UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Centro de Ciências Matemática e da Natureza
Instituto de Matemática

A

CONCURSO PARA PROFESSOR ADJUNTO – DM
Área: Matemática Pura

Nome do Candidato: GARGIULO ACEA, JAVIER NICOLÁS.

Teoria clássica da equação do ~~dolor~~ calor: (5.)

Nós queremos estudar a equação $u_t - \Delta u = 0$ e na versão não homogênea $u_t - \Delta u = f$. onde, $t>0$, $x \in U \subseteq \mathbb{R}^n$ aberto e $u = u(x,t) : \bar{U} \times [0,+\infty) \to \mathbb{R}$ é a variável do problema. Lembra que $\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$ (Laplaciano na variável x)

Observação: A equação é chamada de equação do calor ou de difusão e representa a evolução temporal da distribuição de uma certa quantidade. Por exemplo, a difusão de calor em sólidos. Se $u(x,t)$ representa a temperatura de um sólido num meio homogêneo, a lei de Fourier afirma que $F(x,t) = -K\nabla u(x,t)$ (F fluxo de calor e K constante). Se u é conservada (sem fontes) a equação de conservação correspondente é: $u_t = -\mathrm{div}(F) = K\,\mathrm{div}(\nabla u) = K\Delta u$
a equação do calor mesmo!

Definição 1: A função dada por
$$\Phi(x,t) = \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}} & x \in \mathbb{R}^n,\ t>0 \\ 0 & x \in \mathbb{R}^n,\ t<0 \end{cases}$$
é chamada de solução fundamental ou de núcleo.

Observação 2: O leitor pode verificar que $\Phi_t - \Delta\Phi = 0$ e que $\int_{\mathbb{R}^n} \Phi(x,t)\,dx = 1$ ($\forall t>0$)

Se uma função g representa u em $t=0$ (distribuição inicial), então o problema de Cauchy com condição inicial é

$$\text{(I)} \quad \begin{cases} u_t - \Delta u = 0 & \text{em } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{em } \mathbb{R}^n \times \{0\} \end{cases}$$

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Centro de Ciências Matemática e da Natureza
Instituto de Matemática

CONCURSO PARA PROFESSOR ADJUNTO – DM
Área: Matemática Pura

A ideia da solução é considerar a convolução entre a solução fundamental ϕ com a condição inicial g:

$$u(x,t) = \phi * g(x,t) = \int_{\mathbb{R}^n} \phi(x-y,t)\, g(y)\, dy = \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{\frac{-|x-y|^2}{4t}} g(y)\, dy$$

Teorema 3: Sejam $g \in C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ (contínua e limitada) e $u(x,t) = \phi * g(x,t)$. Então, $u \in C^\infty(\mathbb{R}^n \times (0,+\infty))$ e é uma solução do problema (I) no sentido que $\lim_{(x,t)\to(x_0,0)} u(x,t) = g(x_0)$, $\forall x_0 \in \mathbb{R}^n$.

Prova: Dado que ϕ é de classe C^∞ e tem derivadas limitadas em $\mathbb{R}^n \times [\delta, +\infty)$ com $\delta > 0$, propriedades clássicas do operador convolução implicam que $u \in C^\infty(\mathbb{R}^n \times (0,+\infty))$

Agora, observar que

$$u_t(x,t) - \Delta u(x,t) = \int_{\mathbb{R}^n} (\phi_t - \Delta\phi)(x-y,t)\, g(y)\, dy = 0,$$

onde a primeira igualdade segue de propriedades da derivada com respeito à integral, que são válidas pelas hipóteses dadas sobre g e as propriedades de regularidade de ϕ.

Para provar a última afirmação, vamos usar que $|g(y) - g(x_0)| < \varepsilon$ se $y \in B_\delta(x_0)$. Pela afirmação 2 segue que:

$$|u(x,t) - g(x_0)| = \left| \int_{\mathbb{R}^n} \phi(x-y,t)(g(y) - g(x_0))\, dy \right| \leq \int_{B_\delta(x_0)} \phi(x-y,t)\, |g(y) - g(x_0)|\, dy +$$

$$+ \int_{B_\delta(x_0)^c} \phi(x-y,t)\, |g(y) - g(x_0)|\, dy \leq \varepsilon + 2\|g\|_\infty \int_{B_\delta(x_0)^c} \phi(x-y,t)\, dy$$

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Centro de Ciências Matemática e da Natureza
Instituto de Matemática

CONCURSO PARA PROFESSOR ADJUNTO – DM
Área: Matemática Pura

Agora se $x \in B_{\delta/2}(x_0)$ e $y \in B_\delta(x_0)^c$, pode-se verificar que $|x-y| \geq 2|y-x_0|$. Então

$$\int_{B_\delta(x_0)^c} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x-y|^2}{4t}} dy \leq \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{B_\delta(x_0)^c} e^{-\frac{|y-x_0|^2}{16t}} dy = \frac{C_{TC}}{t^{\frac{n}{2}}} \int_\delta^{+\infty} e^{-\frac{r^2}{16t}} r^{n-1} dr \xrightarrow{t \to 0^+} 0$$

Finalmente, se $|x-x_0| \leq \frac{\delta}{2}$ e t é suficientemente pequeno, temos que

$$|u(x,t) - g(x_0)| < 2\varepsilon \quad \boxed{}$$

Aplicação 4: Se g é contínua, limitada e $g \geq 0$, então a solução u é positiva. Isso pode-se interpretar como que as soluções tem a propriedade de positivo. Isso pode-se interpretar como que as soluções tem a propriedade de velocidade de propagação infinita das perturbações. Ou seja, que uma leve modificação da condição inicial é refletida imediatamente na distribuição em todo o espaço.

