado, e Leapfrog (ponto médio).

Questão 5. **(2.5 Pontos)**

Se aplicarmos um método de passo único ao PVI,

$$u'(t) = \lambda u(t), \quad t > 0, \tag{14}$$

com condição inicial,

$$u(t_0) = u_0, \tag{15}$$

tipicamente obtém-se uma equação da forma,

$$U^{n+1} = R(z)U^n,$$

onde $R(z)$ é alguma função de estabilidade com $z = k\lambda$; tipicamente um polinômio para um método explícito ou uma função racional para um método implícito.

Pede-se, com as devidas justificativas:

(a) Mostre que os dois métodos de "série de Taylor de $4^{\underline{a}}$ ordem" e "Runge-Kutta de $4^{\underline{a}}$ ordem", quando aplicados ao PVI teste (14)-(15), produzem a mesma função $R(z)$.

Dica: O método de Runge-Kutta de $4^{\underline{a}}$ ordem é dado por:

$$\begin{aligned}
Y_1 &= U^n, \\
Y_2 &= U^n + \frac{1}{2}kf(Y_1, t_n), \\
Y_3 &= U^n + \frac{1}{2}kf\left(Y_2, t_n + \frac{k}{2}\right), \\
Y_4 &= U^n + kf\left(Y_3, t_n + \frac{k}{2}\right), \\
U^{n+1} &= U^n + \frac{k}{6}\left[f(Y_1, t_n) + 2f\left(Y_2, t_n + \frac{k}{2}\right)\right. \\
&\qquad \left. + 2f\left(Y_3, t_n + \frac{k}{2}\right) + f(Y_4, t_n + k)\right]
\end{aligned}$$

(b) Se um método é consistente, então $R(z)$ será uma aproximação de e^z para z próximo de zero. Além disso, se esse método é de $p-$ésima ordem, então,

$$R(z) - e^z = O(z^{p+1}), \text{ quando } z \to 0.$$

Mostre que a função $R(z)$ para os métodos do item (a) aproxima e^z a menos de um erro $O(z^5)$, quando $z \to 0$.

Questão 6. **(2.5 Pontos)**

Considere o problema de valor inicial (PVI) formado pela equação diferencial parcial (EDP) de evolução linear, sendo $u = u(x,t)$,

$$\frac{\partial u}{\partial t} + a\,\frac{\partial u}{\partial x} = 0, \qquad x \in \Re, \quad t > 0, \tag{16}$$

com a condição inicial,

$$u(x,t_0) = \eta(x), \qquad x \in \Re, \quad t_0 = 0, \tag{17}$$

sendo a constante ($a \neq 0$). A equação (16) é o exemplo mais simples de uma equação hiperbólica, e a solução exata é dada por,

$$u(x,t) = \eta(x - at). \tag{18}$$

Isto significa que o dado inicial (17) é transladado de forma inalterada e com velocidade a para a direita (se $a > 0$) e para a esquerda (se $a < 0$). Para $a = 0$ a solução é $u(x,t) = \eta(x)$, $\forall\ t > 0$. A EDP (16) é também conhecida como a equação de advecção linear, que modela problemas que exibem propagação de ondas.

Para o PVI (16)-(17), temos ainda que o valor de $u(x,t)$ é constante ao longo de cada curva característica $x - at = x_0$, sendo x_0 o domínio de dependência do ponto (x,t) para a EDP (16). No caso linear aqui discutido tal curva é uma reta com inclinação constante.

Considerando uma discretização no espaço-tempo como segue,

$$t_n = n\,k, \quad n = 0,1,2\cdots \qquad x_j = j\,h, \quad j = \cdots -2,-1,0,1,2,\cdots \tag{19}$$

onde $k = \Delta t$ denota o passo de tempo e $h = \Delta x$ denota o espaçamento da malha, temos então neste caso que sobre um único passo de tempo resulta:

$$u(x_j, t_{n+1}) = u(x_j - ak, t_n). \tag{20}$$

Traçando a característica como indicado em (20) de volta no tempo por um único passo de tempo k, a partir do ponto da malha x_j.Se $0 < ak/h < 1$, então o ponto $x_j - ak$ pertence entre os pontos x_{j-1} e x_j, ver Figura 1(a).

