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

CONCURSO PARA PROFESSOR ADJUNTO – DM
Área: Matemática Pura

F

Nome do Candidato:
LUCCAS CASSIMIRO CAMPOS

TEOREMA DE HAHN-BANACH E APLICAÇÕES

· INTRODUÇÃO

O teorema conhecido como HAHN-BANACH versa, em uma formulação moderna, sobre a extensão de funcionais lineares em um espaço normado. Foi provado por Helly em 1912 em um contexto de espaços separáveis, o que permitiu usar indução finita, e depois generalizado por Hahn e Banach na década de 1920. É um dos teoremas que formam a base da chamada "Análise Funcional" e tem diversas aplicações nesse contexto. Veremos sua demonstração e algumas destas aplicações. Comecemos por algumas definições.

Definição 1: Um funcional linear $f: V \to K$ é uma função cujo domínio é um espaço vetorial V e K é o corpo associado, que assumiremos ser $\mathbb{R}$ ou $\mathbb{C}$, e que satisfaz

$$f(x+\alpha y) = f(x) + \alpha f(y) \quad \forall x,y \in K \quad \forall \alpha \in K$$

Um funcional sublinear é uma função $\rho : V \to \mathbb{R}$, com V espaço vetorial, que satisfaz

$$\rho(x+y) \leq \rho(x) + \rho(y) \qquad \forall\, x, y \in V$$
$$\rho(\alpha x) = \alpha \rho(x) \qquad \forall\, x \in V,\ \forall\, \alpha \geq 0$$

Um funcional sublinear será chamado de seminorma se

$$\rho(\alpha x) = |\alpha|\, \rho(x) \qquad \forall\, x \in V,\ \forall\, \alpha \in \mathbb{R}$$

Note que toda norma é uma seminorma.

Com essas definições, podemos enunciar e provar o teorema principal desta seção:

Teorema (HAHN – BANACH): Seja E um espaço vetorial real e $V \subset E$ um subespaço. Seja $\rho : E \to \mathbb{R}$ um funcional sublinear e suponha que exista $f : V \to \mathbb{R}$ funcional linear tal que

$$f(x) \leq \rho(x) \qquad \forall\, x \in V.$$

ENTÃO existe funcional linear $\hat{f} : E \to \mathbb{R}$ tal que

$$\hat{f}(x) \leq \rho(x) \qquad \forall\, x \in E$$

Tal funcional $\hat{f}$ pode ser escolhido de forma que

$$\hat{f}(x) = f(x) \quad \forall x \in V,$$

ou seja $\hat{f}|_V = f$. Chamaremos $\hat{f}$ de extensão de f.

Observação: O teorema enunciado diz que qualquer funcional (real) em um espaço vetorial limitado por um funcional sublinear pode ser estendido para todo o espaço, mantendo a limitação.

Ideia do Teorema de Hahn Banach:

Seja $x_0 \in E$ tal que $x_0 \notin V$. Queremos definir $\hat{f}(x + t x_0)$, com $x \in V$ e $t \in \mathbb{R}$. Observe que, $\forall x, y \in V$:

$$f(x) + f(y) = f(x+y) \leq \rho(x+y) = \rho(x + x_0 + y - y_0)$$
$$\leq \rho(x + x_0) + \rho(y - y_0)$$

Assim:

$$f(y) - \rho(y - y_0) \leq \rho(x + x_0) - f(x) \qquad \forall x, y \in V.$$

seja α tal que $\sup_{y \in V} \{f(y) - \rho(y - y_0)\} \leq \alpha \leq \inf_{x \in V} \{\rho(x + x_0) - f(x)\}$

e defina $\tilde{f}(x + t x_0) = f(x) + t\alpha$. É claro que $\tilde{f}$ é linear em span$\{V \cup \langle x_0 \rangle\}$. Mostraremos que ainda é cotada superiormente por ρ. Temos, se $t > 0$:

$$\tilde{f}(x + t x_0) = t\,\hat{f}\left(\frac{x}{t} + x_0\right) = t\left[f\left(\frac{x}{t}\right) + \alpha\right]$$
$$\leq t\,\rho\left(\frac{x}{t} + x_0\right)$$
$$= \rho(x + t x_0)$$

