Pontifícia Universidade Católica
do Rio de Janeiro
Rua Marquês de São Vicente, 225
Casa B
22453 - Rio de Janeiro - RJ

ELEMENTOS DE ANÁLISE FUNCIONAL APLICADA

Tatuo Nakanishi
Vitoriano Ruas

INF/003 JUN/79

DIVISÃO DE INTERCÂMBIO E EDIÇÕES

AGRADECIMENTOS

Desejamos externar os nossos profundos agradecimentos a todos aqueles que direta ou indiretamente contribuiram na realização deste trabalho e, de forma especial, ao Prof. Carlos José Pereira de Lucena, Diretor do Departamento de Informática da PUC/RJ por nos haver fornecido os meios de sua divulgação; às Srtas. Vanda Regina Lagame da Silva e Antonia Oka pelo paciente trabalho de datilografia; ao CNPq pelo apoio financeiro consentido e às nossas esposas pelo contínuo encorajamento e compreensão.

PREFÁCIO

Cada vez mais a Análise Funcional vem se impondo como ferramenta básica para o tratamento apropriado de problemas oriundos de diversas áreas que designaríamos como aplicadas tais como a Engenharia, a Pesquisa Operacional, a Estatística e a Informática.

Parece-nos entretanto que, em grande parte, os especialistas nessas matérias carecem de uma formação regular nesse domínio, o que a nosso ver representa um obstáculo para que recolham o máximo proveito de muitos dos métodos modernos e eficientes que têm a sua disposição.

Escrevemos este trabalho, portanto, com o objetivo de levar conceitos básicos de Análise Funcional de maneira tão direta e sucinta quanto possível, ao conhecimento de tantos quanto sejam usuários de técnicas nela fundamentadas, e que não tenham tido a oportunidade de passar por uma formação regular no assunto.

Aliás, gostaríamos de salientar que este texto não pretende de modo algum servir aos interesses de matemáticos, por exemplo. A omissão deliberada de inúmeros conceitos normalmente encontrados em obras clássicas de Análise Funcional, mas que seriam de pouca utilidade para aplicações mais gerais e fundamentais, tornam-no extremamente incompleto para tal fim.

Cremos, no entanto, que a apresentação dos assuntos com a preocupação de exemplificar, além dos itens de relativa simplicidade deixados à reflexão do leitor, deverá dar alguma contribuição à realização dos objetivos daqueles que vêm a compreensão e o manuseio de tais noções apenas como etapa para se lançarem sobre bases mais sólidas à abordagem de problemas em suas áreas respectivas.

É nesse sentido portanto, que este texto não deverá ser encarado como um fim mas tão somente como um preâmbulo a outro assunto de interesse - mesmo da própria Análise Funcional - para o qual seria uma espécie de pré-requisito.

Os tópicos que selecionamos como fundamentais e que tratamos em caráter de interdependência em 4 capítulos são os Espaços Vetoriais, os Espaços Normados, os Espaços Completos e as Aplicações Lineares Contínuas. Para aqueles leitores que experimentem a necessidade de complementar o estudo dos mesmos, recomendamos a leitura de qualquer um dos 6 livros sobre Análise Funcional mencionados na Bibliografia. Para conceitos apresentados no Capítulo I, em particular, sugerimos a leitura de [4].

Esclarecemos, enfim, que os 3 primeiros capítulos foram em sua maior parte regidos pelo primeiro autor, com o segundo participando em alguns sub-capítulos, e o quarto capítulo por este último.

ELEMENTOS DE ANÁLISE FUNCIONAL APLICADA

por TATUO NAKANISHI
e VITORIANO RUAS

CAPITULO I - ESPAÇOS VETORIAIS

I.1 - INTRODUÇÃO E EXEMPLOS

Conjuntos podem ser definidos por propriedades, de que gozam seus elementos. Os espaços vetoriais constituem uma classe particularmente importante de tais conjuntos pois permitiram o desenvolvimento de sólidas e abrangentes teorias que rapidamente se revelaram ser ferramenta extremamente útil em grande parte das aplicações em Matemática.

A propriedade intrínseca a todos os espaços vetoriais é a linearidade que deve ser convenientemente definida com base em duas operações fundamentais:

- a adição de elementos ou simplesmente adição e
- a multiplicação de um elemento por um escalar.

A definição dessas operações bem como a de escalar obedecem certas regras pre-estipuladas sobre as quais faremos as seguintes considerações:

Inicialmente para definir a multiplicação escalar no espaço vetorial necessitamos de um corpo de escalares que usualmente são ou o conjunto $\mathbb{R}$ dos reais ou o conjunto $\mathbb{C}$ dos complexos. Neste texto entretanto, a menos que seja dito explicitamente se tratar de complexos, a denominação escalar será atribuída a números reais. Assim, quando nos referirmos a espaços vetoriais, estaremos normalmente considerando os ditos espaços vetoriais reais.

Definição 1

O espaço vetorial X é um conjunto de elementos chamados vetores, associados à operações de adição e multiplicação por escalar, que têm a seguinte propriedade:

Adição: Se $x,y \in X$ então $x+y \in X$

Multiplicação escalar: Se $x \in X$ e α é escalar então $\alpha x \in X$.

As operações de adição e multiplicação escalar devem satisfazer aos seguintes axiomas:

Sejam x,y e z dois vetores quaisquer de um espaço vetorial X e α e β dois escalares arbitrários.

(1) $x+y = y+x$ (comutatividade)

(2) $(x+y)+z = x+(y+z)$ (associatividade)

(3) $\exists\ \vec{0} \in X \ /\ x+\vec{0} = x$ (vetor nulo)

(4) $\alpha(x+y) = \alpha x + \alpha y$ (distributividade)

(5) $(\alpha\beta)x = \alpha(\beta x)$ (associatividade)

(6) $(\alpha+\beta)x = \alpha x + \beta x$ (distributividade)

(7) $0x = \vec{0}$ e $1x = x$

Por conveniência o vetor $-1x$ é escrito $-x$ e é denominado vetor simétrico de x.

Propriedades

Dos axiomas apresentados, decorrem diferentes propriedades algumas das quais muito importantes como por exemplo:

(8) $x+y = x+z \Rightarrow y=z$

(9) $\alpha x = \alpha y$ e $\alpha \neq 0 \Rightarrow x=y$

(10) $\alpha x = \beta x$ e $x \neq \vec{0} \Rightarrow \alpha=\beta$

(11) $(\alpha-\beta)x = \alpha x - \beta x$

(12) $\alpha(x-y) = \alpha x - \alpha y$

(13) $\alpha\vec{0} = \vec{0}$

O formalismo da definição de Espaços Vetoriais pode parecer desencorajador do ponto de vista prático, porém através de exemplos pode-se mostrar que espaços vetoriais são entes muito comuns em Matemática.

Exemplo 1

O conjunto dos números reais $\mathbb{R}$ é um espaço vetorial real, pois com as definições usuais de adição e multiplicação pode mos verificar todos os axiomas.

O espaço vetorial $\mathbb{R}$ recebe diferentes denominações tais como Espaço dos Números Reais de uma Dimensão, Espaço dos Reais e Reta Real.

Exemplo 2:

O espaço vetorial $\mathbb{R}^n$ é o conjunto de todos os vetores constituidos por uma seqüência de n números reais (n-uplas).

Um vetor arbitrário x de $\mathbb{R}^n$ pode ser denotado da seguinte forma :

$$x = (x_1, x_2, \ldots, x_n)$$

onde x_i é dita a i-ésima componente do vetor x.

Se $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

$y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$

e α um escalar, então por definição:

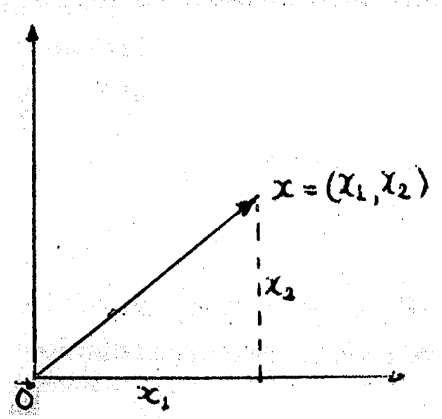
$$x + y = (x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$$

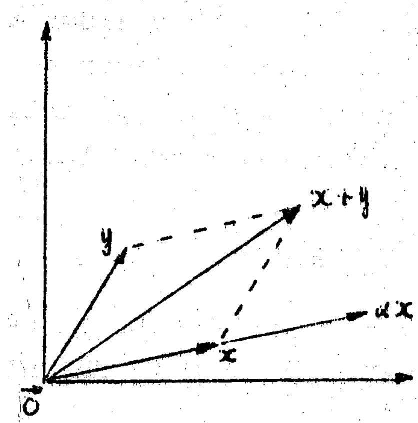
$$\alpha x = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$$

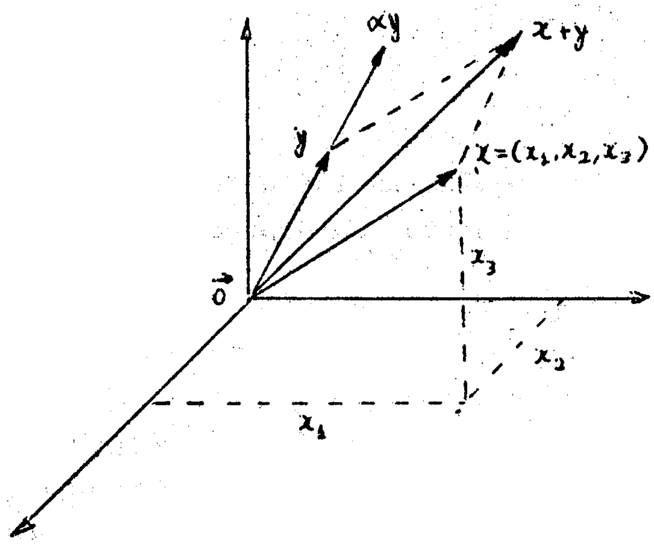
$\vec{0} = (0,0, \ldots, 0)$ é o vetor nulo.