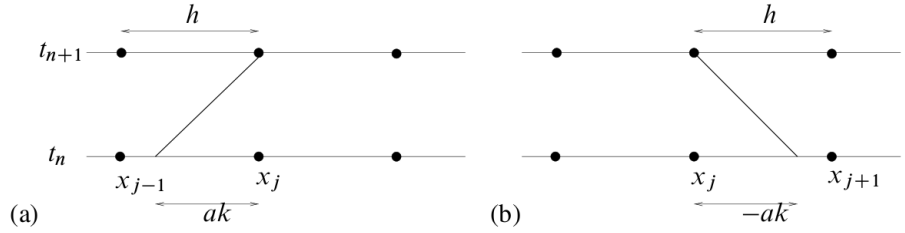

Figura 1: Curva característica "discreta".

Questão 6. (continuação)

Se $-1 < ak/h < 0$, temos o ponto $x_j - ak$ pertence entre os pontos x_j e x_{j+1}, ver Figura $1(b)$. Temos assim que a Figura 1 ilustra como podemos utilizar a solução analítica (18) como base para construção de um método numérico pelo cálculo de uma aproximação para (20) baseada em uma interpolação dos pontos da malha e próximos do ponto U_j^n. Por exemplo, realizando uma interpolação linear simples entre os pontos $(x_{j-1}, U_{j-1}^n$ e (x_j, U_j^n), sendo $a > 0$, resulta em:

$$p(x) = U_j^n + (x - x_j)\left(\frac{U_j^n - U_{j-1}^n}{h}\right). \tag{21}$$

Note que avaliando o ponto $x_j - ak$ na fórmula de interpolação (21) e definindo $p(x_j - ak)$ como U_j^{n+1}, obtemos o método upwind $O(k) + O(h)$:

$$\frac{U_j^{n+1} - U_j^n}{k} + a\,\frac{U_j^n - U_{j-1}^n}{h} = 0, \qquad \text{se } a > 0. \tag{22}$$

Repetindo os passos acima para $a < 0$, $p(x) = U_j^n + (x - x_j)\left(\dfrac{U_{j+1}^n - U_j^n}{h}\right)$, os pontos de interpolação (x_j, U_j^n), (x_{j+1}, U_{j+1}^n) resulta no método upwind $O(k) + O(h)$:

$$\frac{U_j^{n+1} - U_j^n}{k} + a\,\frac{U_{j+1}^n - U_j^n}{h} = 0, \qquad \text{se } a < 0, \tag{23}$$

Para obter uma melhor aproximação para o PVI (16)-(17), pode-se utilizar então a solução analítica (18) como base para construção de um método numérico pelo cálculo de uma melhor aproximação para (20) baseada em uma interpolação com mais pontos da malha e próximos do ponto U_j^n.

Neste contexto, pede-se, com as devidas justificativas,

(a) Supondo que $|ak/h| \leq 1$, defina um polinômio quadrático $p(x)$ pela interpolação dos valores da malha U_{j-1}^n, U_j^n e U_{j+1}^n, e então defina U_j^{n+1} por $p(x_j - ak)$ para obter o método de Lax-Wendroff $O(k^2) + O(h^2)$ para a equação de advecção linear (16).

(b) Suponha agora que $0 \leq ak/h \leq 2$ (i.e., a¿0) Defina então um outro polinômio quadrático $p(x)$ pela interpolação dos valores da malha U_{j-2}^n, U_{j-1}^n e U_j^n, defina U_j^{n+1} por $p(x_j - ak)$ para obter o método de Beam-Warming $O(k^2) + O(h^2) + O(h\,k)$ para o caso $a > 0$, para a equação de advecção linear (16).