Aplicação 5: Se u inicialmente em $t=0$ tem uma "descontinuidade" "ou singularidade" então ela vira-se imediatamente suave para todo $t>0$. Por exemplo se dois corpos isolados* são colocados um no extremo de outro, então "rapidamente" a temperatura nos pontos de conexão vai ser $\frac{T_1+T_2}{2}$, e a gráfica da temperatura vai ser suave para $t>0$. Então, a equação tem a propriedade de "regularizar" funções.

* com temperaturas uniformes T_1 e T_2

Vamos considerar agora o problema não homogêneo:

$$(\mathrm{II}) \begin{cases} u_t(x,t) - \Delta u(x,t) = f(x,t) & \text{em } \mathbb{R}^n \times (0,+\infty) \\ u = 0 & \text{em } \mathbb{R}^n \times \{0\} \end{cases}$$

onde podemos pensar a f como uma fonte de calor. Os resultados anteriores implicam que a função $u(x,t;s) = \int_{\mathbb{R}^n} \Phi(x-y,t-s) f(y,s)\,dy$ $(s>0)$ é solução de

$$\begin{cases} u_t(\cdot,\cdot;s) - \Delta u(\cdot,\cdot;s) = 0 & \text{em } \mathbb{R}^n \times (s,+\infty) \\ u(\cdot,s;s) = f(\cdot,s) & \text{em } \mathbb{R}^n \times \{t=s\} \end{cases}$$

É um princípio geral (Princípio de Duhamel) que a função $u(x,t) = \int_0^t u(x,t;s)\,ds$ deveria ser uma solução de (II)! De fato,

<u>Teorema 6</u>: Sejam $f \in C_1^2(\mathbb{R}^n \times [0,+\infty))$ (com suporte compacto) e $u(x,t) = \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s) f(y,s)\,dy\,ds$. Então $u \in C_1^2(\mathbb{R}^n \times (0,+\infty))$ e é uma solução do problema (II).

(As ferramentas da prova são similares as utilizadas na prova do Teorema 3)

<u>Corolário 7</u>: Dos Teoremas 3 e 6 segue que a função

$$u(x,t) = \int_{\mathbb{R}^n} \Phi(x-y,t) g(y)\,dy + \int_0^t \int_{\mathbb{R}^n} \Phi(x-y,t-s) f(y,s)\,dy\,ds$$

é uma solução do problema geral:

$$\begin{cases} u_t - \Delta u = f & \text{em } \mathbb{R}^n \times (0,+\infty) \\ u = g & \text{em } \mathbb{R}^n \times \{0\} \end{cases}$$

Observação 8: É uma pergunta interessante saber se a solução é única. Nesse contexto existem soluções não triviais do problema homogêneo ($f=0$ e $g=0$). O que vai ser verdadeiro é que a solução descrita é a única fisicamente razoável!

Vamos descrever agora o resultado no caso de domínios limitados. Sejam $U \subseteq \mathbb{R}^n$ aberto limitado e $T>0$. Vamos considerar $U_T = U \times (0,T]$ e $\partial_p U_T = \overline{U_T} \setminus U_T = U \times \{0\} \cup \partial U \times [0,T]$ (fronteira parabólica).

Proposição 9 (princípio fraco do máximo em U_T). Seja $u \in C_1^2(U_T) \cap C(\overline{U_T})$ solução de $u_t - \Delta u = 0$ em U_T. Então, $\max_{\overline{U_T}} u = \max_{\partial_p U_T} u$.

Prova: Primeiro, vamos supor a situação $u_t - \Delta u < 0$ em U_T e que existe $(x_0,t_0) \in U_T$ tal que $u(x_0,t_0) = \max_{\overline{U_T}} u$. É claro que $\frac{\partial^2 u}{\partial x_i^2}(x_0,t_0) \leq 0$. E que se $t_0 < T$, então $u_t(x_0,t_0) = 0$, e se fosse $t_0 = T$, $u_t(x_0,t_0) \geq 0$. Em qualquer caso, segue que $u_t(x_0,t_0) - \Delta u(x_0,t_0) \geq 0$, o que é uma contradição.

Em nosso caso, sabemos que $u_t - \Delta u = 0$ em U_T. Então, vamos considerar $u^\varepsilon(x,t) = u(x,t) - \varepsilon t$, com $\varepsilon > 0$. Segue que $u^\varepsilon_t - \Delta u^\varepsilon < 0$ em U_T. Pelo provado acima sabemos que $\max_{\overline{U_T}} u^\varepsilon = \max_{\partial_p U_T} u^\varepsilon$, e nosso resultado segue de que $u^\varepsilon \rightrightarrows u$ ☒

Corolário 10: No máximo há uma solução $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ do problema

$$\begin{cases} u_t - \Delta u = f & \text{em } U_T \\ u = g & \text{em } \partial_p U_T \end{cases}$$

Prova: Segue de ~~[illegible]~~ aplicar a Proposição 9 nas funções $u_1 - u_2$ e $u_2 - u_1$ com u_1 e u_2 duas soluções.

Também, pode-se provar um princípio forte do máximo e existem dois caminhos. Um é provar uma fórmula do valor médio para nosso problema, e outro é utilizar a desigualdade de Harnack. Vamos seguir esse último um e provar a correspondente.