E, se $t < 0$:

$$\hat{f}(x + t x_0) = -t\,\hat{f}\left(\frac{-x}{t} - x_0\right) \leq -t\left(\rho\left(\frac{x}{-t} - x_0\right)\right) = \rho(x + t x_0)$$

Logo, provamos que podemos estender f para span$\{V \cup \langle x_0 \rangle\}$.
A extensão no espaço todo E segue por indução transfinita (comumente existe nos livros de Análise Funcional como um "axioma" equivalente ao axioma de escolha, sob a forma do Lema de Zorn. Banach, no entanto, utilizou indução transfinita diretamente em sua prova)

$\square$

Veremos alguns corolários do teorema de Hahn – Banach. O primeiro deles versa sobre a extensão de funcionais lineares limitados.

Definição: Um funcional linear $f: E \to K$ em um espaço normado E é dito limitado se

$$\sup_{\|x\| \leq 1} |f(x)| < +\infty.$$

Para tais funcionais, definimos sua norma:

$$|||f||| = \sup_{\|x\| \leq 1} |f(x)| = \inf \{ c > 0 \ ; \ |f(x)| \leq c\|x\| \ \forall x \in E \}.$$

Corolário 1: Seja $V \subset E$ subespaço de um espaço vetorial normado e $f: V \to \mathbb{R}$ funcional linear limitado. Então existe extensão $\tilde{f}: E \to \mathbb{R}$ tal que $|||\tilde{f}||| = |||f|||$.

Prova: Defina o funcional sublinear $p: E \to \mathbb{R}$ como

$$p(x) = |||f||| \cdot \|x\|$$

Pela definição, vale $|f(x)| \leq p(x) \quad \forall x \in V$

Por Hahn - Banach, f pode ser estendido a $\hat{f} : E \to \mathbb{R}$ tal que $\hat{f}(x) \leq \rho(x) \quad \forall\, x \in E$. Note que isso implica $-\hat{f}(x) = \hat{f}(-x) \leq \rho(-x) = |||f|||\ \|x\| \quad \forall x \in E$, o que garante

$$|\hat{f}(x)| \leq |||f|||\,.\,\|x\| \qquad \forall\, x \in E,$$

ou seja, $|||\hat{f}||| \leq |||f|||$. A desigualdade contrária segue do fato que $\hat{f}$ é extensão de f. $\square$

<u>Observação</u>: O corolário 1 discorre apenas sobre funcionais em um espaço vetorial real, mas a prova pode ser estendida para funcionais complexos em espaços vetoriais complexos da seguinte forma: se $f : V \to \mathbb{C}$ é funcional linear em um subespaço normado complexo, seja $g(x) = \operatorname{Re} f(x)$, e veja $V \subset E$ como um espaço vetorial real, da maneira natural. Estenda g a $\hat{g} : E \to \mathbb{R}$ e recupere $\tilde{f}(x) := \tilde{g}(x) - i\tilde{g}(ix)$.

Como Hahn-Banach nos dá informação sobre o espaço dos funcionais lineares limitados em um espaço vetorial normado, é natural definir:

Definição: O dual E^* de um espaço vetorial normado E é o conjunto dos funcionais lineares limitados $f: E \to K$.

É fácil ver que E^* é, de fato, um espaço vetorial e o fato de que $|(f+g)(x)| = |f(x)+g(x)| \leq (|||f|||+|||g|||)\|x\|$ faz ver que $|||\cdot|||$ satisfaça a desigualdade triangular. Como $|||\alpha f||| = |\alpha|\,|||f|||$ $\forall \alpha \in K$, $\forall f \in E^*$, e $|||f||| = 0$ se, e só se $|f(x)| = 0$ $\forall x \in E$ faz com que $(E^*, |||\cdot|||_{E^*})$ seja um espaço normado. Hahn-Banach nos dá, portanto, uma injeção $(V^*, |||\cdot|||_{V^*}) \to (E^*, |||\cdot|||_{E^*})$ que preserva as respectivas normas. Veremos, como mais uma aplicação do Teorema, que o dual de um espaço é bastante rico:

Corolário 2: O dual E^* de E separa pontos, isto é, se $f(x_0) = 0 \ \forall f \in E^*$, então $x_0 = 0$.