(c) Para cada um dos esquemas explícitos de Lax-Wendroff (item a) e Beam-Warming (item b) é possível escolher parâmetros k (passo de tempo) e h (espaçamento da malha) de tal forma que a solução obtida concorda exatamente com a solução analítica (18). Isso não é um "chute", mas uma escolha adequada com base no resultado (20), como rigorosamente previsto pela teoria (18). Explique qual é essa escolha adequeda para k e h para cada um dos esquemas explícitos de Lax-Wendroff e de Beam-Warming. Faça também uma interpretação geométrica (via curva característica no plano x-t) do mesmo tipo como na Figura 1.

Questão 7. **(2.5 Pontos)**

Considere o problema de valor inicial (PVI) formado pela equação diferencial parcial (EDP) de evolução linear, sendo $u = u(x,t)$,

$$\frac{\partial u}{\partial t} + a\,\frac{\partial u}{\partial x} = 0, \qquad x \in \Re, \quad t > 0, \tag{24}$$

com a condição inicial,

$$u(x,t_0) = \eta(x) = e^{-20(x-2)^2} + e^{-(x-5)^2}, \qquad x \in \Re, \quad t_0 = 0, \tag{25}$$

sendo a constante ($a \neq 0$). A EDP (24) é conhecida como a equação de advecção linear, que é o exemplo mais simples de uma equação hiperbólica, cuja solução exata é dada por,

$$u(x,t) = \eta(x - at). \tag{26}$$

A solução exata (26) revela que o dado inicial (25) é transladado de forma inalterada e com velocidade a para a direita (se $a > 0$) e para a esquerda (se $a < 0$).