Proposição 11 (Desigualdade de Harnack) Seja $u \in C^2_1(U_T)$, com $u \geq 0$, uma solução de $u_t - \Delta u = 0$ em U_T. Se $V \subseteq \bar{V} \subseteq U$ e $0 < t_1 < t_2 \leq T$, então existe uma constante C (que só depende de $dist(V, \partial V)$, t_1, t_2 e n) tal que: $\max_{\bar{V}} u(\cdot, t_1) \leq C \min_{\bar{V}} u(\cdot, t_2)$.

Como consequência temos que:

Teorema 12 (princípio forte do máximo em U_T) Seja $u \in C^2_1(U_T) \cap C(\bar{U}_T)$ uma solução de $u_t - \Delta u = 0$ em U_T. Se existe $(x_0, t_0) \in U_T$ tal que $u(x_0, t_0) = \max_{\bar{U}_T} u$, então u é constante em $U_{t_0} = U \times (0, t_0]$.

Prova: Segue de utilizar a desigualdade de Harnack para estudar a função $v(x,t) = u(x_0, t_0) - u(x,t)$.

Aplicação 13: O princípio forte implica que as soluções do problema em domínios limitados também tem a propriedade de velocidade de propagação infinita das perturbações (Ap. 4).

O Teorema 12 é o fenômeno fundamental no caso de um princípio do máximo em todo $\mathbb{R}^n$, se supomos que as soluções tem crescimento limitado para $|x|>>0$.

Teorema 14 (Princípio do máximo em $\mathbb{R}^n$) Seja $u \in C_1^2(\mathbb{R}^n\times(0,T]) \cap C(\mathbb{R}^n\times[0,T])$ uma solução do problema

$$\begin{cases} u_t - \Delta u = 0 & \text{em } \mathbb{R}^n\times(0,T) \\ u = g & \text{em } \mathbb{R}^n\times\{0\} \end{cases}$$

Se existem constantes $A>0$ e $a>0$ tais que: $|u(x,t)| \leq Ae^{a|x|^2}$ $\forall (x,t)\in \mathbb{R}^n \times[0,T]$

[Propriedade (*)]

Então, $\sup_{\mathbb{R}^n\times[0,T]} u = \sup_{\mathbb{R}^n} g$

Como consequência podemos deduzir a unicidade de soluções fisicamente razoáveis do problema de Cauchy:

Aplicação 15: Sejam $f \in C(\mathbb{R}^n\times[0,T])$ e $g\in C(\mathbb{R}^n)$, então existe uma única solução $u\in C_1^2(\mathbb{R}^n\times(0,T])\cap C(\mathbb{R}^n\times[0,T])$ do problema:

$$\begin{cases} u_t - \Delta u = f & \text{em } \mathbb{R}^n \times (0,T) \\ u = g & \text{em } \mathbb{R}^n \times \{0\} \end{cases}$$

satisfazendo a propriedade (*) do Teorema anterior.

$\boxed{\text{Aplicação / Comentário Geral}}$ A equação do calor surge na modelagem de muitos fenômenos probabilísticos. Por exemplo, a solução fundamental (caso $n=1$) coincide (no limite) com a função de densidade da variável de posição de uma partícula que faz um passeio aleatório na reta. Além das dimensões e para isso utiliza o Teorema Central do Limite.

Similarmente, a equação é muito utilizada no estudo de movimentos Brownianos. Também surge na modelagem de variáveis financeiras, como o reconhecido modelo de Black-Scholes.

Seguindo a aplicação 5, a equação também é utilizada na teoria de processamento de imagens.

Teorema de Hahn-Banach: (6.)

(I) Versão analítica do Teorema: É um resultado importante e clássico da análise funcional, e a ferramenta fundamental na prova de muitos outros teoremas. Nesta versão o teorema descreve a ~~trivial~~ possível extensão de uma forma linear definida num subespaço vetorial para todo o espaço.

Teorema 1 (HB) Sejam V um espaço vetorial sobre $\mathbb{R}$ e $P: V \to \mathbb{R}$ uma aplicação que satisfaz as condições:

$$1)\ P(\lambda x) = \lambda P(x) \quad \forall \lambda \in \mathbb{R}, \forall x \in V$$
$$2)\ P(x+y) \leq P(x) + P(y), \ \forall x, y \in V$$
(função sublinear ou subnorma)

Se $W \subseteq V$ é um subespaço vetorial e $g: W \to \mathbb{R}$ uma aplicação linear tal que $g(x) \leq P(x), \forall x \in W$, então existe $f: V \to \mathbb{R}$ linear que estende a g: $f(x) = g(x)\ \forall x \in W$, e tal que $f(x) \leq P(x), \forall x \in V$.

Vamos precisar do seguinte resultado clássico:

Lema de Zorn: Seja $P \neq \emptyset$ um conjunto parcialmente ordenado, tal que para todo subconjunto $Q \subseteq P$ totalmente ordenado, existe uma cota superior $c \in P$. Então, existe um elemento $m \in P$ maximal, ou seja, $m \leq x$ implica que $m = x$.