Prova: Seja $V = \text{span}\{x_0\}$ e defina

$$g : V \to \mathbb{R}$$
$$g(t x_0) = t \|x_0\|$$

É imediato que g é funcional linear e que é limitado por $p : E \to \mathbb{R}$ (ou $\mathbb{C}$)
$x \mapsto \|x\|$.

Como $|g(x_0)| = \|x_0\|$ e $|g(t x_0)| \leq \|t x_0\| \ \forall t \in \mathbb{R}$, temos que $|||g||| = \|x_0\|$. Estende g para $\tilde{g} : E \to \mathbb{R}$ (ou $\mathbb{C}$), de acordo com o corolário 1. Por hipótese, vale

$$0 = \tilde{g}(x_0) = g(x_0) = \|x_0\|,$$ o que implica $\|x_0\| = 0$ $\square$

Uma observação importante diz respeito à unicidade da extensão. Defina

$$\ell^1(\mathbb{N}) = \{ \{x_n\}_{n\in\mathbb{N}} \; ; \; \sum_n |x_n| < +\infty \},$$

com a norma $\|\{x_n\}_n\|_{\ell^1} = \sum_n |x_n|$.

Seja $V = \{ \{x_n\} \in \ell^1(\mathbb{N}) \; ; \; x_1 = 0 \}$, isto é, as sequências em $\ell^1(\mathbb{N})$ cuja primeira coordenada é zero. Seja $f : V \to \mathbb{C}$
$$f(\{x_n\}) = x_2 .$$

Note que $f(0,1,0,0,0\ldots) = 1$ e que $|f(x)| \leq \|x\|$, de forma que $\|| f \|| = 1$. Defina

$$\tilde{f}_1 : \ell^1(\mathbb{N}) \to \mathbb{C} \qquad \text{e} \qquad \tilde{f}_2 : \ell^1(\mathbb{N}) \to \mathbb{C}$$
$$\tilde{f}_1(\{x_n\}) = x_2 \qquad\qquad \tilde{f}_2 = x_1 + x_2$$

Temos que ambos $\tilde{f}_1$ e $\tilde{f}_2$ estendem f, mantendo a mesma norma, mostrando que há mais de uma extensão.
De fato, se $\lambda \in [0,1]$, $\tilde{f}_\lambda := (1-\lambda)\tilde{f}_1 + \lambda \tilde{f}_2$

também é extensão de norma unitária, mostrando que, de fato, há infinitas delas nesse caso. De fato, a escolha de λ na prova do Teorema pode não ser única, e no exemplo dado a unicidade falha pela ausência de reflexividade, ou de alguma condição extra sobre a norma sobre a bola unitária.

Como comentário final, vemos que o teorema de Hahn–Banach não exige que o espaço seja normado. De fato, há versões para espaços vetoriais topológicos localmente convexos, nos quais o conceito de norma pode não estar bem-definido, mas que ainda se possa adaptar os corolários para a linguagem desses espaços. Os esforços de pesquisa atuais vão na direção de provar Hahn-Banach em outros corpos (como corpos não-arquimedianos) e para funções lineares $f: E \to K^N$, por exemplo.

Teoria Clássica da Equação do Calor e Aplicações

Introdução

A equação do Calor (também conhecida como Equação da Difusão) é um modelo para a propagação do calor em um meio, por meio da condução. Também pode ser usada para modelar difusão de massa, e é um rico modelo matemático de Equações Diferenciais Parciais (EDPs). Apesar de clássica, serve de inspiração até mesmo na atualidade, como na prova da classificação das variedades compactas em dimensão dois dada por Perelman, que utilizou uma espécie de "Difusão de curvatura", em algumas partes similar ao modelo da Equação do Calor. Vejamos sua definição e alguns resultados:

A Equação do Calor.