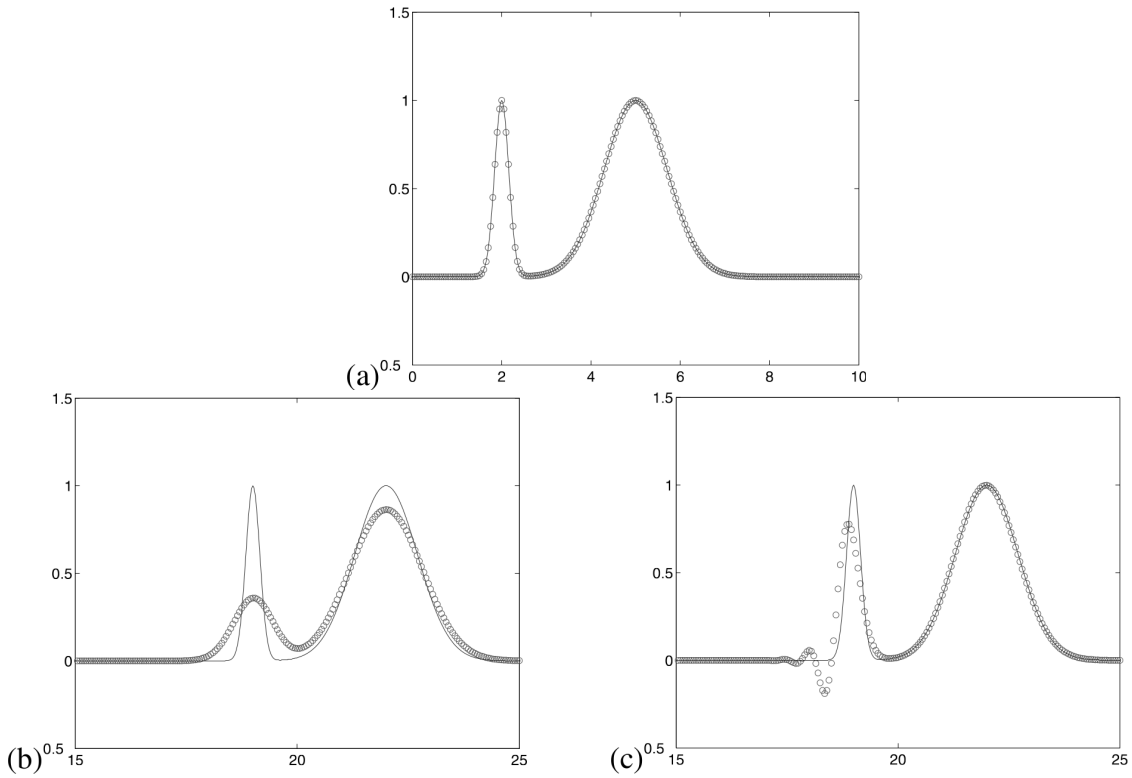

Figura 2: (a) Condição inicial $u(x,0) = e^{-20(x-2)^2} + e^{-(x-5)^2}$. Para um tempo $t^* > 0$, temos as soluções numéricas: (b) Upwind e (c) Lax-Wendroff. A solução exata (26) é também mostrada em linha sólida nos mesmos gráficos (b) e (c).

A Figura 2(a) mostra o dado inicial (25). Desta forma, um bom método numérico para o PVI (24)-(25) deve ter a propriedade de preservar a "forma" da condição inicial (25) com

Questão 7. (continuação)

"velocidade" de propagação a para qualquer tempo $t > 0$. Considerando uma discretização no espaço-tempo como segue,

$$t_n = n\,k, \quad n = 0, 1, 2 \cdots \qquad x_j = j\,h, \quad j = \cdots -2, -1, 0, 1, 2, \cdots \tag{27}$$

onde $k = \Delta t$ denota o passo de tempo e $h = \Delta x$ denota o espaçamento da malha, e aplicando ao PVI (24)-(25) aproximações de diferenças finitas no tempo e no espaço, pode-se obter (supondo $a > 0$) os seguintes métodos explícitos Upwind e Lax-Wendroff, respectivamente:

$$U_j^{n+1} = U_j^n - \frac{ak}{h}\left(U_j^n - U_{j-1}^n\right), \tag{28}$$

e

$$U_j^{n+1} = U_j^n - \frac{ak}{2h}\left(U_{j+1}^n - U_{j-1}^n\right) + \frac{a^2k^2}{2h^2}\left(U_{j-1}^n - 2U_j^n + U_{j+1}^n\right). \tag{29}$$

Na Figura 2(b) é mostrada a solução numérica do PVI (24)-(25) para um tempo $t^* > 0$ com o método Upwind (28), enquanto que na Figura 2(c) é mostrada a solução numérica para o mesmo problema com o método Lax-Wendroff (29). Para os dois métodos utilizou-se uma velocidade de propagação $a = 1$ com os parâmetros de malha $h = 0.05$ e $k = 0.8h$, o que resulta em um número de Courant $ak/h = 0.8$.

Os resultados mostrados na Figura 2 revelam claramente que o método Upwind (28) exibe um caráter de uma equação de advecção-difusão (Figura 2(b)), que é diferente do método Lax-Wendroff (29), que exibe um caráter de uma equação de advecção-dispersão (Figura 2(c)).

Neste contexto, pede-se, com as devidas justificativas,

(a) Determine a equação diferencial parcial modificada associada ao método Upwind (28), mantendo os termos $O(k)$, para k/h fixo.

(b) Determine a equação diferencial parcial modificada associada ao método Lax-Wendroff (29), mantendo os termos $O(k^2)$, para k/h fixo.

(c) Realize uma análise via transformada de Fourier das equações modificadas obtidas no itens (a) e (b), imediatamente acima, para obter a relação de dispersão associada a cada um dos métodos Upwind (28) e Lax-Wendroff (29). Utilize então esses resultados para explicar a natureza dos resultados mostrados na Figura 2.