Verifica-se facilmente que as operações de adição e multiplicação escalar assim definidas satisfazem aos axiomas (1)-(7)

O espaço vetorial $\mathbb{R}$ é um caso particular de $\mathbb{R}^n$ com n=1, e os espaços $\mathbb{R}^2$ e $\mathbb{R}^3$ são os espaços vetoriais estudados em Geometria Analítica.







Exemplo 3:

O conjunto de todas as seqüências infinitas de reais $\mathbb{R}^{\infty}$, $x \in \mathbb{R}^{\infty}$

$$x = (x_1, x_2, \ldots, x_i, \ldots)$$

é um espaço vetorial com as operações de adição e multiplicação por escalar definidas através das operações de componentes como no exemplo anterior.

Exemplo 4:

O conjunto de todas as seqüências limitadas constitui um espaço vetorial. Uma seqüência $\{x_i\}$ é uma seqüência limitada se existe uma constante M tal que $|x_i| < M$, $\forall i$.

A adição de seqüências limitadas, assim como a multiplicação escalar são definidas de forma idêntica à do exemplo 2.

Esse espaço vetorial será denotado por ℓ^{∞}.

Exemplo 5:

O conjunto de todas as seqüências infinitas de números reais, tendo somente um número finito de têrmos diferente de zero é um espaço vetorial com as operações de adição e multiplicação escalar definidas como no Exemplo 2. Seqüências diferentes do espaço podem ter diferentes números de têrmos diferentes de zero.

Exemplo 6:

O conjunto de todas as seqüências infinitas de números reais que convergem para zero constitui um espaço vetorial. Denotá-lo-emos por c_o.

Exemplos 7:

O conjunto $C^o[a,b]$ das funções contínuas com valores reais, definidas no intervalo $[a,b]$, é um espaço vetorial onde $[a,b]$ é um intervalo fechado limitado da reta real. Relembramos que uma função f é contínua num intervalo $I \subset \mathbb{R}$ se para todo $x \in I$, dado um $\varepsilon>0$, arbitrário, podemos encontrar um $\delta>0$, dependendo de ε e eventualmente de x, tal que

$$|x-y|< \delta \Rightarrow |f(x)-f(y)|<\varepsilon, \ y\in I$$

Assim se $f(x)$ e $g(x)$, $a\leq x\leq b$, são duas funções contínuas em $[a,b]$ e α um escalar, podemos definir as operações fundamentais com
$$\begin{cases} (f+g)(x) = f(x)+g(x), & \forall x \in [a,b] \\ (\alpha f)(x) = \alpha f(x) \quad, & \forall x \in [a,b] \end{cases}$$

Além disso temos

$f = g \Rightarrow f(x) = g(x)$, $\forall x \in [a,b]$ e

$\vec{0}(x)=0$, $\forall x \in [a,b]$ é a função identicamente nula.

Analogamente poderíamos definir o espaço vetorial $C^o(a,b)$ de funções contínuas no intervalo aberto limitado (a,b). Entretanto tais funções não têm, em geral, a propriedade de serem uniformemente contínuas, como as de $C^o[a,b]$, sendo a continuidade uniforme num intervalo I assim definida:

$\forall x \in I$ e $\forall \varepsilon>0$, $\exists \delta>0$, independente de x tal que $|f(x)-f(y)|< \varepsilon$ se $|x-y|< \delta$, $y\in I$.

Por exemplo, a função $f(x)=1/x$ é contínua em $(0,1)$ mas não em $[0,1]$.

Note-se que essa função não é uniformemente contínua em $(0,1)$.

Exemplo 8:

O conjunto $\mathcal{P}$ de todos os polinômios com coeficientes reais é um espaço vetorial como se pode constatar sem dificuldades. O vetor nulo, a adição e a multiplicação escalar são definidos como no Exemplo 7.

Comumente define-se também o espaço $\mathcal{P}[a,b]$ dos polinômios restritos ao intervalo limitado $[a,b]$, o mesmo valendo quanto ao vetor nulo e às operações fundamentais.

I.2 - PRODUTO CARTESIANO

Definição 2

Sejam X e Y espaços vetoriais. Chamamos de "produto Cartesiano "$X \times Y$ ao espaço vetorial constituído de todos os pares ordenados (x,y) onde $x \in X$ e $y \in Y$.

A adição e a multiplicação escalar são definidas da seguinte forma:

$$(x_1,y_1) + (x_2,y_2) = (x_1+x_2, y_1+y_2)$$

$$\alpha(x,y) = (\alpha x, \alpha y)$$

O vetor nulo de $X \times Y$ é o vetor $\vec{0} = (\vec{0}_x, \vec{0}_y)$ onde $\vec{0}_x$ e $\vec{0}_y$ são os vetores nulos de X e Y respectivamente.

A definição de produto cartesiano pode ser generalizada para n espaços vetoriais $X_1, X_2, \ldots, X_n$, como sendo o conjunto de n-tuplas ordenadas.

$$(x_1, x_2, \ldots, x_n) \text{ onde } x_i \in X_i.$$

Se $X_1 \equiv X_2 \equiv \ldots \equiv X_n$ denotamos o produto por X^n, i.e. o produto de um espaço por ele mesmo n vezes.

Exemplo 9:

$\mathbb{R}^2$ = produto cartesiano $\mathbb{R} \times \mathbb{R}$

$\mathbb{R}^n$ = produto cartesiano $\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$

I.3 - SUBESPAÇOS

Subespaço é a denominação atribuida a um subconjunto M de um espaço vetorial X, que seja também um espaço vetorial com as operações de adição e multiplicação escalar definidas em X. Por razões práticas escolheremos como definição de subespaço, o que alguns autores consideram uma propriedade, que é o critério de subespaço.

Definição 3

Seja M um subconjunto não vazio de um espaço vetorial X e x e y dois vetores quaisquer de M. Dado dois escalares quaisquer α e β, se $\alpha x + \beta y$ ainda é um vetor de M, então M é um subespaço do espaço vetorial X.

Fazendo $\beta=0$ na definição de subespaço, deduzimos que αx pertence a M para todo escalar α e todo $x \in M$.

No caso particular em que $\alpha=0$ e x um vetor qualquer de M teremos

$$x = 0x = \vec{0}.$$

Portanto todo subespaço contem necessàriamente o vetor nulo.

Qualquer espaço vetorial é um subespaço do mesmo espaço vetorial. Quando um subespaço M é diferente do correspondente espaço vetorial X, êle é chamado um subespaço próprio de X.

O subespaço constituido apenas pelo vetor nulo é chamado subespaço trivial.

Exemplo 10:

No espaço $\mathbb{R}^2$ qualquer reta que passa pela origem é um subespaço.

Exemplo 11:

O subconjunto M de $\mathbb{R}^3$, consistindo de todos os vetores cuja terceira componente é zero, é um subespaço de $\mathbb{R}^3$. Se cada vetor de $\mathbb{R}^3$ é visto como suas componentes ordinárias no espaço tridimensional (x,y,z) então M é exatamente o plano (x,y)

Exemplo 12:

O conjunto das funções contínuas que assumem o valor zero no ponto 1/2 é um subespaço do espaço $C^0[0,1]$

Exemplo 13:

Retas e planos que passam pela origem são subespaços próprios, não triviais, de $\mathbb{R}^3$.

Exemplo 14:

O conjunto $\mathcal{P}^n$ de todos os polinômios de coeficientes reais e de grau menor ou igual a n é um subespaço do espaço de todos os polinômios. Se $p \in \mathcal{P}^n$

$$p = a_0 + a_1 x + \ldots + a_n x^n$$

onde a_i são números reais

Se n é um número inteiro fixo, naturalmente, o vetor nulo de $\mathcal{P}^n$ é o polinômio identicamente nulo.

Pode-se mostrar facilmente que $\mathcal{P}^n$ é também um subespaço de $\mathcal{P}^m$ sempre que $m \geq n$.

Como no Exemplo 8, podemos definir também $\mathcal{P}^n[a,b]$ subespaço de $\mathcal{P}[a,b]$.

Proposição 1:

Se M e N são subespaços de um espaço vetorial X, então o conjunto de vetores que constituem a interseção $M \cap N$ é um subespaço de X.

Demonstração:

Sejam x e y dois vetores quaisquer da interseção $M \cap N$ e

α,β dois escalares arbitrários. $\alpha x + \beta y$ pertencerá ao subespaço M e também ao subespaço N, logo pertencerá à intersecção $M \cap N$. Portanto é um subespaço. c.q.d.

O leitor poderá verificar que uma proposição semelhante não é válida para a união $M \cup N$.

I.4 - GERAÇÃO DE SUBESPAÇOS

Definição 4:

Uma combinação linear de um conjunto de vetores $\{x_1, x_2, \ldots, x_n\}$ de um espaço vetorial X, é uma soma da forma

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$$

onde os α_i são escalares.

Pela aplicação recursiva do critério de subespaço verifica-se que a combinação linear pertence ainda ao espaço X.

De forma inversa poderemos construir um subespaço a partir de um conjunto de vetores de um espaço vetorial, através de combinações lineares desses vetores.