Prova do Teorema 1: Vamos considerar

$$X = \{ h \mid h: D(h) \to \mathbb{R} \text{ é uma aplicação linear com } D(h) \subseteq V \text{ tal que } W \subseteq D(h),\ h \text{ estende } e\ g \text{ e } h(x) \leq p(x)\ \forall x \in D(h) \}$$

O conjunto X é ordenado com a relação:

$$h_1 \leq h_2 \iff D(h_1) \subseteq D(h_2) \text{ e } h_2 \text{ estende a } h_1$$

Seja $Q = \{h_i : D(h_i) \to \mathbb{R}\}_{i \in I} \subseteq P$ um subconjunto totalmente ordenado. Vamos considerar:

$$D(h) = \bigcup_{i \in I} D(h_i) \text{ e } h: D(h) \to \mathbb{R} \text{ dado por } h(x) = h_i(x) \text{ se } x \in D(h_i).$$

Que Q seja totalmente ordenado implica que $D(h) \subseteq V$ é subespaço e que h seja bem definida. Pode-se verificar facilmente que $W \subseteq D(h)$, h estende a g e que $h(x) \leq p(x)\ \forall x \in D(h)$. De fato, h é cota superior de Q. Então, pelo Lema de Zorn, existe $(b: D(b) \to \mathbb{R}) \in P$ maximal para P. Vamos provar que $D(b) = V$. Se fosse $D(b) \subsetneq V$, existe $x_0 \notin D(b)$ e poderíamos considerar:

$$D(h) = D(b) \oplus \langle x_0 \rangle \text{ e } h(x + t x_0) = b(x) + t\alpha \ ,\ x \in D(b),$$

com α a ser definido. Em particular, precisamos que $b(x) + t\alpha \leq p(x + t x_0)$ $\forall x \in D(b)$. Pelas propriedades de P, isso é equivalente a:

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Centro de Ciências Matemática e da Natureza
Instituto de Matemática

CONCURSO PARA PROFESSOR ADJUNTO – DM
Área: Matemática Pura

$$\begin{cases} b(x) - \alpha \leq p(x - x_0) \\ b(x) + \alpha \leq p(x + x_0) \end{cases}$$

Então, podemos escolher α com: $\sup_{x \in D(b)} b(x) - p(x - x_0) \leq \alpha \leq \inf_{x \in D(b)} p(x + x_0) - b(x)$

Segue que $h \in P$ e que $b < h$, o que é uma contradição. Então, $D(b) = V$. $\blacksquare$

Vamos desenvolver agora alguns corolários e aplicações desta versão do Teorema.

Definição: Dado $(E, \|\ \|)$ um espaço vetorial normado, vamos definir

$E' = \{ b : E \to \mathbb{R} \mid b$ é linear e contínua$\}$ (dual topológico de E)

e para cada $b \in E$, a norma:

$$\|b\|_{E'} = \sup_{\substack{x \in E \\ \|x\| \leq 1}} |b(x)| = \sup_{\substack{x \in E \\ \|x\| = 1}} |b(x)|.$$

Também, vamos escrever $\langle b, x \rangle = b(x)$.

Corolário 1: Sejam $G \subseteq E$ subespaço e $g \in G'$. Então existe $b \in E'$ tal que b estende a g e $\|g\|_{G'} = \|b\|_{E'}$.

Prova: Segue de aplicar o Teorema 1 com $p(x) = \|g\|_{G'} \|x\|$.

Corolário 2: Dado $x_0 \in E$, existe $b_0 \in E'$ tal que $\|b_0\|_{E'} = \|x_0\|$ e $\langle b_0, x_0 \rangle = \|x_0\|^2$.

Prova: Basta aplicar o Corolário 1 para $G = \langle x_0 \rangle$ e $g(tx_0) = t\|x_0\|^2$.

Corolário 3: Para cada $x \in E$, $\|x\| = \sup_{\substack{b \in E' \\ \|b\| \leq 1}} |\langle b, x\rangle| = \max_{\substack{b \in E' \\ \|b\| = 1}} |\langle b, x\rangle|$.

Prova: É claro que $\sup_{\|b\| \leq 1} |\langle b, x\rangle| \leq \|x\|$. Agora, pelo Corolário 2 existe $b_0 \in E'$ tal que $\|b_0\| = \|x\|$ e $\langle b_0, x\rangle = \|x\|^2$. Então, se $x \neq 0$ e $b_1 = \frac{1}{\|x\|} b_0$ segue que $\|b_1\| = 1$ e $\langle b_1, x\rangle = \|x\|$.

(II) Versão geométrica do teorema: Neste contexto o teorema trabalha sobre a separação de conjuntos convexos. Agora, $E = (E, \|\,\|)$ sempre vai ser um espaço normado.

Definição: Sejam $b: E \to \mathbb{R}$ linear e $\alpha \in \mathbb{R}$. O conjunto $H = \{x \in E / b(x) = \alpha\}$ é chamado de hiperplano de E.

Observação: Um hiperplano $H \subseteq E$ é um conjunto fechado se e somente se b é contínua ($b \in E'$).

Definição: Sejam $A, B \subseteq E$ dois subconjuntos. Vamos dizer que um hiperplano $H = \{x \in E / b(x) = \alpha\}$ separa A e B no sentido estrito se existe $\varepsilon > 0$ tal que:

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
Centro de Ciências Matemática e da Natureza
Instituto de Matemática

CONCURSO PARA PROFESSOR ADJUNTO – DM
Área: Matemática Pura

$$b(x) \leq \alpha - \varepsilon \ \forall x \in A \text{ e } b(x) \geq \alpha + \varepsilon \ \forall x \in B.$$

E separa no sentido amplo se:

$$b(x) \leq \alpha \ \forall x \in A \text{ e } b(x) \geq \alpha \ \forall x \in B.$$

Também, vamos lembrar que $C \subseteq E$ é convexo se:

$$tx + (1-t)y \in C, \ \forall x, y \in C, \ \forall t \in [0,1].$$

Teorema 2 (HB) Sejam $A, B \subseteq E$ subconjuntos não vazios, convexos e tais que $A \cap B = \emptyset$. Suponhamos que A é aberto. Então, existe um hiperplano fechado H que separa A e B no sentido amplo.