(d) Para cada uma das equações diferenciais parciais modificadas obtidas nos itens (a) e (b) é possível escolher um número de Courant adequado de tal forma que essas equações modificadas concordem exatamente com a EDP original (24). Assim, exiba e explique qual é essa escolha adequada para o número de Courant. Note que isso também não é um "chute", mas uma escolha conveniente dos parâmetros k e h que pode ser totalmente justificada com base no resultado obtido no item (c).

Esquema de Beam-Warming para $(a > 0)$

$$U_j^{n+1} = U_j^n - \frac{ak}{2h}\left(3U_j^n - 4U_{j-1}^n + U_{j-2}^n\right) + \frac{a^2k^2}{2h^2}\left(U_j^n - 2U_{j-1}^n + U_{j-2}^n\right).$$

Interpolação polinomial de grau n - método de Lagrange

$$p(x) = \sum_{k=0}^{n} f(x_k)\left[\prod_{i=0;i\neq k}^{n} \frac{(x - x_i)}{(x_k - x_i)}\right]$$

UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Matemática, Estatística e Computação Científica
MT571 – Aprendizado de Máquinas: Aspectos Teóricos e Práticos

ALUNO	RA

MT571 – Aprendizado de Máquinas: Aspectos Teóricos e Práticos

Exame de Qualificação – 29/08/2025

INSTRUÇÕES

SERÃO CONSIDERADAS SOMENTE AS QUESTÕES ESCRITAS DE FORMA CLARA E DEVIDAMENTE JUSTIFICADAS

INFORMAÇÕES ÚTEIS

$$\text{Precisão} = \frac{V_p}{V_p + F_p} \qquad \text{e} \qquad \text{Revocação} = \frac{V_p}{V_p + F_n}. \tag{1}$$

Regressão linear, ridge, lasso e elastic net: Considere $h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \theta_0 + \sum_{j=1}^{n} \theta_j x_j$, a regressão linear, ridge, lasso e elastic net são obtidas minimizando as seguintes função perda:

$$\mathcal{L}_{\texttt{MSE}}(\boldsymbol{\theta}) = \frac{1}{m} \sum_{i=1}^{m} (h_{\boldsymbol{\theta}}(\boldsymbol{x}_i) - y_i)^2. \tag{2}$$

$$\mathcal{L}_{\texttt{Ridge}}(\boldsymbol{\theta}) = \frac{1}{m} \sum_{i=1}^{m} (h_{\boldsymbol{\theta}}(\boldsymbol{x}_i) - y_i)^2 + \alpha \sum_{j=1}^{n} \theta_j^2. \tag{3}$$

$$\mathcal{L}_{\texttt{Lasso}}(\boldsymbol{\theta}) = \frac{1}{m} \sum_{i=1}^{m} (h_{\boldsymbol{\theta}}(\boldsymbol{x}_i) - y_i)^2 + \alpha \sum_{j=1}^{n} |\theta_j|. \tag{4}$$

Máquina de Suporte Vetorial Linear:

$$\begin{cases} \text{argmin}_{\mathbf{w},b,\boldsymbol{\xi}} & \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{C}{m}\sum_{i=1}^{m}\xi_i \\ \text{sujeito à} & y_i(\mathbf{w}^T\boldsymbol{x}_i + b) \geq 1, \quad \forall i = 1, \ldots, m, \\ & \xi \geq 0, \quad \forall i = 1, \ldots, m, \end{cases} \tag{5}$$

$$\begin{cases} \text{argmax}_{\boldsymbol{\alpha}} & \displaystyle\sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} y_i y_j \alpha_i \alpha_j \boldsymbol{x}_i^T \boldsymbol{x}_j \\ \text{sujeito à} & \displaystyle\sum_{i=1}^{m} y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad \forall i = 1, \ldots, m, \end{cases} \tag{6}$$

A função decisão $h : \mathbb{R}^n \to \mathbb{R}$ da SVM é

$$h(\boldsymbol{x}) = \mathbf{w}^T\boldsymbol{x} + b = \sum_{i=1}^{m_S} \alpha_{(i)} y_{(i)} \langle \boldsymbol{x}_i, \boldsymbol{x} \rangle + b. \tag{7}$$

Questão 1. Considere uma rede neural que possui a função softmax na camada de saída. Lembre-se que a função softmax é dada pela equação

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}, \quad \forall i = 1, \ldots, n, \tag{8}$$

em que $z_1, \ldots, z_n$ e $p_1, \ldots, p_n$ denotam as entradas e saídas, respectivamente.