Definição 5:

Seja M um conjunto de vetores de um espaço vetorial X. Chamamos de subespaço gerado por M o conjunto de todas as combinações lineares dos elementos de M, que denotaremos por [M] .

Se M é um conjunto de vetores $\{x_1, x_2, \ldots, x_n, \ldots\}$ do espaço vetorial X então cada vetor do subespaço [M] é dado por

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n \qquad \text{para inteiros n e}$$

para α_i escalares.

Propriedade:

[M] é o menor subespaço que contém M, pois todos os subespaços que contém M, contém todas as combinações lineares dos vetores de M. Como [M] é o conjunto dessas combinações então ele é o menor deles.

Exemplo 15:

Se M é constituido de um único vetor x, não nulo; de $\mathbb{R}^2$ então [M] é a reta que passa pela origem determinada por x.

Exemplo 16:

Sejam x e y dois vetores, não nulos, do espaço $\mathbb{R}^2$. Se x e y são colineares eles geram um subespaço que é a reta que os contém. Senão, eles geram um subespaço que é o próprio espaço $\mathbb{R}^2$.

Qualquer vetor de $\mathbb{R}^2$ neste caso, pode ser obtido por meio de uma combinação linear de x e y.

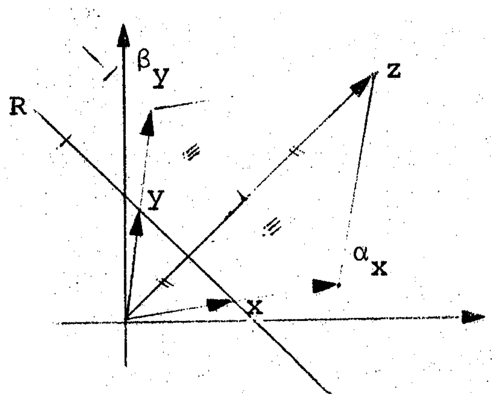
Exemplo 17:

No espaço $\mathbb{R}^3$ o subespaço gerado por uma reta que não passa pela origem é um plano que passa pela origem e contém a reta (verificar).

Exemplo 18:

No espaço $\mathbb{R}^2$ o subespaço gerado por uma reta R que não passa pela origem é o próprio espaço $\mathbb{R}^2$.

$z = \alpha x + \beta y$



I.5 - SOMA DE DOIS SUBESPAÇOS

Definição 6:

Sejam M e N dois subespaços de um espaço vetorial X. Chamamos de soma dos subespaços M e N o conjunto de todos os vetores da forma $x = m+n$ onde $m \in M$ e $n \in N$, que denotaremos por $M+N$.

Proposição 2:

A soma de dois subespaços M e N de um espaço vetorial X é um subespaço.

Demonstração:

Se x e y são vetores de $M + N$, então existem os vetores $m_1, m_2 \in M$ e $n_1, n_2 \in N$ tais que:

$$x = m_1 + n_1$$
$$y = m_2 + n_2$$

Logo:

$$\alpha x + \beta y = \alpha m_1 + \alpha n_1 + \beta m_2 + \beta n_2 = \alpha m_1 + \beta m_2 + \alpha n_1 + \beta n_2$$

Como M e N são subespaços

$m = \alpha m_1 + \alpha m_2 \in M$

$n = \alpha n_1 + \beta n_2 \in N$

Logo

$x + y = m + n \in M + N$ e $M+N$ é um espaço vetorial

Analogamente, $\alpha x = \alpha m + \alpha n$ e $\alpha x \in M+N$ $\forall \alpha \in \mathbb{R}$ c.q.d.

Exemplo 19:

De acordo com a Proposição 2 podemos dizer que o espaço $C^0[a,b]$ das funções contínuas é a soma de dois espaços vetoriais $C^+[a,b]$ e $C^-[a,b]$, onde:

$C^+[a,b]$ é o conjunto das funções pares $(f(x)=f(-x), \ \forall x \in [a,b])$

$C^-[a,b]$ é o conjunto das funções ímpares $(f(x) = -f(-x), \ \forall x \in [a,b])$

De fato se $f(x) \in C^0[a,b]$ é possível escrever

$$f(x) = f_1(x) + f_2(x)$$

onde

$f_1(x) \in C^+[a,b]$ e

$f_2(x) \in C^-[a,b]$, senão vejamos

Definimos $f_1(x) = \dfrac{f(x) + f(-x)}{2}$ e constatamos que $f_1(x)$ é par pois

$$f_1(-x) = \frac{f(-x) + f(-(-x))}{2} = f_1(x)$$

Agora definimos

$$f_2(x) = \frac{f(x) - f(-x)}{2}$$

e constatamos que $f_2(x)$ é ímpar:

$$-f_2(-x) = -\frac{f(-x) - f(-(-x))}{2} = f_2(x)$$

Como

$f(x) = f_1(x) + f_2(x)$, temos o resultado desejado.

Da própria construção acima, se poderá verificar que tal representação de elementos de $C^0[a,b]$ como soma de funções de $C^+[a,b]$ e $C^-[a,b]$ é única. No entanto, podemos considerar $C^0[a,b]$ igualmente como soma dos subespaços $C_E[a,b]$ e $C_D[a,b]$ onde:

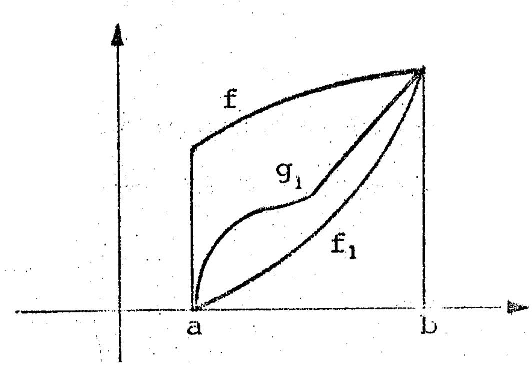
$$C_E[a,b] = \{f/f \in C^0[a,b] \ \text{ e } \ f(a) = 0\}$$

$$C_D[a,b] = \{f/f \in C^0[a,b] \ \text{ e } \ f(b) = 0\}$$

Neste caso podemos encontrar uma infinidade de decomposições possíveis $f(x) = f_1(x) + f_2(x)$ com $f_1 \in C_E[a,b]$ e $f_2 \in C_D[a,b]$ como ilustrado abaixo:

$f = f_1 + f_2$ com $f_2 = f - f_1$

$f = g_1 + g_2$ com $g_2 = f - g_1$



I.6 - VARIEDADE LINEAR

Definição 7

Seja M um subespaço de um espaço Vetorial X e x_o um vetor de X, $x_o \neq \vec{0}$. O conjunto de vetores

$$V = \{v / v = x_o + m, \text{ onde } m \in M\}$$

é denominado uma variedade linear de X.

Uma variedade linear V pode ser escrita sob a forma

$$V = x_o + M$$

Podemos interpretar uma variedade linear como a translação de um subespaço M de "módulo" x_o.

Uma propriedade fundamental das variedades lineares é que:

Se $x, y \in V$ então $x - y \in M$

De fato, temos

$$x \in V \quad \exists\, m_1 \in M \quad \text{tal que} \quad x = x_o + m_1$$

$$y \in V \quad \exists\, m_2 \in M \quad \text{tal que} \quad y = x_o + m_2$$

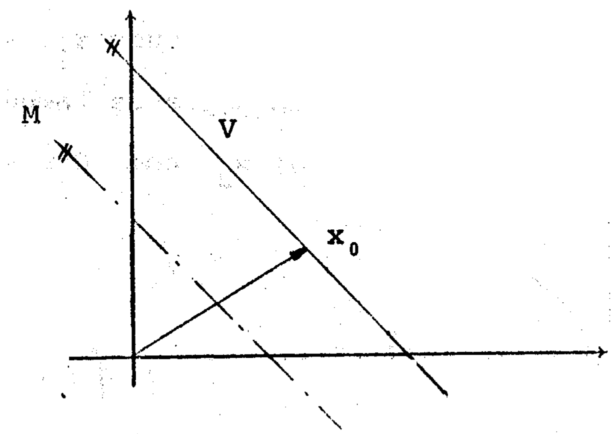
Logo:

$$x - y = m_1 - m_2 \in M$$

Dada uma variedade linear, qualquer um de seus elementos pode servir de módulo x_o.

Exemplo 20

No espaço $\mathbb{R}^2$, uma reta que não passa pela origem é uma variedade linear. O subespaço M é a reta paralela que passa pela origem.



Exemplo 21

Em $\mathbb{R}^3$, um plano que não contém a origem é uma variedade linear

Exemplo 22

Consideremos o subespaço de $C^0[\pi/6, \pi/4]$ gerado por $S = \{\text{sen}x, \cos x\}$. $V = S + f$, sendo f uma função de $C^0[\pi/6, \pi/4]$, $f \notin S$, é uma variedade linear de módulo $x_0 = f$

Se tomarmos o vetor $x_0 = f + \text{sen}x$, $f \in V$, obteremos a mesma variedade linear.

Exemplo 23

Considerando agora o subespaço de $C^0[0,1]$ de funções tais que $f(0) = 0$, o conjunto de funções contínuas cujo valor para $x=0$ é constante igual a c é uma variedade linear de $C^0[0,1]$. Podemos tomar $x_0 = x+c$, por exemplo.