Vamos precisar dos seguintes lemas:

Lema 1 (Funcional de Mincowski) Seja $C \subseteq E$ aberto e convexo, com $0 \in C$. Para cada $x \in E$ vamos considerar $P(x) = \inf\{\alpha \in \mathbb{R}_{>0} / \alpha^{-1} x \in C\}$. Então, a aplicação P satisfaz as condições 1) e 2) do Teorema 1 (HB) e as propriedades:

3) Existe M constante tal que $P(x) \leq M \|x\|$;

4) $C = \{x \in E / P(x) < 1\}$.

FAVOR NÃO ESCREVER NO VERSO!
NÃO ULTRAPASSAR A LINHA DA MARGEM!

Prova: Por simplicidade e tempo só somos para 3) e 4). Dado que $0 \in C$ e C é aberto, existe $r>0$ tal que $B(0,r) \subseteq C$. Segue que $P(x) \leq \frac{1}{r} \|x\|$.

Agora, só para 4) observar que se $x \in C$, então $(1+\varepsilon)x \in C$ para $\varepsilon>0$ pequenho. Então, $P(x) \leq \frac{1}{1+\varepsilon} < 1$. Por outro ~~lado~~ lado se $P(x)<1$, existe α com $0<\alpha<1$ tal que $\alpha^{-1}x \in C$. Pela convexidade de C, segue que $\alpha\alpha^{-1}x + (1-\alpha)0 = x \in C$. $\boxtimes$

Lema 2: Seja $C \subseteq E$ não vazio, aberto e convexo. Então, para cada $x_0 \in E \setminus C$, existe $f \in E'$ tal que $f(x) < f(x_0)$, $\forall x \in C$. Em particular, $H = \{x \in E / f(x) = f(x_0)\}$ separa $\{x_0\}$ e C no sentido amplo.

Prova: Por ~~transl~~ translação podemos supor que $0 \in C$. Seja P o funcional definido no lema anterior para C. Vamos considerar $G = \langle x_0 \rangle$ e $g(Tx_0) = T$. É claro que $g(x) \leq P(x)$, $\forall x \in G$ (pois $g(x_0) = 1 \leq P(x_0)$). Pelo Teorema 1, existe $f: E \to \mathbb{R}$ linear que estende a g e tal que $f(x) \leq P(x)$, $\forall x \in E$. Pelas propriedades 3) e 4) de P segue que f é ~~con~~ contínua (limitado) e que $f(x) \leq P(x) < 1 = f(x_0)$.

$\boxtimes$

Prova do Teorema 2: Seja $C = A - B$. Pode-se verificar que C é não vazio, aberto e convexo. De fato, $0 \notin C$. Então, pelo Lema 2 existe $f \in E'$ tal que $f(z) < 0$, $\forall z \in C$. Em particular, $f(x) < f(y)$, $\forall x \in A$ e $\forall y \in B$. Vamos escolher α com a propriedade: $\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y)$. Segue que o hiperplano $H = \{x \in E \,/\, f(x) = \alpha\}$ separa A e B no sentido amplo. $\square$

Também, existe uma versão do Teorema para separar no sentido estrito:

Teorema 3 (HB) Sejam $A, B \subseteq E$ subconjuntos não vazios, convexos e disjuntos. Suponhamos que A é fechado e B é compacto. Então, existe um hiperplano fechado H que separa A e B no sentido estrito.

Ideia da prova: Vamos considerar $A_\varepsilon = A + B(0,\varepsilon)$ e $B_\varepsilon = B + B(0,\varepsilon)$, com $\varepsilon > 0$. Pode-se verificar que A_ε e B_ε são abertos e convexos. Além disso, pode-se provar que para ε pequeno, são disjuntos. Para isso, é fundamental a hipótese de A fechado e B compacto. Finalmente, pelo Teorema 2 existe um hiperplano fechado H que separa A_ε e B_ε no sentido amplo, e que também vai separar A e B no sentido estrito. $\square$

Aplicações e comentários finais:

(A) O seguinte resultado é muito útil para provar que um subespaço é denso.

Aplicação: Seja $F \subseteq E$ subespaço tal que $\bar{F} \neq E$. Então, existe $b \in E'$ $(b \neq 0)$ tal que $b(x) = 0$ $\forall x \in F$.

Prova: Se ~~x~~ $x_0 \in E \setminus \bar{F}$, então $\{x_0\}$ e $\bar{F}$ podem ser separados no sentido estrito por $H = \{x \in E / b(x) = \alpha\}$. Então, $b(x) < \alpha < b(x_0)$, $\forall x \in F$. Em particular, $\lambda b(x) < \alpha$, $\forall \lambda \in \mathbb{R}, \forall x \in F$. O que implica $b(x) = 0$, $\forall x \in F$.

Como consequência:

Corolário: Sejam $F \subseteq E$ subespaço e ~~...~~ $x \in E$. Então, $x \in \bar{F}$ se e somente se para cada $b \in E'$ com $b|_F \equiv 0$, segue que $b(x) = 0$.