(a) Mostre que as saídas $p_1, \ldots, p_n$ da função softmax são todas positivas e somam 1. Como essa observação é utilizada na interpretação da saída da rede neural em problemas de classificação?

(b) Mostre que $\partial p_i / \partial z_k$ é positivo se $i = k$ e negativo se $i \neq k$. Como esse resultado é interpretado no aprendizado da rede neural?

(c) Mostre que a entrada z_i pode ser expressa em termos do logaritmo da saída p_i adicionando uma constante C que não depende do índice i.

(d) Qual função perda (loss function) é recomendada para treinar a rede neural num problema de classificação? Justifique sua resposta.

Questão 2. Considere um comitê com 10 classificadores binários $\mathcal{E} = \{\psi_1, \ldots, \psi_{10}\}$, em que $\psi_i : \mathcal{X} \to \{0, 1\}$ é definido pela equação $\psi_i(\boldsymbol{x}) = \chi[f_i(\boldsymbol{x}) \geq 0.5]$ usando uma função-de-decisão $f_i : \mathcal{X} \to [0, 1]$ que fornece a probabilidade de uma amostra pertencer à classe 1 (como, por exemplo, a regressão logística). Aqui, χ denota a função indicadora, ou seja, $\chi[P] = 1$ se a proposição P é verdadeira e $\chi[P] = 0$ caso contrário. Suponha que os 10 classificadores forneceram as probabilidades

$$\begin{array}{lllll} f_1(\boldsymbol{x}) = 0.1, & f_2(\boldsymbol{x}) = 0.15, & f_3(\boldsymbol{x}) = 0.2, & f_4(\boldsymbol{x}) = 0.2, & f_5(\boldsymbol{x}) = 0.55, \\ f_6(\boldsymbol{x}) = 0.6, & f_7(\boldsymbol{x}) = 0.6, & f_8(\boldsymbol{x}) = 0.65, & f_9(\boldsymbol{x}) = 0.7, & f_{10}(\boldsymbol{x}) = 0.75, \end{array} \tag{9}$$

para uma amostra $\boldsymbol{x} \in \mathcal{X}$.

(a) Qual será a classe estimada de $\boldsymbol{x} \in \mathcal{X}$ por um ensemble $\psi_E : \mathcal{X} \to \{0, 1\}$ obtido usando o sistema de votos pela maioria?

(b) Diferente do voto pela maioria, a classe de $\boldsymbol{x} \in \mathcal{X}$ pode ser estimada por um ensemble $\psi_F : \mathcal{X} \to \{0, 1\}$ definido como segue usando a chamada votação flexível:

$$\psi_F(\boldsymbol{x}) = \chi\left[\frac{1}{10}\sum_{i=1}^{10} f_i(\boldsymbol{x}) \geq 0.5\right]. \tag{10}$$

Usando a votação flexível, qual será a classe de $\boldsymbol{x} \in \mathcal{X}$?

(c) Compare os resultados fornecidos pelo sistema de votos pela maioria e a votação flexível.

Questão 3. Considere uma extensão do algoritmo PCA padrão que incorpora regularização. O algoritmo modifica a matriz de covariância da seguinte forma:

$$C_{\text{reg}} = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)(x_i - \mu)^T + \lambda I,$$

em que λ é o parâmetro de regularização e I é a identidade.