I. 7 - CONVEXIDADE E CONES

Definição 8

Um conjunto S de um espaço vetorial é dito ser convexo se, dado dois pontos quaisquer $x_1, x_2 \in S$, o segmento de reta por eles determinado pertence também ao conjunto S, ou seja, se $x = \alpha x_1 + (1 - \alpha) x_2$ com $0 \leq \alpha \leq 1$, então $x \in S$.

CONVEXO NÃO CONVEXO

O conjunto vazio é considerado convexo.

Subespaços e variedades lineares são convexos (verificar)

Proposição 3

Sejam S e T conjuntos convexos de um espaço vetorial. Então:

a) $\alpha S = \{ x/x = \alpha s \in S \}$ é convexo para qualquer escalar α

b) $S+T = \{x/x = s+t$ com $s \in S$ e $t \in T\}$ é convexo

Demonstração

a) Seja: $x_1 = \alpha s_1 \quad , s_1 \in S$

$x_2 = \alpha s_2 \quad , s_2 \in S \quad e \quad 0 \leq \beta \leq 1$

Temos: $\beta x_1 + (1-\beta)\, x_2 = \beta\, \alpha s_1 + (1-\beta)\, \alpha s_2$
$= \alpha\, [\beta s_1 + (1-\beta) s_2]$

Como S é convexo $\beta s_1 + (1-\beta)\, s_2 \in S$

Logo

$\beta x_1 + (1-\beta) x_2 \in \alpha S$

b) Seja $x_1, x_2 \in S + T$. Então $\exists s_1, s_2 \in S$ e $\exists\, t_1, t_2 \in T$ tais que:

$$x_1 = s_1 + t_1$$
$$x_2 = s_2 + t_2$$

Seja $0 \leq \alpha \leq 1$: $y = \alpha x_1 + (1-\alpha) x_2$
$= \alpha s_1 + \alpha t_1 + (1-\alpha) s_2 + (1-\alpha) t_2$
$= \alpha s_1 + (1-\alpha) s_2 + \alpha t_1 + (1-\alpha) t_2$

S convexo $\Rightarrow s = \alpha s_1 + (1-\alpha) s_2 \in S$

T convexo $\Rightarrow t = \alpha t_1 + (1-\alpha) t_2 \in T$. Logo

$y = s + t$ onde $s \in S$ e $t \in T$

e $y = \alpha x_1 + (1-\alpha) x_2 \in S+T$, $\forall x_1, x_2 \in S+T$; portanto $S+T$ é convexo.

c.q.d.

Proposição 4:

Qualquer interseção de conjuntos convexos é um convexo.

Demonstração

Seja X uma coleção arbitrária de convexos e $K = \bigcap_{C \subset X} C$ a interseção de todos os convexos $C \subset X$.

Se $x_1, x_2 \in K$ então $x_1, x_2 \in C \quad \forall C \subset X$ e como C é convexo

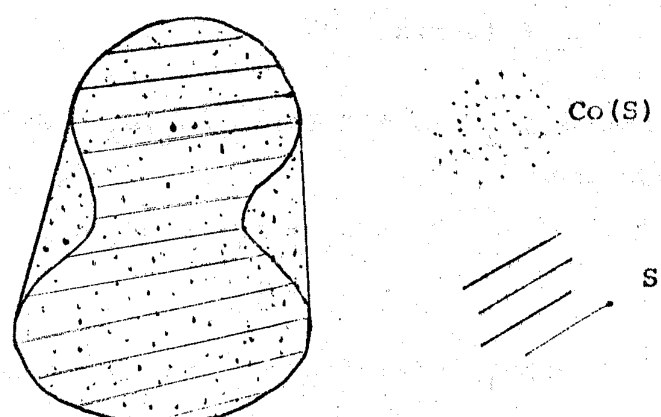
$$\alpha x_1 + (1-\alpha) x_2 \in C \qquad \forall C \subset X$$
$$0 \leq \alpha \leq 1$$

então $\alpha x_1 + (1-\alpha) x_2 \in K$ e K é convexo.

c.q.d.

Definição 9:

Seja S um conjunto arbitrário de um espaço vetorial X. O "casco" convexo de S", denotado $co(S)$, é o menor convexo que contém S. Em outras palavras, é a intersecção de todos os conjuntos convexos que contém S.



Proposição 5:

Seja S um conjunto de vetores $\{x_1, x_2, \ldots, x_n, \ldots\}$ de um espaço vetorial X. Se K é o conjunto de todas as combinações lineares da forma

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n$$

com $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 1$ e $\alpha_i \geq 0 \quad i = 1, 2, \ldots, n$ ditas combinações convexas, então $K = co(S)$

Demonstração:

Seja

$$S = \{x_1, x_2, \ldots, x_n, \ldots\} \text{ e}$$
$$K = \{y / y = \sum_{i=1}^{n} \alpha_i x_i, \quad \sum_{i=1}^{n} \alpha_i = 1, \quad \alpha_i \geq 0\}$$

Mostraremos inicialmente que K está contido na intersecção de todos os convexos que contém S e, em seguida, que K é convexo. Daí se concluirá portanto: $K= co(S)$

Seja C um convexo qualquer que contém S e K_n conjunto das combinações convexas de n vetores de S. Mostraremos por indução finita que C contém K_n $\forall n$.

Seja $n=2$ e α_1, α_2 escalares com $\alpha_1 + \alpha_2 = 1$, $\alpha_1,\alpha_2 \geq 0$.

Como C é convexo, então

$$\alpha_1 x_{i_1} + \alpha_2 x_{i_2} \in C, \quad i_1 \neq i_2, \text{ ou seja } K_2 \subset C$$

Suponhamos agora que $K_{n-1} \subset C$, isto é:

$$\alpha_1 x_{i_1} + \alpha_2 x_{i_2} + \ldots + \alpha_{n-1} x_{i_{n-1}} \in C, \quad \sum_{i=1}^{n-1} \alpha_i = 1$$

$$i_k \neq i_\ell \quad \text{se} \quad k \neq \ell$$

Provemos que $K_n \subset C$:

Seja então a combinação

$$\sum_{k=1}^{n} \alpha_k x_{i_k} \quad \text{com} \quad \sum_{k=1}^{n} \alpha_k = 1, \ \alpha_k \geq 0 \quad \forall k$$

Podemos excluir os casos $\alpha_k = 0,1$ para algum k por recairem ou no caso anterior ou no caso trivial $x_{i_k} \in C$.

Por conseguinte, definindo-se y, tal que:

$$y = \frac{\alpha_2}{1-\alpha_1} x_{i_2} + \frac{\alpha_3}{1-\alpha_1} x_{i_3} + \ldots + \frac{\alpha_n}{1-\alpha_1} x_{i_n}$$

temos: $$\sum_{i=2}^{n} \frac{\alpha_i}{1-\alpha_1} = \frac{1-\alpha_1}{1-\alpha_1} = 1$$

Logo, pela hipótese de indução $y \in C$.

Como C é convexo

$$\alpha_1 x_{i_1} + (1-\alpha_1) y \in C, \text{ ou seja}$$

$$\alpha_1 x_{i_1} + \alpha_2 x_{i_2} + \ldots + \alpha_n x_{i_n} \in C$$

Enfim, como $K = \bigcup_{n=1}^{\infty} K_n$ temos $K \subseteq C$, donde $K \subseteq Co(S)$.

Resta provar que K é convexo.

Seja $x,y \in K$, logo

$$x = \sum_{i=1}^{n} \alpha_i x_i \quad \text{e} \quad y = \sum_{i=1}^{n} \beta_i x_i$$

para certo n bem escolhido, com $\sum_{i=1}^{n} \alpha_i = 1 = \sum_{i=1}^{n} \beta_i$

Seja agora $0 \leq \gamma \leq 1$ e

$$\gamma x + (1-\gamma)y = \gamma \sum_{i=1}^{n} \alpha_i x_i + (1-\gamma) \sum_{i=1}^{n} \beta_i x_i$$

$$= \sum_{i=1}^{n} [\gamma\alpha_i + (1-\gamma)\beta_i]x_i$$

Porém

$$\sum_{i=1}^{n} [\gamma\alpha_i + (1-\gamma)\beta_i] = \gamma \sum_{i=1}^{n} \alpha_i + (1-\gamma) \sum_{i=1}^{n} \beta_i = \gamma + 1 - \gamma = 1$$

Logo $\gamma x + (1-\gamma)y \in K$ e K é convexo c.q.d.

Definição 10:

Um conjunto convexo C de um espaço vetorial é dito ser um cone convexo de vértice na origem se $x \in C$ implicar em $\alpha x \in C$ para todo $\alpha \geq 0$.

Um cone C_p com vértice p é definido como uma translação $p + C$ de um cone C com vértice na origem. Nesse caso verifica-se a condição se transforma em: se $x \in C_p$, $\alpha x + (1-\alpha)p \in C_p \quad \forall \alpha \geq 0$

Exemplo 24:

O conjunto P no espaço $\mathbb{R}^n$, tal que

$P = \{x/x = (x_1, x_2, \ldots, x_n),\ x_i \geq 0,\ \forall_i\}$ é um cone convexo.

Exemplo 25:

O conjunto de todas as funções contínuas não negativas é um cone convexo no espaço das funções contínuas.

I.8 - INDEPENDÊNCIA LINEAR

Definição 11:

Um vetor x é dito linearmente dependente de um conjunto de vetores $S = \{x_1, x_2, \ldots, x_n\}$ se ele pode ser expresso como uma combinação linear dos vetores de S, ou seja:

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n, \quad \alpha_1, \alpha_2, \ldots, \alpha_n \text{ escalares}$$

Se por outro lado não existe tal relação, diz-se que x é linearmente independente do conjunto S de vetores.