Seja $\Omega \subset \mathbb{R}^n$ um aberto convexo. A Equação do Calor é dada por, se $T>0$:

$$\begin{cases} \partial_t u = \Delta u & \text{em } \Omega \times (0,T) \\ u(x,0) = f(x) & \text{em } \Omega \end{cases}$$

Se f é bem-comportada, como por exemplo se pertencer à Classe de Schwartz das funções suaves que decaem super-polinomialmente, podemos utilizar a Transformada de Fourier para escrever, em $\Omega = \mathbb{R}^n$:

$$\hat{u}(\xi, t) = e^{-4\pi^2|\xi|^2 t}\hat{f}(\xi)$$

Nesse caso, temos $u(x,t) = K_t * f(x)$, em que

$$K_t(x) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}}$$

é o chamado núcleo do calor. Note que $K_t \in C^\infty(\mathbb{R}^n)$, decai exponencialmente se $|x| \to +\infty$ e vale

$$\int_{\mathbb{R}^n} K_t \, dx = \hat{K}_t(0) = 1. \text{ Além disso:}$$

$$K_t(x) = \frac{1}{t^{n/2}} K_1\left(\frac{x}{\sqrt{t}}\right).$$

Mostraremos que, mesmo que f não pertença à classe de Schwartz, ainda podemos definir uma solução usando K_t.

Lema: Seja $f \in L^p(\mathbb{R}^n)$. Então, para $1 \leq p < +\infty$,

$$K_t * f \xrightarrow{L^p} f, \quad \text{se } t \to 0^+.$$

Prova. Observe:

$$\| K_t * f(x) - f(x) \|_{L^p_x} \leq \int_{\mathbb{R}^n} \| f(x-y) - f(x) \|_{L^p_x} K_t(y)\, dy$$

$$= \int_{\mathbb{R}^n} \| f(x - \sqrt{t}y) - f(x) \|_{L^p_x} K_1(y)\, dy,$$

em que trocamos a variável na última igualdade. Como

$$\| f(x - \sqrt{t}y) - f(x) \|_{L^p_x} \leq 2 \|f\|_{L^p} \quad \text{e}$$

$$\| f(x - \sqrt{t}y) - f(x) \|_{L^p_x} \to 0 \quad \text{se } t \to 0^+,$$

o resultado segue por Convergência Dominada. $\square$

Veremos também que, se f for apenas contínua e limitada, também podemos obter uma solução

Lema 2: Se f é contínua e limitada em $\mathbb{R}^n$, então $u(x,t) = K_t * f(x)$ é contínua em $\Omega \times [0, T)$.

Prova: Seja $x_0 \in \mathbb{R}^n$ e $r_0 \geq 0$. Defina

$$V = \{x \in \mathbb{R}^n, t \geq 0 \ ; \ \|x - x_0\| \leq r \text{ e } |t| \leq r\}.$$

Como $K_1 \in L^1(\mathbb{R}^n)$, dado $\delta > 0$, seja $C_\delta \subset \mathbb{R}^n$ compacto tal que $\int_{\mathbb{R}^n \setminus C_\delta} K_1(y)\, dy < \delta$.

Daí:

$$\sup_{\substack{x \in V \\ y \in C_\delta}} |f(x - \sqrt{t}\, y) - f(x)| \longrightarrow 0, \text{ se } t \to 0^+.$$

Logo, se $x \in V$,

$$|u(x,t) - f(x)| \leq \int_{\mathbb{R}^n} |f(x - \sqrt{t}\, y) - f(x)|\, K_1(y)\, dy$$

$$\leq \sup_{\substack{x \in V \\ y \in C_\delta}} |f(x - \sqrt{t}\, y) - f(x)| \cdot + 2\|f\|_\infty \cdot \delta,$$

o que termina a prova □

Observação: se f não for limitada, mas satisfizer $|f(x)| \leq M e^{a|x|^2}$ para $M, a > 0$, a prova do lema 2 mostra que $K_t * f$ está bem-definida para $t < \frac{1}{4a}$, e satisfaz as mesmas propriedades de continuidade desse intervalo, bastando trocar $K_t(x)$ por $M e^{a|x|^2} K_t$ na prova.