(a) Dado o dataset bidimensional:

x_1	2.0	1.0
x_2	1.0	2.0
x_3	0.0	0.0

Com $\lambda = 0.5$:

- Calcule a matriz de covariância C e a matriz regularizada C_{reg}
- Encontre os autovalores e autovetores de ambas
- Compare os erros de reconstrução com $k = 1$ componente

(b) Mostre que os autovalores da matriz regularizada são:

$$\lambda_{i,\text{reg}} = \lambda_i + \lambda$$

onde λ_i são os autovalores originais.

(c) Derive uma expressão matemática de como o erro de reconstrução é afetado pela regularização.

Questão 4. Considere uma versão do algoritmo K-means que usa um kernel Gaussiano para capturar relações não-lineares. A distância entre um ponto x_i e o centroide do cluster k é calculada no espaço de características como:

$$d_k(x_i) = K(x_i, x_i) - \frac{2}{|C_k|} \sum_{x_j \in C_k} K(x_i, x_j) + \frac{1}{|C_k|^2} \sum_{x_j, x_l \in C_k} K(x_j, x_l)$$

onde $K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$ e C_k é o conjunto de pontos no cluster k.

(a) Dado o dataset:

x_1	0.0	0.0
x_2	1.0	1.0
x_3	3.0	0.0
x_4	4.0	1.0

Com $\sigma = 1.0$ e $K = 2$ clusters:

- Calcule a matriz kernel $K_{4\times4}$
- Inicie com clusters $C_1 = \{x_1, x_2\}$ e $C_2 = \{x_3, x_4\}$
- Calcule a distância kernelizada de cada ponto para cada cluster
- Determine se há reatribuição de pontos

(b) Explique por que $K(x_i, x_i) = 1$ no kernel Gaussiano e como isso simplifica a fórmula da distância.

(c) Mostre matematicamente que quando $\sigma \to \infty$, a distância kernelizada converge para a distância Euclidiana padrão do K-means.

(d) Derive como valores pequenos de σ (kernel "estreito") versus valores grandes de σ (kernel "largo") afetam a formação dos clusters.

FOLHA ADICIONAL

FOLHA ADICIONAL

Boa Prova!

UNIVERSIDADE ESTADUAL DE CAMPINAS
Instituto de Matemática, Estatística e Computação Científica
MT580 – Conjuntos e Lógica Fuzzy: Teoria e Aplicações

ALUNO	RA

MT580 – Conjuntos e Lógica Fuzzy: Teoria e Aplicações

Exame de Qualificação – 29/08/2025

INSTRUÇÕES

SERÃO CONSIDERADAS SOMENTE AS QUESTÕES ESCRITAS DE FORMA CLARA E DEVIDAMENTE JUSTIFICADAS

INFORMAÇÕES ÚTEIS

Definição 1 (Função de pertinência triangular)**.** *A função de pertinência triangular* $T(a,m,b):\mathbb{R}\to[0,1]$*, caracterizada pelos parâmetros* $a,m,b\in\mathbb{R}$*, é dada por*

$$T(a,m,b)(x)=\begin{cases}(x-a)/(m-a), & a<x<m,\\ (b-x)/(b-m), & m<x<b,\\ 0, & \textit{caso contrário},\end{cases}$$

Definição 2 (Relação de Equivalência)**.** *Uma relação clássica* $E\subseteq U\times U$ *é uma relação de equivalência em* U *se:*

(i) $(u,u)\in E$ *para todo* $u\in U$*.*

(ii) $(u,v)\in E$ *implica* $(v,u)\in E$*.*

(iii) $(u,v)\in E$ *e* $(v,w)\in E$ *implica* $(u,w)\in E$*.*

Definição 3 (Relação de Equivalência Fuzzy)**.** *Uma relação fuzzy* $\mathcal{R}\in\mathcal{F}(U\times U)$ *é uma relação de equivalência fuzzy se:*

(a) $\mathcal{R}(u,u)=1$ *para todo* $u\in U$*.*

(b) $\mathcal{R}(u,v)=\mathcal{R}(v,u)$*, para todo* $u,v\in U$*.*

(c) $\mathcal{R}(u,w)\geq\mathcal{R}(u,v)\wedge\mathcal{R}(v,w)$*,* $\forall u,v,w\in U$*.*

Questão 1.