Portanto, se x é linearmente dependente de S, ele pertence ao espaço $[S]$ gerado por S. Notemos também que o vetor nulo é linearmente dependente de qualquer vetor.

De acordo com a Definição 11, dizemos que o próprio S constitui um conjunto de vetores é linearmente independentes se nenhum de seus vetores linearmente dependente dos restantes.

Teorema 1

Uma condição necessária e suficiente para que $S = \{x_1, x_2, \ldots, x_n\}$ seja um conjunto de vetores linearmente independentes é que:

$$\sum_{k=1}^{n} \alpha_k x_k = \vec{0} \Rightarrow \alpha_k = 0 \quad \forall k$$

Demonstração

Seja $\{x_1, x_2, \ldots, x_n\}$ um conjunto de vetores linearmente independentes.

$$\text{Se} \quad \sum_{k=1}^{n} \alpha_k x_k = \vec{0}$$

e existe um certo $r, 1 \leq r \leq n$ tal que $\alpha_r \neq 0$ então podemos escrever:

$$\alpha_r x_r + \sum_{\substack{k=1 \\ k \neq r}}^{n} \alpha_k x_k = \vec{0}$$

Logo:

$$x_r = \sum_{\substack{k=1 \\ k \neq r}}^{n} \left(\frac{-\alpha_k}{\alpha_r}\right) x_k$$

ou seja, x_r é linearmente dependente dos restantes o que é um absurdo.

Temos agora que provar que dado um conjunto de vetores $\{x_1, x_2, \ldots, x_n\}$, então

$$\sum_{k=1}^{n} \alpha_k x_k = \vec{0} \Rightarrow \alpha_k = 0 \quad \forall_k$$

implica que nenhum vetor x_k é linearmente dependente dos demais.

De fato, se existisse um x_r tal que $x_r = \sum_{\substack{k=1 \\ k\neq r}}^{n} \alpha_k x_k$

então existiria uma combinação linear nula dos n vetores com $\alpha_r = 1$, o que contrária a hipótese.

Logo $\{x_1, x_2, \ldots, x_n\}$ é um conjunto de vetores linearmente independentes. c.q.d.

<u>Exemplo 26</u>:

O Teorema 1 é normalmente utilizado na verificação da independência linear de vetores.

Por exemplo: $x_1 = (1,0,0)$, $x_2 = (0,1,0)$, $x_3 = (2,3,5)$ são linearmente independentes porque $\sum_{k=1}^{3} \alpha_k x_k = 0$ implica que

$$\begin{aligned} \alpha_1 + 2\alpha_3 &= 0 \\ \alpha_2 + 3\alpha_3 &= 0 \\ 5\alpha_3 &= 0 \end{aligned} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \alpha_3 = 0$$

Uma conseqüência direta do Teorema 1 é o resultado seguinte cuja veracidade se poderá verificar sem problemas:

Se $x_1, x_2, \ldots, x_n$ são linearmente independentes então:

$$\sum_{k=1}^{n} \alpha_k x_k = \sum_{k=1}^{n} \beta_k x_k \Longrightarrow \alpha_k = \beta_k \quad \forall k.$$

<u>Definição 12</u>:

Um conjunto de vetores (finito ou não) é dito <u>livre</u> se todo seu subconjunto finito for de vetores linearmente independentes.

É evidente que todo conjunto de vetores linearmente independentes é livre.

I.9 - DIMENSÃO E BASE

Consideremos um conjunto S de vetores que geram um espaço $[S]$. Se eliminarmos sucessivamente de S todos os vetores linearmente dependentes dos restantes até obtermos um conjunto S' do qual não é possível se extrair nenhum vetor linearmente dependente dos demais, S' continuará gerando o mesmo espaço $[S]$.

Nesse caso $[S]$ é gerado por um conjunto livre de vetores S' que será dito uma base de $[S]$. Em geral temos:

Definição 13:

Um conjunto livre S de vetores de um espaço vetorial X, é dito uma base de X, se $[S] = X$.

Quando o conjunto livre S tem um número de vetores finito com $[S] = X$, então X é um espaço de dimensão finita e, em caso contrário diz-se que é de dimensão infinita. No caso em que a dimensão é finita escrevemos $\dim X = n$ onde n é o número de vetores da base que é único de acordo com o

Teorema 2

Se um espaço vetorial X tem uma base contendo n vetores então todas as outras bases de X contêm n vetores.

Demonstração:

Mostraremos que se $\{x_1, x_2, \ldots, x_n\}$ e $\{y_1, y_2, \ldots, y_m\}$ são bases de um espaço vetorial X então $m = n$.

Vamos supor que $m \geq n$.

Cada vetor y_1 pode ser escrito como combinação linear de vetores da primeira base:

$$y_1 = \sum_{i=1}^{n} \alpha_i x_i$$

Como $y_1 \neq \vec{0}$ pelo menos um dos α_i é diferente de zero. Reindexando os x_i, se necessário, podemos considerar que $\alpha_1 \neq 0$ e então:

(14) $$x_1 = \alpha_1^{-1} y_1 + \sum_{i=2}^{n} \alpha_1^{-1} \alpha_i x_i$$

Substituindo x_1 na base $\{x_1, x_2, \ldots, x_n\}$ por (14), concluiremos $\{y_1, x_2, x_3, \ldots, x_n\}$ gera o espaço X.

Usemos agora indução finita. Suponhamos que tenha sido possível substituir os k-1 primeiros vetores x_i pelos k-1 primeiros y_i obtendo-se ainda um conjunto que gera X. O vetor y_k pode assim ser expresso como uma combinação linear de $y_1, y_2, \ldots, y_{k-1}, x_k, \ldots, x_n$:

$$y_k = \sum_{i=1}^{k-1} \alpha_i y_i + \sum_{i=k}^{n} \beta_i x_i$$

Como $y_1, y_2, \ldots, y_k$ são linearmente independentes, temos:

$$\sum_{i=k}^{n} \beta_i x_i \neq \vec{0}.$$

Logo nem todos os β_i podem ser nulos. Agora, reindexando-se $x_k, x_{k+1}, \ldots, x_n$, se necessário, podemos ter $\beta_k \neq 0$. Então:

$$x_k = \sum_{i=1}^{k} \beta_k^{-1} \alpha_i y_i + \sum_{i=k+1}^{n} \beta_k^{-1} \beta_i x_i$$

O novo conjunto de vetores

$\{y_1, y_2, \ldots, y_k, x_{k+1}, \ldots, x_n\}$ gera ainda o espaço X.

Logo pela indução concluimos que podemos substituir todos os n x_i pelos n primeiros y_i, formando um conjunto de n vetores y_i linearmente independentes que gera o espaço X. Então $y_1, y_2, \ldots, y_n$ é uma base de X e qualquer vetor y_{n+1} pode ser escrito como uma combinação linear desses n vetores.

Logo m deve ser igual a n.

c.q.d.

Um resultado, algo evidente, que o leitor poderá provar é o seguinte:

Se M é um subespaço do espaço vetorial X de dimensão n, então a dim$M \leq n$.

<u>Observação</u>:

O que se pode concluir também de tal propriedade é que todo subespaço M não trivial de um espaço X de dimensão n, tem uma base de dimensão $m \leq n$. Se $m=n$ então a base de M será também uma base de X e portanto $M=X$. Por outro lado se $m<n$ en-

tão M é um subespaço próprio de X (verificar).

Exemplo 27:

O conjunto de vetores $\{e_1, e_2, \ldots, e_n\}$ com:

$$\begin{aligned} e_1 &= (1,0,\ldots,0) \\ e_2 &= (0,1,\ldots,0) \\ &\vdots \\ e_n &= (0,0,\ldots,1) \end{aligned}$$

é chamado a base canônica de $\mathbb{R}^n$.

Qualquer vetor $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ pode ser escrito como a combinação linear desse conjunto de vetores, dada por

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$$

Exemplo 28

O conjunto livre $\{1, x, x^2, \ldots, x^n, \ldots\}$ é uma base do espaço $\mathcal{P}$.

Exemplo 29

$b_1 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, $b_2 = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ formam uma base de $\mathbb{R}^2$.

De fato, $\alpha_1 b_1 + \alpha_2 b_2 = \vec{0} \Longrightarrow \alpha_1 = \alpha_2 = 0$ (verificar)

Assim qualquer vetor de $\mathbb{R}^2$ tem que ser uma combinação linear desses dois vetores linearmente independentes. Nesse caso:

$$x = (x_1, x_2) = \frac{x_1 + x_2}{\sqrt{2}} b_1 + \frac{x_1 - x_2}{\sqrt{2}} b_2$$

Exemplo 30

O conjunto $S = \{e_1, e_2, \ldots, e_n, \ldots\}$ com $\begin{cases} e_i^n = 1 \text{ se } n=i \\ e_i^n = 0 \text{ se } n \neq i \end{cases}$

forma uma bse do espaço X de sequencias com um número finito de componentes não nulas.

De fato, S é livre e além disso $[S] = X$.

I.10 - ESPAÇOS QUOCIENTE

Definição 14:

Seja M um subespaço próprio de X. Podemos construir uma infinidade de variedades lineares V de X arbitrando-se diferentes $x_o \in X$, $x_o \notin M$. O conjunto dessas variedades lineares é chamado quociente de X por M e denotado X/M.

X/M é um espaço vetorial como será mostrado mais adiante.