(B) Agora, seja $E = (E, \| \ \|)$ um espaço de Banach (normado e completo)

Definição: O espaço E é chamado de reflexivo se a aplicação linear $\varphi: E \to E'' = (E')'$ é um isomorfismo (se sempre é injetiva, e pelos corolários do teorema 1 resulta uma isometria)

O Teorema de Hahn-Banach é a ferramenta fundamental na prova de:

Teorema: Sejam E um espaço de Banach reflexivo e $F \subseteq E$ subespaço fechado. Então, F também é reflexivo.

(C) Comentário / Aplicação: O Teorema de Hahn-Banach também tem aplicações na teoria de equações diferenciais em derivadas parciais. Por exemplo, é uma ferramenta fundamental na prova de existência de uma solução fundamental para todo operador diferencial com coeficientes constantes, (Teorema de Malgrange - Ehrenpreis), e também na prova de existência de uma função de Green para o biharmônico no método de Lax - Garabedian.

(D) Uma outra interpretação interessante do Teorema, no contexto da Álgebra homológica, o teorema é que o operador "dualizar" preserva sequências exatas curtas de espaços de Banach. Ou seja, se

$$0 \to E_1 \xrightarrow{\varphi} E_2 \xrightarrow{\psi} E_3 \to 0$$

é uma sequência exata de espaços de Banach E_i, então

$$0 \to E_3' \xrightarrow{\psi'} E_2' \xrightarrow{\varphi'} E_1' \to 0$$

também é exata. O teorema permite provar exatamente que φ' é sobrejetiva.

Forma canônica de Jordan: 7.

Seja V um espaço vetorial sobre $K = \mathbb{R}$ ou $\mathbb{C}$ de dimensão finita n.

Observação: Não todo operador linear $T: V \to V$ é diagonalizável, ou seja, nem sempre existe uma base B de V, tal que $[T]_B^B$ é diagonal. Mas, para nossas aplicações é suficiente que exista B tal que $[T]_B^B$ tenha uma forma bastante próxima da forma diagonal, que chamaremos de forma canonica de Jordan.

Definição 1: Uma matriz quadrada de $m \times m$ da forma $J_\lambda = \begin{pmatrix} \lambda & & 0 \\ 1 & \ddots & \\ 0 & 1 & \lambda \end{pmatrix}$ é chamada de bloco de Jordan associado a $\lambda \in K$. E uma matriz J de $n \times n$ está na forma canonica de Jordan, se ela é da forma

$$J = \begin{pmatrix} J_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_L} \end{pmatrix}$$ com J_{λ_K} blocos de Jordan.

Definição 2: Seja $T: V \to V$ linear. Um $v \in V$ ($v \neq 0$) é chamado de autovetor generalizado de T associado a $\lambda \in K$, se existe $k \in \mathbb{N}$ tal que $(T - \lambda I_d)^k (v) = 0$.

Se existe uma base $B = \{ v_{i_K}^{(K)} \}_{i_K = 1 \ldots P_K}^{K = 1 \ldots L}$ de V tal que $[T]_B^B$ está na forma canônica de Jordan, então:

\# 18

$$(*) \quad \begin{cases} T(v_{i_k}^{(k)}) = \lambda_k v_{i_k}^{(k)} + v_{i_k+1}^{(k)} \text{ se } i_k < p_k \\ T(v_{p_k}^{(k)}) = \lambda_k v_{p_k}^{(k)} \end{cases}$$

(Bloco de Jordan de tamanho p_k associado a λ_k)

Essas relações implicam que:

$$\begin{cases} (T - \lambda_k I_d)(v_{p_k}^{(k)}) = 0 \\ \vdots \\ (T - \lambda_k I_d)^{p_k}(v_1^{(k)}) = 0 \end{cases}$$

Então, nosso problema é achar uma base de V formada por autovetores generalizados de T satisfazendo (*).

Definição 3: O subespaço $W_\lambda = \{v \in V / \exists k \in \mathbb{N} \text{ com } (T-\lambda I_d)^k(v) = 0\}$ é chamado de autoespaço generalizado de T associado a $\lambda \in K$.

Vamos precisar dos seguintes resultados sobre autoespaços generalizados. Não vamos incluir as provas correspondentes, mas em geral são elementares.

Proposição 4: a) Se $\lambda \in K$ é autovalor de T, então W_λ é T invariante ($T(W_\lambda) \subseteq W_\lambda$).

b) Se $\lambda \neq \mu$ são autovalores, $W_\lambda \cap W_\mu = \{0\}$.

Proposição 5: Seja T tal que, se $P_T(x) = \det(T - x\,Id)$ é seu polinômio característico, $P_T(\lambda) = 0$ implica $\lambda \in K$. Seja $\lambda \in K$ uma raiz de P_T de multiplicidade m, então

a) $\dim(W_\lambda) \leq m$ b) $W_\lambda = \ker(T - \lambda I_d)^m$.

Teorema 6: Seja $T: V \to V$ linear tal que seu polinômio característico é da forma $P_T(x) = (-1)^n (x-\lambda_1)^{m_1} \cdots (x-\lambda_l)^{m_l}$ com λ_k autovalores distintos. Se $W_1 \ldots W_l$ são os autoespaços associados a cada λ_k, então temos que:

a) $V = W_1 \oplus \cdots \oplus W_l$ b) $\dim W_k = m_k$, $k = 1 \ldots l$.