Por último, observamos que $K_t * f(x)$ é suave para $t > 0$ e que $\Delta K_t = \partial_t K_t$. Obtemos, portanto, uma solução da equação do calor nos casos acima pela fórmula

$$u(x,t) = K_t * f(x) .$$

Discutiremos agora a unicidade. Seja

$$g(t) = \begin{cases} e^{-t^{-2}/2}, & t > 0 \\ 0, & t \leq 0 . \end{cases}$$

Por indução, vale que $|g^{(k)}(t)| \leq \frac{k!}{(\theta t)^k} e^{-t^{-2}/2}$, para algum $\theta > 0$

Defina, para $x \in \mathbb{R}$,

$$v(x,t) = \sum_{k=0}^{\infty} \frac{g^{(k)}(t)\, x^{2k}}{(2k)!}.$$

Note que

$$|v(x,t)| \leq \sum_{k=0}^{\infty} \frac{|g^{(k)}(t)|\, |x|^{2k}}{(2k)!} \leq \sum_{k=0}^{\infty} \frac{1}{k!} \frac{|x|^{2k}}{(\theta t)^k} e^{-t^{-2}/2}$$

$$\leq e^{-\frac{1}{t}\left(\frac{x^2}{\theta} - \frac{1}{2t}\right)},$$

com desigualdades análogas para as derivadas de g,

de forma que v é suave e satisfaz

$$\lim_{t \to 0} v(x,t) = 0.$$

Além disso, $\partial_t v = \partial_{xx} v$, o que mostra que, sem condições extras sobre a solução, a unicidade pode não ocorrer. Veremos agora, no entanto, que em domínios limitados a unicidade ocorre.

Lema 3: Princípio do Máximo em domínios limitados:

Seja $\Omega_T = \Omega \times (0,T)$ e $\partial_p \Omega_T = \Omega \times \{0\} \cup \partial\Omega \times [0,T]$.

Se $u \in C^{1,2}(\Omega_T) \cap C(\bar{\Omega}_T)$ então; para Ω aberto limitado:

(i) Se $\partial_t u - \Delta u \leq 0$, vale

$$\max_{\bar{\Omega}_T} u \leq \max_{\partial_p \Omega_T} u$$

(ii) Se $\partial_t u - \Delta u \geq 0$, vale

$$\min_{\bar{\Omega}_T} u \geq \min_{\partial_p \Omega_T} u$$

Prova: (i): Se o máximo de u em $\bar{\Omega}_T$ é atingido em $\bar{\Omega}_T \setminus \partial_p \Omega_T$, em (x_0, t_0), vale, pelas condições de primeira e segunda ordem para máximo:

$$\Delta u(x_0,t_0) \leq 0 \quad \text{e} \quad \partial_t u(x_0,t_0) \geq 0.$$

Logo, $\partial_t u(x_0,t_0) - \Delta u(x_0,t_0) \geq 0$.

Desta forma, se $\partial_t u - \Delta u < 0$, deve valer necessariamente que o máximo é atingido em $\partial_p \Omega_T$. Defina agora

$$v_K(x,t) := u(x,t) - Kt.$$

Temos $\partial_t v_K - \Delta v_K = \partial_t u - \Delta u - K < 0$.

Logo:

$$\max_{\bar\Omega_T} v_K = \max_{\partial_p \Omega_T} v_K.$$

Assim:

$$\max_{\bar\Omega_T} u = \max_{\bar\Omega_T} v_K + KT \leq \max_{\partial_p\Omega_T} u + KT$$

Como $v_K - u$ converge a zero uniformemente se $K \to 0^+$, obtemos o item (i). O item (ii) é provado substituindo u por $-u$ em (i).

$\square$

Como corolário, obtemos a unicidade da solução em domínios limitados:

Corolário: O PVI

$$\begin{cases} \partial_t u - \Delta u = g & \text{em } \Omega_T \\ u(x,0) = f & \text{em } \Omega \end{cases}$$

possui no máximo uma solução, se Ω é limitado.