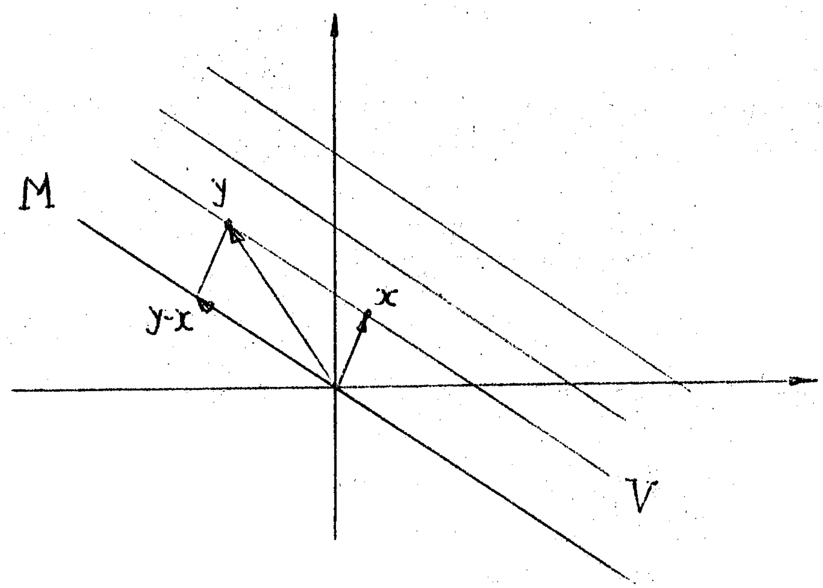
Uma variedade linear V do espaço X/M pode ser representada por qualquer um de seus vetores $x \in V$, que passa a ser denominado um representante da classe V, sendo esta denotada por $[x]$.

Deve ficar subentendido que se $x,y \in V$ (isto é, $x-y \in M$) então $[x] = [y] = V$ em X/M.

Observe-se que a condição necessária e suficiente para que se obtenha duas variedades lineares V_1 e $V_2 \in X/M$ diferentes é que os dois vetores escolhidos para gerá-las x_1 e $x_2 \in X$ sejam tais que $x_1-x_2 \notin M$.

Exemplo 31:

O conjunto de retas de $\mathbb{R}^2$ paralelas a uma reta M que passa pela origem é o espaço $\mathbb{R}^2/M$. Evidentemente podemos representar a reta V por $[x]$ ou $[y]$ se $x,y \in V$, pois $x-y \in M$



Exemplo 32:

Se $X = C^{0}[0,1]$ e se tomarmos $M = \mathcal{P}^{n}[0,1]$, cada variedade linear do espaço quociente X/M é constituida de elementos que diferem entre si por um polinômio de grau $\leq n$.

$$[\text{sen}x] = [\text{sen}x+1] = [\text{sen}x + x^{n}-1]$$

Propriedade

X/M é um espaço vetorial com as operações definidas por:

- adição:

$$[x_1] + [x_2] = [x_1 + x_2]$$

- multiplicação escalar:

$$\alpha[x] = [\alpha x]$$

cuja consistência verificamos com:

$$x_1 - y_1 \in M \Rightarrow x_1, y_1 \in [x_1]$$

$$x_2 - y_2 \in M \Rightarrow x_2, y_2 \in [x_2]$$

Agora como

$(x_1 + x_2) - (y_1 + y_2) \in M$ então

$(x_1 + x_2), (y_1 + y_2) \in [x_1 + x_2]$

Se $x, y \in [x]$, $x-y \in M$ então

$\alpha(x-y) \in M$

ou seja: $\alpha x - \alpha y \in M$.

Logo $\alpha x, \alpha y \in [\alpha x]$

CAPITULO II - ESPAÇOS VETORIAIS NORMADOS

II.1. DEFINIÇÃO E EXEMPLOS

No estudo elementar de Algebra Linear os conceitos geométricos de distância em $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$ são apresentados de modo intuitivo e formalizados com a introdução da norma. Esta possibilita estender a "medição de distâncias" ao espaço mais geral $\mathbb{R}^n$.

Por outro lado, em Cálculo elementar, a noção de limite nos reais é ferramenta básica para a definição de propriedades fundamentais como a convergência de seqüências e a continuidade de funções.

Esses conceitos podem ser estendidos a espaços vetoriais com a introdução de uma métrica que tem basicamente a função de "medir distâncias" no espaço vetorial que passa então a ser dotado de uma estrutura de espaço métrico.

No nosso caso trataremos unicamente de um tipo particular de métrica que é a associada a uma norma. Diremos então que um Espaço Vetorial Normado é um espaço vetorial X no qual é definido uma Norma.

<u>Definição 1</u>: Uma norma é uma aplicação que faz corresponder a cada elemento $x \in X$ um número real positivo $\|x\|$

$$\begin{aligned} \|\cdot\| : X &\to \mathbb{R}^+ \\ x &\mapsto \|x\| \end{aligned}$$

denominado norma de x, devendo satisfazer aos seguintes axiomas:

(I) $\|x\| \geq 0 \quad , \forall x \in X$

$\|x\| = 0 \quad$, se e somente se $x = \vec{0}$

(II) $\|x+y\| \leq \|x\| + \|y\| \quad , \forall x,y \in X$

(III) $\|\alpha x\| = |\alpha| \cdot \|x\| \quad , \forall x \in X$ e α escalar.

<u>Exemplo 1</u>: O espaço vetorial $C^0[a,b]$ de todas as funções contínuas no intervalo $[a,b]$ se torna um espaço vetorial normado se a ele associarmos a norma

$$\| f \| = \max_{0 \leq x \leq b} |f(x)|$$

denominada "norma do máximo".

<u>OBS</u>: Alguns autores usam também o símbolo $\| \cdot \|_{0,\infty}$ para a norma do máximo, onde o 1º índice corresponde a ordem de deriva - das envolvidas na definição do espaço, tendo o 2º índice a conotação de máximo.

Verificação dos axiomas:

(I) Positividade:

$$\max_{a \leq x \leq b} |f(x)| \geq 0 \quad \Rightarrow \quad \| f \| \geq 0$$

$\| f \| = 0$ se e somente se f é a função identicamente nula.

(II) Desigualdade do Triângulo

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \qquad \forall x \in [a,b] \text{ , donde}$$

$$\max_{a \leq x \leq b} |f(x) + g(x)| \leq \max_{a \leq x \leq b} |f(x)| + \max_{a < x < b} |g(x)|$$

(III) Homogeneidade:

$$|\alpha f(x)| = |\alpha| \, |f(x)| \qquad \forall x \in [a,b] \text{ , donde}$$

$$\max_{a \leq x \leq b} |\alpha \, f(x)| = |\alpha| \, . \max_{a \leq x \leq b} |f(x)|$$

<u>Exemplo 2</u>: O espaço das funções contínuas no intervalo [a,b] se torna um espaço vetorial normado com a norma

$$\| f \|_{0,1} = \int_a^b |f(x)| dx$$

Devemos observar que este espaço normado e diferente do espaço normado do exemplo 1.

O primeiro índice de $\|\cdot\|_{0,1}$ diz respeito à ordem de derivadas envolvidas no caso e o segundo ao expoente da função a integrar.

Exemplo 3: O espaço vetorial $C^1[a,b]$ constituido de todas as funções contínuas que tem derivada de primeira ordem contínua no intervalo $[a,b]$ é um espaço vetorial normado com a norma

$$\| f\|_{1,\infty} = \max_{a\le x\le b} |f(x)| + \max_{a\le x\le b} |f'(x)|$$

Exemplo 4: A reta $\mathbb{R}$ passa a ser um espaço vetorial normado com a norma do valor absoluto

$$\| x\| = |x| \quad \forall x \in \mathbb{R}$$

Exemplo 5: O espaço Euclidiano é um espaço constituido de n-uplas com a norma de um elemento

$$x = \{x_1, x_2, \ldots, x_n\}$$

definida por

$$\| x\|_2 = \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2}$$

isto é, o espaço vetorial $\mathbb{R}^n$ associado à norma $\|\cdot\|_2$ dita norma euclidiana.

OBS: Para a verificação do axioma II ver a desigualdade de Minkowski apresentada mais adiante.

Exemplo 6: O espaço das seqüências infinitas que tem finitas componentes diferentes de zero é um espaço vetorial normado com a definição de normas apropriadas. Por exemplo, se

$$x = \{x_1, x_2, \ldots, x_n, 0, 0, \ldots\}$$

é uma seqüência do espaço então

$$\| x\|_1 = \sum_{i=1}^{n} |x_i|$$

é uma norma.

Analogamente definindo-se

$$\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$$

teremos outro espaço vetorial normado. Se n é fixo, obtemos duas normas para $\mathbb{R}^n$, respectivamente.

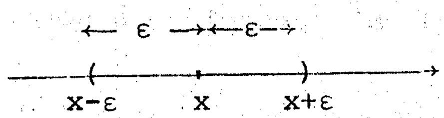
II.2. CONJUNTO ABERTO E FECHADO

Para a análise de limites e de convergência em espaços vetoriais normados é muito útil o conceito de "bola aberta".

<u>Definição 2</u>: Em um espaço vetorial normado X, chamamos de bola aberta de centro em $x \in X$ e raio $\varepsilon > 0$, e denotada por $B(x,\varepsilon)$, ao conjunto de pontos $y \in X$ que satisfazem à condição

$$\|x-y\| < \varepsilon$$

A bola nos fornece a idéia de vizinhança em relação a um determinado elemento x de um espaço vetorial normado. Uma bola aberta no espaço $\mathbb{R}$ seria um intervalo aberto como ilustrado:



No espaço $\mathbb{R}^2$ uma bola seria um círculo de raio ε e no espaço $\mathbb{R}^3$ uma esfera de raio ε. Em geral, porém ela não tem necessariamente um aspecto geométrico, tudo depende da natureza do espaço e de sua norma.