Prova: Vamos provar o caso $l = 2$. A prova do caso geral é parecido, só que precisa de um argumento indutivo. Coloquemos $P_1(x) = (x-\lambda_1)^{m_1}$ e $P_2(x) = (x-\lambda_2)^{m_2}$. Dado que P_1 e P_2 são primos entre si, existem polinômios q_1 e q_2 tais que: $1 = q_1 P_1 + q_2 P_2$. Então, para cada $v \in V$ temos que:

$$v = q_1(T) \circ P_1(T)(v) + q_2(T) \circ P_2(T)(v) = w_1 + w_2.$$

Pelo Teorema de Hamilton-Cayley ($P_T(T) = 0$) segue que:

$$(T - \lambda_1 I_d)^{m_1}(w_2) = q_2(T) \circ (-1)^n P_T(T)(v) = 0.$$

Portanto, $w_2 \in W_1$ (e $w_1 \in W_2$). Pela Proposição 4 também sabemos que $W_1 \cap W_2 = \{0\}$. O que implica que $V = W_1 \oplus W_2$ e que $\dim W_k = m_k$, pois $\dim W_k \leq m_k$ e $m_1 + m_2 = n$. ☒

Pelo Teorema 6 para terminar de resolver nosso problema precisamos que $T|_{W_\lambda}$ entre na forma canônica de Jordan.

Definição 7: Sejam $\lambda \in k$ autovalor de T, $v \in W_\lambda$ e $p \in \mathbb{N}$ o menor número natural tal que $(T-\lambda I_d)^p(v) = 0$. O conjunto:

$$C_v = \{v, (T-\lambda I_d)(v), \dots, \underbrace{(T-\lambda I_d)^{p-1}(v)}_{v^b}\}$$

é chamado de ciclo de autovetores generalizados associados a $\lambda \in k$.

~~Observ~~ Observação 8: O vetor final do ciclo v^b sempre é um autovetor de T associado a $\lambda \in k$.

Proposição 9: Sejam $\lambda \in k$ autovalor e $\{C_i\}_{i=1}^{q}$ ciclos de autovetores generalizados associados a λ tais que $\{v_i^b\}_{i=1}^{q}$ são linearmente independentes (LI). Então, $C = \bigcup_{i=1}^{q} C_i$ é um conjunto LI.

Prova: Seja $C_i = \{v_i, \dots, (T-\lambda I_d)^{p_i-1}(v_i)\}$. Vamos fazer indução em $\#C = p = \sum p_i$. Se $p=1$ o resultado é claramente certo. Então,

Vamos supor que $p>1$ e que o resultado é verdadeiro se $\#C<p$.
Consideremos $x_{ij}\in K$ tais que:

$$\sum_{i=1}^{q}\sum_{j=0}^{p_i-1}(T-\lambda I_d)^j(v_i)=0.$$

Aplicando $(T-\lambda I_d)$ na igualdade temos que:

$$\sum_{i=1}^{q}\sum_{j=0}^{p_i-2}(T-\lambda I_d)^{j+1}(v_i)=0.$$

A expressão é uma combinação de vetores em $C'=\bigcup_{i=1}^{q}C_i'$, com $C_i'=\{(T-\lambda I_d)(v_i),\dots,(T-\lambda I_d)^{p_i-2}((T-\lambda I_d)(v_i))\}$. Observe que os vetores além dos C_i' são os mesmos do que os de C_i, mas $\#C_i'=\#C_i-1$. Pela hipótese segue que $x_{ij}=0$ para $i=1..q$ e $j=0..p_i-2$. Mas, como $\{v_i^b\}_{i=1}^{q}$ são LI (caso $j=p_i-1$) também segue que $x_{i(p_i-1)}=0$. ☑

Definição 10: Um operador linear $T:V\to V$ é chamado de nilpotente de índice p se: $T^p=0$ e $T^{p-1}\neq 0$.

Agora, vamos provar o teorema:

↓

Teorema 11: Seja $T: V \to V$ linear tal que seu polinômio característico é da forma $p_T(x) = (-1)^n (x-\lambda_1)^{m_1} \cdots (x-\lambda_l)^{m_l}$, com λ_k autovalores distintos.
Então,
a) Para cada $1 \leq k \leq l$, existe uma base de W_{λ_k} formada por ciclos de autovetores generalizados de T associados a λ_k.
b) Existe uma base C de V formada por ciclos de autovetores generalizados, tal que $[T]_C^C$ está na forma canônica de Jordan.

Prova: Fixemos $\lambda \in K$ autovalor, e consideremos $U: W_\lambda \to W_\lambda$ com $U = (T - \lambda Id)|_{W_\lambda}$. É claro que $v \in W_\lambda$ se e somente existe $k \in \mathbb{N}$ com $U^k(v) = 0$. Então, a prova reduz-se ao caso nilpotente.

Seja $p = \dim(W_\lambda)$. Vamos usar que o resultado é verdadeiro quando a dimensão do autoespaço é menor do que p.

A imagem $Im(U)$ é U-invariante, então vamos considerar $U_I = U|_{Im(U)}$. Dado que $\dim Ker(U) > 0$, segue que $\dim Im(U) < p$. Portanto, pela hipótese, existem $C_1' \dots C_k'$ ciclos de autovetores generalizados de U_I (associados a 0) tais que $C' = \bigcup_{i=1}^{k} C_i'$ é uma base de $Im(U)$.