Neste ano, celebramos os 60 anos da teoria dos conjuntos *fuzzy*. Esta teoria surgiu com o artigo intitulado "Fuzzy Sets" no qual L. A. Zadeh observou o seguinte:

```
Essentially, fuzzy sets in X
constitute a distributive lattice with
a 0 and 1.
```

(a) Quando podemos dizer que $\mathbb{L} \neq \emptyset$ com 2 operações binárias $\curlywedge, \curlyvee : \mathbb{L}^2 \to \mathbb{L}$ representa um reticulado? No caso afirmativo, como definimos uma ordem parcial em $\mathbb{L}$?

(b) Quais são os operadores de ínfimo e supremo de $A, B \in \mathcal{F}(X)$ no reticulado ao qual Zadeh se referiu?

(c) Considere a operação binária $\vee$ (máximo) em $\mathbb{R}$. Defina $\sqcup : \mathcal{F}(\mathbb{R}) \times \mathcal{F}(\mathbb{R}) \to \mathcal{F}(\mathbb{R})$ como sendo a extensão de Zadeh de $\vee : \mathbb{R}^2 \to \mathbb{R}$.

(d) Sejam $\chi_{[1,2]\cup[5,6]}$ e $\chi_{[3,4]}$ as funções de pertinência de $A \in \mathcal{F}(\mathbb{R})$ e $B \in \mathcal{F}(\mathbb{R})$, respectivamente. Determine $A \sqcup B$.

Dica: Sabemos que para qualquer operação contínua $\star : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ e números fuzzy $A, B \in \mathcal{F}(\mathbb{R})$ temos $[A \star B]^\alpha = [A]^\alpha \star [B]^\alpha$.

Questão 2. Considere a base de regras *fuzzy*:

$$\begin{cases} \textbf{SE } x \textbf{ é } A_1 \textbf{ E } y \textbf{ é } A_2, \textbf{ ENTÃO } z \textbf{ é } B_1, \\ \textbf{SE } x \textbf{ é } A_2 \textbf{ E } y \textbf{ é } A_3, \textbf{ ENTÃO } z \textbf{ é } B_2, \\ \textbf{SE } x \textbf{ é } A_3 \textbf{ E } y \textbf{ é } A_1, \textbf{ ENTÃO } z \textbf{ é } B_3. \end{cases} \tag{1}$$

em que

$$A_1 = T(-20, -10, 0), \quad A_2 = T(-10, 0, 10), \quad A_3 = T(0, 10, 20). \tag{2}$$

$$B_1 = T(-2, -1, 0), \quad B_2 = T(-1, 0, 1) \quad \text{e} \quad B_3 = T(0, 1, 2). \tag{3}$$

Esboce o conjunto *fuzzy* deduzido pela base de regras acima usando o método de inferência de Mamdani em $x = 2$ e $y = 3$. Justifique sua resposta explicitando os cálculos.

Questão 3. Mostre que $\mathcal{R} \in \mathcal{F}(U \times U)$ é uma relação de equivalência fuzzy se, e somente se, seus α-níveis $[\mathcal{R}]^{\alpha}$ são relações de equivalência em U para todo $\alpha \in (0, 1]$.

Questão 4.

(a) Sob quais condições $A \in \mathcal{F}(\mathbb{R})$ é um número *fuzzy*?

(b) A afirmação "Se A e B são números *fuzzy* triangulares, então o produto de A e B é também um número *fuzzy* triangular." é verdadeira? Justifique sua resposta.

FOLHA ADICIONAL

FOLHA ADICIONAL

Boa Prova!