Prova: se u e v são soluções, vale, para $w := u - v$:

$$\begin{cases} \partial_t w = \Delta w & \text{em } \Omega_T \\ w(x,0) = 0 & \text{em } \Omega, \end{cases}$$

o que implica $w \equiv 0$. $\square$

Um resultado interessante, no entanto, diz respeito à unicidade passada:

Proposição: Se $\partial_t u - \Delta u = 0$ em $\mathbb{R}^n \times [0,T)$ e existe $t_0 \in (0,T)$ com $u(x,t_0) \equiv 0$, então $u \equiv 0$, se $u \in L^2(\mathbb{R}^n)$ $\forall t \in [0,T)$ e $|\nabla u| \in L^2(\mathbb{R}^n)$, $|\Delta u| \in L^2(\mathbb{R}^n)$.

Prova

Defina a energia $E(t) = \frac{1}{2}\int |u|^2\, dx$. Temos

$$E'(t) = \int u\,\Delta u = -\int |\nabla u|^2$$

$$E''(t) = 2\int (\Delta u)^2$$

Note que $|E'(t)| \leq \sqrt{E(t)}\,\sqrt{E''(t)}$, por Cauchy-Schwarz.

Definindo $e(t) = \ln E(t)$, vemos que $e(t)$ é convexa e assim, se t_1 e t_2 são tais que $E(t_1)$, $E(t_2) \neq 0$, vale

$$E(t) \leq E(t_1)^{\frac{t_2 - t}{t_2 - t_1}}\, E(t_2)^{\frac{t - t_1}{t_2 - t_1}}$$

$\square$

Por fim, provaremos a unicidade impondo certas condições sobre o crescimento da solução

Proposição: A solução de

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

é única se existem $a, M > 0$ tais que $|u(x,t)| \leq Me^{a|x|^2}$.

Prova: Seja $y \in \mathbb{R}^n$ e $\mu > 0$. Defina, para $0 < \varepsilon < T-t$:

$$v_{\mu,y}(x,t) = u(x,t) - \frac{\mu}{[4\pi(T-t-\varepsilon)]^{n/2}} e^{\frac{|x-y|^2}{4(T-t-\varepsilon)}}$$

Note que $\partial_t v_{\mu,y} = \Delta v_{\mu,y}$. Defina também, para $\rho > 0$, $\hat{B}_T = B_\rho(y) \times [0,T]$. Vale em $\hat{B}_T$:

$$\max_{\overline{B_\rho(y)} \times [0,T]} v_{\mu,y} \leq \max_{\partial_p B_{\rho,T}} v_{\mu,y}.$$

Em $B_\rho \times \{0\}$, vale

$$v(x,0) \leq u(x,0) \leq \sup_{\mathbb{R}^n} |f|$$

em $\partial B_\rho(y) \times [0, T]$:

$$v_{\mu,y}(x,t) \leq u(x,t) - \frac{\mu}{(4\pi T)^{n/2}} e^{+\rho^2/4(T-t-\varepsilon)}$$

$$\leq M e^{a(|x|^2+|y|^2)} - \frac{\mu}{(4\pi T)^{n/2}} e^{\rho^2/4(T-\varepsilon)}$$

$$\leq e^{a\rho^2}\left(M e^{|y|^2}\right) - \frac{\mu}{(4\pi T)^{n/2}}\left(e^{\frac{\rho^2}{4(T-\varepsilon)}}\right)$$

ESCOLHENDO $\varepsilon < T - \frac{1}{4a}$ e ρ grande o suficiente,

temos

$$\sup_{\partial B_\rho(y)} v_{\mu,y}(x,t) \leq \sup_{\mathbb{R}^n} |f| - \varepsilon$$

Por fim, como $v_{\mu,y}(y,t) \leq u(y,t) - \frac{\mu e^0}{4\pi(T-t-\varepsilon)}$,

obtemos

$$\sup_{\mathbb{R}^n \times [0,T)} |u| \leq \sup_{\mathbb{R}^n} |f|.$$

Logo, se u e v são soluções com crescimento no máximo exponencial, $w := u - v \equiv 0$ $\square$

Forma Canônica de Jordan e Aplicações

A forma canônica de Jordan é uma espécie de generalização do conceito de diagonalização, que pode ser realizada em qualquer matriz quadrada. Para apresentá-la, precisamos de algumas definições.