Escrevendo $C_i' = \{w_i, \dots, U^{l_i}(w_i)\}$, e dado que $w_i \in \mathrm{Im}(U)$, existem $v_i \in W_\lambda$ tais que $C_i' = \{U(v_i), \dots, U^{l_i+1}(v_i)\}$. Vamos escrever $C_i = C_i' \cup \{v_i\}$. Observe que os vetores livres $U^{l_i+1}(v_i)$ estão no núcleo de U. Completemos a uma base $\{U^{l_i+1}(v_i)\}_{i=1}^{k} \cup \{v_{k+1}, \dots, v_r\}$ de $\ker U$. Agora, vamos definir:

$$C = \bigcup_{i=1}^{k} C_i \cup \{v_{k+1}, \dots, v_r\}$$

Pode-se verificar facilmente que $\# C = \dim \ker U + \dim \mathrm{Im} U = \dim(W_\lambda)$. Pela Proposição 9, segue que C é uma base de W_λ formada por ciclos de autovetores generalizados.

b) Utilizando e) e que $V = W_{\lambda_1} \oplus \cdots \oplus W_{\lambda_\ell}$ (Teorema 6) existe uma base de V:

$$C = \bigcup_{k=1}^{\ell} \bigcup_{j=1}^{r_k} C_j^{(k)}$$

formada por ciclos da forma

$$C_j^{(k)} = \{v_j^{(k)}, \dots, (T - \lambda_k I_d)^{m_j^k}(v_j^{(k)})\} \subseteq W_{\lambda_k}.$$

É fácil verificar que $[T]_C^C$ está na forma canônica de Jordan, usando que os vetores de cada ciclo satisfazem as equações $(*)$ de página 19.

Comentários e aplicações breves:

Observação 12: Para cada autovalor λ_k existem r_k ciclos $C_1^{(k)} \dots C_{r_k}^{(k)}$ que induzem r_k blocos de tamanhos $m_1^k \dots m_{r_k}^k$. Pode-se verificar que $\sum_j m_j^k = m_k$ é a multiplicidade de λ_k no polinômio característico p_T, e que $\max_j m_j^k$ é a multiplicidade no polinômio mínimo de T: m_T.

Muitas aplicações do Teorema precisam da existência e da unicidade da forma canônica de Jordan. A prova da unicidade é uma consequência do seguinte resultado:

Lema 13: Seja $T: V \to V$ um operador linear nilpotente de índice p. Então, para cada $1 \leq i \leq p$, a quantidade de blocos de Jordan de tamanho $i \times i$ é $\operatorname{rank}(T^{i+1}) - 2\operatorname{rank}(T^{i}) + \operatorname{rank}(T^{i-1})$.

Se nós reduzirmos do caso geral para o caso nilpotente, podemos provar que:

Teorema 14: Nas hipóteses do Teorema 11, existe uma única forma de Jordan $J \in K^{n \times n}$ (a menos da ordem dos blocos) tal que para alguma base B de V, $[T]_B^B = J$.

Uma aplicação interessante é na classificação de matrizes por conjugação. Se $T, T' \in K^{n\times n}$, vamos dizer que: $T \sim T'$ se existe C inversível tal que $T' = CTC^{-1}$.

Aplicação/Teorema: Sejam $T, T' \in \mathbb{C}^{n\times n}$ e J_T e $J_{T'}$ suas formas de Jordan associadas. Então, $T \sim T' \iff J_T = J_{T'}$ (a menos de ordem)

Prova: $\Leftarrow$) Segue de que "$\sim$" é uma relação de equivalência e de que $T \sim J_T$ e $T' \sim J_{T'}$.
$\Rightarrow$) Se $T \sim T'$ então $J_T \sim J_{T'}$, e serem duas formas canônicas de Jordan para a mesma aplicação.

Aplicação/Exemplo: Se $T, T' \in \mathbb{C}^{3\times 3}$, então $T \sim T' \iff P_T = P_{T'}$ e $m_T = m_{T'}$

(Segue de que os autovalores e os tamanhos dos blocos podem ser lidos, neste caso especial $n=3$, de informações que só dependem de P_T ~~era~~ e de m_T.

Observação: Uma interpretação importante do teorema de Jordan, é que (a menos de uma mudança de base) determina uma decomposição de T

$$CTC^{-1} = J_T = D + N,$$ D diagonal e N nilpotente.

E isso é muito útil para calcular as potências de uma matriz, utilizando uma fórmula do Binômio de Newton para matrizes que comutam.

Exemplo: Com essa mesma ideia, é possível definir e descrever a exponencial de uma matriz como: $e^T = \sum_{k \geq 0} \frac{T^k}{k!}$

Além das questões de convergência, obtemos a seguinte:

Aplicação: Seja $\vec{x}'(s) = T \cdot \vec{x}(s)$ um sistema de equações diferenciais lineares induzido por $T \in K^{n \times n}$ (nas hipóteses do Teorema). Então, $\vec{x}(s) = \vec{x}(0)\, e^{sT}$ determina uma solução fundamental, onde

$$e^{sT} = C^{-1} e^{sJ_T} C = C^{-1} e^{sD} e^{sN} C =$$

$$= C^{-1} \begin{pmatrix} e^{s\lambda_1} & & & & 0 \\ & \ddots & & & \\ & & e^{s\lambda_1} & & \\ & & & \ddots & \\ 0 & & & & e^{s\lambda_l} \end{pmatrix} \left(\sum_{j=0}^{p} \frac{s^j}{j!} N^j \right) C.$$