Uma primeira noção de limite num espaço vetorial normado é dada na:

<u>Definição 3</u>: Seja M um subconjunto de um espaço vetorial normado X. Um ponto $x \in X$ é denominado "Ponto de Acumulação" de M se $\forall \varepsilon > 0$ existe pelo menos um ponto de M contido em $B(x,\varepsilon)$.

No intervalo fechado $M = [a,b]$, a e b são pontos de acumulação de M e pertencem ao conjunto M.

No intervalo aberto $M = (a,b)$, a e b também são pontos de acumulação de M porém neste caso eles não pertencem a M. É evidente que todos os pontos de um conjunto são seus pontos de acumulação. A recíproca no entanto nem sempre é verdadeira.

<u>Definição 4:</u> Seja A um subconjunto de um espaço vetorial normado X. A é um conjunto aberto se para qualquer elemento $x \in A$ pode-se construir uma bola de raio $\varepsilon > 0$ totalmente contida em A.

Um conjunto vazio é por definição aberto. Um espaço vetorial normado é um conjunto aberto.

O intervalo da reta $A = (a,b)$ é um conjunto aberto ($x \in A \Rightarrow a < x < b$).

Nota-se imediatamente que um conjunto aberto não vazio pode possuir pontos de acumulação que não lhe pertencem.

<u>Definição 5:</u> Seja M um subconjunto de um espaço vetorial normado X. Chamamos de "fecho de M" o conjunto de todos os pontos de acumulação de M que denotamos por $\overline{M}^X$ (ou simplesmente, $\overline{M}$ quando não houver possibilidade de confusão).

<u>Definição 6:</u> Um subconjunto F de um espaço vetorial X é um conjunto fechado se ele contém todos os seus pontos de acumulação. Podemos dizer portanto que F é fechado se $\overline{F}^X = F$.

Um conjunto vazio é também por definição um conjunto fechado.

O intervalo da reta $F = [a,b]$ ($x \in F \Leftrightarrow a \leq x \leq b$) é um conjunto fechado.

Uma esfera unitária de $\mathbb{R}^3$ constituida de todos os pontos x com $|x| \leq 1$ é um conjunto fechado.

Um conjunto constituído de um único elemento é fechado.

<u>Proposição 1:</u> O complemento de um conjunto aberto é um fechado e o complemento de um fechado é um aberto.

Demonstração:

Seja A um conjunto aberto de um espaço vetorial normado X e $\bar{A}$ o seu complemento ($\forall x \in X$, $x \in \bar{A} \iff x \notin A$).

Se $x \in A$ e A é aberto, x não é ponto de acumulação de $\bar{A}$, pois se for, existe $B(x,\varepsilon)$ contida inteiramente em A e qualquer que seja essa bola, ela contém necessariamente pontos de $\bar{A}$, então $A \cap \bar{A} \neq \emptyset$, o que é absurdo. Logo $\bar{A}$ contém todos os seus pontos de acumulação e é portanto fechado.

Seja F um conjunto fechado de X e $\bar{F}$ o seu complemento.

Em qualquer ponto $x \in \bar{F}$ ($x \notin F$) pode-se construir $B(x,\varepsilon)$ totalmente contida em $\bar{F}$, pois se isso não fosse possível qualquer $B(x,\varepsilon)$ conteria pelo menos um ponto F. Isto seria absurdo porque nesse caso x seria um ponto de acumulação de F e conseqüentemente teríamos $x \in F$. Isso está em contradição com $x \in \bar{F}$.

c.q.d.

Proposição 2: A interseção de um número finito de conjuntos abertos é um conjunto aberto.

Demonstração:

Seja inicialmente A a inserteção não vazia de n conjuntos abertos A_i, $i = 1,2,\ldots,n$.

Em um ponto qualquer x de A construimos n bolas $B_i(x,\varepsilon_i)$ com centro em x e raio ε_i, cada uma delas contida em A_i. ($B_1(x,\varepsilon_1) \subset A_1$, $B_2(x,\varepsilon_2) \subset A_2,\ldots$). Tomando-se a bola $B(x,\varepsilon)$ onde $\varepsilon = \min(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$, ela estará contida em todos os conjuntos A_i, e portanto também na interseção A. Logo em um ponto qualquer $x \in A$ existe uma bola $B(x,\varepsilon)$ totalmente contida em A e assim A é um conjunto aberto.

Enfim, se A é vazio ele é aberto por definição.

c.q.d.

Proposição 3: Qualquer união de conjuntos abertos é um conjunto aberto.

Demonstração:

Seja A a união não vazia de um número qualquer de conjuntos abertos A_i.

Um ponto qualquer x de A pertence necessariamente a pelo menos um dos conjuntos abertos A_j e podemos construir a bola $B(x,\varepsilon_j)$ contida totalmente em A_j. Estando contida em A_j estará contida em A, portanto A é um conjunto aberto.

Se A é vazio ele é aberto por definição.

c.q.d.

<u>Proposição 4</u>: A interseção F de um número arbitrário de conjuntos fechados F_i é um conjunto fechado.

<u>Demonstração</u>:

Se A é a união de um número arbitrário de conjuntos abertos A_i, de acôrdo com a Proposição 3, A é um conjunto aberto.

Tomemos agora o complemento de F e de todos os conjuntos F_i. De acordo com a Proposição 1 o complemento de um conjunto aberto é um conjunto fechado e vice-versa.

Como $\bar{F} = \overline{\bigcap_i F_i} = \bigcup_i \bar{F}_i$

e pela Proposição 3, $\bigcup_i \bar{F}_i$ é aberto, está provada a proposição.

c.q.d.

<u>Proposição 5</u>: A união F de um número finito de conjuntos fechados F_i é um conjunto fechado.

<u>Demonstração</u>:

De acordo com a proposição 1, o complemento de um conjunto aberto é um fechado. Tomando-se agora os complementos dos conjuntos F_i e do conjunto F, pela Proposição 2, concluimos como na Proposição 4, que a união de um número finito de conjuntos fechados é um conjunto fechado.

c.q.d.

II.3. CONVERGÊNCIA EM ESPAÇOS VETORIAIS NORMADOS

Vejamos agora como podemos utilizar a norma para introduzir noções de limite e convergência em espaços vetoriais.

Definição 7: Uma seqüência infinita $\{x_n\}$ de vetores de um espaço vetorial normado X, converge para $x \in X$, se para qualquer $\varepsilon>0$, existe um inteiro $N(\varepsilon)$ tal que para todo $n>N(\varepsilon)$ tenhamos $x_n \in B(x,\varepsilon)$.

Utilizaremos a notação $x_n \to x$, se $\{x_n\}$ converge para x, que é dito limite da seqüência.

Decorre imediatamente da definição, que se $x_n \to x$ então $\forall \varepsilon>0$, o número de vetores da seqüência contidos na bola $B(x,\varepsilon)$ é infinito, e, inversamente, é finito o número de vetores de $\{x_n\}$ não contidos em $B(x,\varepsilon)$.

Proposição 6: Se $x_n \to x$ então a seqüência de números reais

$\{\| x_n - x \|\}$ converge para zero.

Demonstração:

Se $\{\| x_n - x\|\}$ não tende para zero quando $x_n \to x$, então existe um $\varepsilon>0$, tal que, para qualquer N, $n>N$ implica em

$$\| x_n - x \| \geq \varepsilon$$

Em outras palavras, existe uma bola $B(x,\varepsilon)$, $\varepsilon>0$, que não contém nenhum ponto da seqüência, o que é absurdo pois $x_n \to x$.

c.q.d.

Pode-se demonstrar sem dificuldades que a recíproca da Proposição 6 também é verdadeira.

Proposição 7: Em um espaço vetorial normado X

$$\| x \| - \| y \| \leq \| x - y \|$$

para quaisquer vetores $x, y \in X$.

Demonstração:

$$\|x\| - \|y\| = \|x-y+y\| - \|y\|$$

$$\leq \|x-y\| + \|y\| - \|y\|$$

$$\leq \|x-y\|$$

c.q.d.

Proposição 8: Em um espaço vetorial normado se $x_n \to x$ então $\|x_n\| \to \|x\|$

Demonstração:

Sabemos que: $\|x_n\| - \|x\| \leq \|x_n - x\|$ e

$\|x\| - \|x_n\| \leq \|x_n - x\|$

Então: $\left| \|x_n\| - \|x\| \right| \leq \|x_n - x\|$

Se $x_n \to x$, $\|x_n - x\| \to 0$ logo

$\left| \|x_n\| - \|x\| \right| \to 0$ e

$\|x_n\| \to \|x\|$

c.q.d.

Proposição 9: Se uma seqüência $\{x_n\}$ converge para x, o limite x é único.

Demonstração:

Vamos supor que: $x_n \to x$ e $x_n \to y$.

$$\|x-y\| = \|x - x_n + x_n - y\|$$

$$\leq \|x - x_n\| + \|x_n - y\|$$

Se $x_n \to x$, $\|x - x_n\| \to 0$

Se $x_n \to y$, $\|x_n - y\| \to 0$

logo $\|x-y\| = 0$

isto é: $x = y$.

c.q.d.