Def: Em um espaço vetorial de dimensão finita E, dizemos que $\lambda \in \mathbb{C}$ é um autovalor de $T: E \to E$ linear se existe $v \neq 0$ em E tal que

$$Tv = \lambda v.$$

O vetor v é chamado de autovetor associado a λ. Note que não há unicidade para a escolha de v.

Até o fim do texto, trabalharemos apenas em espaços vetoriais de dimensão finita

Def: O λ - autoespaço (generalizado) associado ao autovalor λ é o conjunto dos $v \in E$ tais que existe $j > 0$ com $(T - \lambda I)^j v = 0$.

Observação: Os autovalores de T são sempre um conjunto não-vazio, pelo teorema fundamental da Álgebra para polinômios, pois λ é autovalor se, e só se, satisfaz $\det(T - \lambda I) = 0$. Além disso, se λ é autovalor, o λ - autoespaço generalizado é sempre não-trivial.

Def: $v \neq 0$ é um autovetor generalizado de T se existe $j \geq 1$ tal que $(T - \lambda I)^j v = 0$.

Observação: Se $K_j = \ker((T - \lambda I)^j)$, vale que $K_j \subset K_{j+1}$ $\forall j$. Logo, como a dimensão do espaço é finita, para cada autovalor λ de T, existe $j_\lambda \in \mathbb{N}$ tal que $K_{j_\lambda+1} = K_{j_\lambda}$, e assim $K_m = K_{j_\lambda}$ $\forall\, m \geq j_\lambda$. Chamamos j_λ de o índice de λ.

Proposição (1) Seja λ autovalor de T e V_λ o seu autoespaço generalizado. Se o índice de λ é j_λ, existem $\{w_i\}_{i=1}^{j_\lambda} \neq 0$ tais que $w_{i+1} = (T - I\lambda) w_i$ e o conjunto dos w_i é linearmente independente.

Prova: Seja $v \neq 0$ tal que $(T-\lambda I)^{j_\lambda} v = 0$ e $(T-\lambda I)^{j_\lambda - 1} v \neq 0$. Defina $w_1 = v$ e $w_{i+1} = (T-\lambda I) w_i$. Suponha $\sum_{i=a}^{b} c_i w_i = 0$, com $1 \leq a < b \leq j_\lambda$. Aplicando $(T-\lambda I)^{j_\lambda - a+1}$, obtemos a independência linear. □

Dessa forma, vemos que um procedimento iterativo pode gerar autovetores generalizados, tal que

$$T w_i = \lambda w_i + w_{i+1}$$

o que diz que, se tomarmos w_i e w_{i+1} como parte da base, a matriz associada terá, em parte

$$[T] = \begin{array}{c} \\ i \\ \\ \end{array} \begin{bmatrix} & \vdots & \\ & \lambda & \\ & 1 & \\ & \vdots & \end{bmatrix}$$

Dessa forma, definimos os blocos de Jordan associados a λ como (sub) matrizes quadradas J_λ tais que $(J_\lambda)_{i,i} = \lambda$ e $(J_\lambda)_{i+1,i} = 0$ ou 1. O fato de 0 poder aparecer abaixo da diagonal em vez de 1 está relacionado à dimensão de $(K_{i+1})/(K_i)$ ser ou não ser 1.

Com essa definição em mãos, podemos enunciar:

Teorema (Forma canônica de Jordan): Todo operador linear $T: E \to E$ admite uma base na qual sua representação matricial é

$$[T] = \begin{bmatrix} J_{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & J_{\lambda_m} \end{bmatrix}$$

Longue os λ_i são reais autovalores e os J_{λ_i} são blocos de Jordan.

Prova: É imediato ver que os autoespaços generalizados associados a cada autovalor são invariantes por T, e que, por contagem de dimensão e pelo fato de que vetores não-nulos em autoespaços generalizados diferentes são L.I., que $E = \bigoplus_{\lambda_j} V_{\lambda_j}$. Podemos, portanto, nos restringir a cada autoespaço generalizado V_λ. ~~Ab~~ ~~autoespaço~~ ~~...~~ ~~generalizado~~. Basta agora, de forma indutiva, tomar uma sequência finita de $\{w_i\}$ como na proposição ①. Se os $\{w_i\}$ não exaurirem V_λ, podemos reaplicar a proposição ① a ~~...~~ $V_\lambda/\{w_i\}$, restringimos o operador a esse espaço e obtendo uma nova sequência de w_i. Esse processo para em um número finito de etapas gerando um bloco de Jordan para o autovalor em questão, finalizando a prova.

$\square$

Encerraremos com algumas aplicações:

~~[illegible]~~

Corolário: Se existe $k \in \mathbb{N}$ tal que $T^k = I$, $k \neq 0$

então T é diagonalizável.

Prova: Considere o polinômio $p(t) = (t^k - 1)$

Se λ é autovalor de T, vale $\lambda^k = 1$. Fatorando

$$p(t) = t^k - 1 = (t - \lambda)\left(t^{k-1} + \lambda t^{k-2} + \dots + \lambda^{k-1}\right),$$

Temos:

$$0 = T^k - I = (T - \lambda I)\left(T^{k-1} + \lambda T^{k-2} + \dots + \lambda^{k-1} I\right)$$

Se v é autovetor ~~associado~~ generalizado associado a λ,

e $(T - \lambda I)v \neq 0$,

mas $(T-\lambda I)^2 v = 0$, defina $w = (T-\lambda I)v$.

Lema

$$0 = (T^{k-1} + \lambda T^{k-2} + \cdots + \lambda^{k-1} I) w$$

$$= k \lambda^k w .$$

Como $\lambda \neq 0 \Rightarrow$ $w = 0$

Logo, qualquer autovetor generalizado é autovetor, e a forma canônica de Jordan de T não tem elementos não-nulos fora da diagonal principal

$\square$

Outra aplicação interessante da forma canônica de Jordan é na resolução de sistemas de EDOs.

O sistema $\begin{cases} X' = AX, \\ X(0) = X_0 \end{cases}$ com A matriz quadrada tem a solução única

$$X(t) = e^{tA} x_0 \quad ,$$

em que $e^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}$, que é

absolutamente convergente, portanto convergente no espaço de Banach dos operadores limitados de $\mathbb{R}^n$ em $\mathbb{R}^n$.

É imediato ver que, se $A = PBP^{-1}$, $e^{tA} = Pe^{tB}P^{-1}$)
então podemos assumir que A está em uma forma de Jordan. Além disso, escrevendo $A = D + N$, em que D é a diagonal de A, vemos (por indução) que N é nilpotente, i.e., $N^k = 0$ para algum $k \in \mathbb{N}^*$ e que $DN = ND$. Isso implica que (por indução)

$$e^{tA} = e^{tD} e^{tN} = e^{tD} \cdot \left(\sum_{i=0}^{k} (tN)^i \right) \quad ,$$

com a última soma sendo finita, pois N é nilpotente. Além disso, se $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$,

$$e^{tD} = \begin{pmatrix} e^{t\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{t\lambda_n} \end{pmatrix},$$ e assim temos

explicitamente a solução de $\begin{cases} X' = AX \\ X(0) = 0 \end{cases}$

Mais ainda, essa representação implica na classificação topológica de sistemas lineares de EDO's a coeficientes constantes, bastando para tal olhar apenas para as partes real e imaginária dos autovalores.

Como exemplo um pouco mais concreto, assuma que a parte real de todos os autovalores de A seja negativa.

Então

$$\|X(t)\| = \|e^{tA}x_0\| \leq e^{-ct} \to 0, \text{ se } t \to \infty,$$

com $c < \max_i \{\mathrm{Re}(\lambda_i)\}$

Por outro lado, se a parte real de algum autovalor é positiva, existe x_0^* (autovetor associado ao autovalor λ_0^* com parte real positiva) tal que

$$\|e^{tA} x_0^*\| = e^{\mathrm{Re}(\lambda_0^*)t} \|x_0\| \to +\infty \quad \text{se } t \to +\infty.$$

FIM