Teorema 1:

Um subconjunto M de um espaço vetorial normado X é um conjunto fechado se e somente se qualquer seqüência $\{x_n\}$ de elementos de M, que converge, tiver limite em M.

$$M \text{ fechado} \iff (\forall \{x_n\} \subset M / x_n \to x \Rightarrow x \in M)$$

Demonstração:

Seja $\{x_n\}$ uma seqüência qualquer de elementos de M.

a) Hipótese: $\begin{cases} x_n \to x\ ,\ \{x_n\} \subset M \\ M \text{ é fechado} \end{cases}$

Tese: $x \in M$

Se $x_n \to x$, pela definição de convergência, toda bola $B(x,\varepsilon)$, $\varepsilon>0$, contém elementos de $\{x_n\}$, e por conseguinte de M. Logo x é ponto de acumulação de M.

Como M é fechado, M contém todos os seus pontos de acumulação, donde $x \in M$.

b) Hipótese: $\begin{cases} \forall \{x_n\} \text{ tal que } x_n \to x \\ \text{temos } x \in M \end{cases}$

Tese: M é fechado

Vamos supor que M não é fechado e mostrar que isto contraria a hipótese estabelecida.

Se M não é fechado, existe um ponto de acumulação $x \in X$, tal que, $x \notin M$. Vamos agora provar que nesse caso podemos construir uma seqüência de elementos de M que converge para x.

Por exemplo, usando as bolas $B(x,1/n)$, $n = 1,2,3,\ldots\ldots$ que contêm pelo menos um elemento de M, selecionamos em cada bola um elemento x_n. Obtemos uma seqüência $\{x_n\} \subset M$ que estará convergindo para x, que não pertence ao conjunto M. Isto contraria a nossa suposição e portanto M só pode ser um conjunto fechado.

C.q.d.

II.4. SUPREMO E ÍNFIMO

Antes de prosseguir tratando a convergência e limites em espaços vetoriais quaisquer, vamos nos restringir por um momento aos reais. Trata-se de introduzir conceitos extremamente importantes em Análise Funcional relativos a subconjuntos de $\mathbb{R}$ que usaremos a seguir.

<u>Definição 8</u>: Seja P um subconjunto não vazio de $\mathbb{R}$. Dizemos que P é majorado por um número <u>a</u>, se <u>a</u> é maior ou igual a todos os elementos de P. Nesse caso <u>a</u> é dito majorante de P. Ao menor dos majorantes de P dá-se o nome de <u>Supremo de P</u> e escreve-se:

$$s = \sup P \text{ ou } s = \sup_{x \in P} x$$

<u>Observações</u>:

a) Se $P = [a,b]$
$\sup P = b$, e neste caso $\sup P \quad P$

b) Se $P = (a,b)$
$\sup P = b$ e $\sup P \notin P$

c) No caso em que $s = \sup P$ e $s \in P$, s é denominado também de máximo de P.

d) O supremo é uma grandeza que sempre existe, podendo inclusive assumir valores infinitos. Porém não se pode dizer o mesmo do máximo.

<u>Propriedade</u>: $\forall x \in P, \quad x \leq \sup P$
decorrência direta da definição de supremo.

<u>Proposição 10</u>: Para qualquer $\varepsilon > 0$, existe $x \in P$, tal que: $s - \varepsilon < x \leq s$, onde $s = \sup P$.

OBS: Daí ocorre que, se s for finito, s é um ponto acumulação de P.

Demonstração

Se para certo $\varepsilon>0$ não existe $x\in P$, tal que, $s-\varepsilon<x \leq s$, então existe um pouco $s'<s$ $(s'=s-\varepsilon)$ que majora P. Absurdo pois $s = \sup P$ é o menor dos majorantes de P. c.q.d.

Proposição 11: Se $s = \sup P < \infty$ existe uma seqüência $\{x_n\}$ de elementos de P que converge para s, com $x_{n+1} \geq x_n$ $\forall n$

Demonstração

Pela proposição 10, nesse caso, $s = \sup P$, é um ponto de acumulação de P. Poderemos então construir bolas com centros em s e diferentes raios $\varepsilon_n>0$, que conterão pontos de P.

Tomemos uma seqüência de bolas $B(s,\varepsilon_n)$ com raios decrescentes, $\varepsilon_n \to 0$ e $\varepsilon_{n+1} \leq \varepsilon_n$ $\forall n$. Para cada uma delas podemos selecionar um ponto x_n com $x_n \in P$ e $x_n = s - \varepsilon_n$.

Essa seqüência $\{x_n\}$ de elementos de P converge para $s = \sup P$ com $x_n \leq x_{n+1}$, $\forall n$.

Definição 9: Seja P um sub-conjunto não vazio de $\mathbb{R}$. Dizemos que P é minorado por um número b se b é menor ou igual a todos os elementos de P. b é denominado um minorante de P. Ao maior dos minorantes de P chamamos de ínfimo de P e denotamos

$$m = \inf P \quad \text{ou}$$

$$m = \inf_{x \in P} x$$

Se $m = \inf P$ e $m \in P$, m é denominado também de mínimo de P.

Propriedade: $\forall x \in P$, $x \geq \inf P$.

Proposição 12: Se $m = \inf P$ for finito então m é um ponto de acumulação de P.

Demonstração:

É análoga à da Proposição 10.

Proposição 13: Se $m = \inf P$, existe uma seqüência $\{x_n\}$ de elementos de P que converge para m com $x_{n+1} \le x_n$, $\forall n$

Demonstração:

É análoga à da Proposição 11.

II. ESPAÇOS ℓ^p

Vamos agora estudar, em detalhes, espaços de seqüências que se prestam para ilustração de alguns conceitos importantes em Análise Funcional.

Definição 10: Chamamos de "espaço ℓ^p" o conjunto constituido de todas as seqüências infinitas de reais

$$x = \{x_1, x_2, \ldots, x_i, \ldots\ldots\}$$

para as quais

$$\sum_{i=1}^{\infty} |x_i|^p < \infty, \quad 1 \le p < \infty$$

A rigor seria preciso verificar que ℓ^p é efetivamente um espaço vetorial. A dificuldade essencial é provar que

$$x, y \in \ell^p \Rightarrow x+y \in \ell^p$$

que é uma conseqüência imediata da desigualdade de Minkowski provada na Proposição 16. Supondo no momento que o resultado é verdadeiro, ℓ^p se torna um espaço vetorial normado com a norma:

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p}$$

Definição 11: O espaço ℓ^∞ é o conjunto constituido de seqüências infinitas limitadas, isto é, para as quais existe $M>0$, tal que $|x_i| < M$, $\forall i$.

Pode-se facilmente verificar que ℓ^∞ é um espaço vetorial normado com

$$\|x\|_\infty = \sup_i |x_i|$$

<u>Proposição 14</u>: Se x pertence ao espaço ℓ^{p_0}, x pertencerá a todos os espaços ℓ^p com $p \geq p_0$.

<u>Demonstração</u>:

Seja $x = \{x_1, x_2, \ldots, x_i, \ldots\} \in \ell^{p_0}$ $1 \leq p_0 \leq \infty$

Temos então que:

$$\sum_{i=1}^{\infty} |x_i|^{p_0} < \infty$$

e portanto existe um escalar $M>0$ tal que

$$\sup_i |x_i| = M < \infty$$

Dividindo-se todos os componentes x_i de x pelo escalar M, o valor absoluto de cada elemento resultante será:

$$\frac{|x_i|}{M} \leq 1$$

Então poderemos escrever para $1 \leq p < \infty$; $p \geq p_0$.

$$\sum_{i=1}^{\infty} \left(\frac{|x_i|}{M}\right)^p \leq \sum_{i=1}^{\infty} \left(\frac{|x_i|}{M}\right)^{p_0} \quad \text{para} \quad \forall p \geq p_0$$

$$\frac{1}{M^p} \sum_{i=1}^{\infty} |x_i|^p \leq \frac{1}{M^{p_0}} \sum_{i=1}^{\infty} |x_i|^{p_0} .$$

Como $\frac{1}{M^{p_0}}$ e $\sum_{i=1}^{\infty} |x_i|^{p_0}$ são finitos, temos que

$$\frac{1}{M^p} \cdot \sum_{i=1}^{\infty} |x_i|^p < \infty.$$

Como $\frac{1}{M^p}$ é finito, temos finalmente

$$\sum_{i=1}^{\infty} |x_i|^p < \infty, \text{ i.e., } x \in \ell^p \quad \forall p \geq p_0, \ p < \infty .$$

Enfim como $x \in \ell^{p_0}$, $\sup_i |x_i| < \infty$. Temos $x \in \ell^\infty$. c.q.d.

Observa-se que a recíproca da proposição 13 não é verdadeira, isto é, $x \in \ell^{p_0}$ não se pode afirmar que $x \in \ell^p$ se $p<p_0$

Podemos mostrar por exemplo que

$$x = \{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\ldots\}$$

pertence ao espaço ℓ^2 porém não pertence ao espaço ℓ^1. De fato ,

$$\sum_{n=1}^{\infty} \left(\frac{1}{n}\right)^2 \leq \int_1^{\infty} \frac{1}{x}\,dx + 1$$

$$\leq \frac{1}{x}\Big|_1^{\infty} + 1$$

$$\leq 2$$

$$< \infty \qquad \forall x \in \ell^2$$

Por outro lado temos:

$$\sum_{n=1}^{\infty} \frac{1}{n} \geq \int_1^{\infty} \frac{1}{x}\,dx$$

$$\geq \mathrm{Ln}\,x \Big|_1^{\infty} = \infty$$

Logo $x \notin \ell^1$

Vejamos agora duas desigualdades fundamentais válidas para seqüências de ℓ^p.

<u>Proposição 15</u>: (Desigualdade de Hölder)

Seja

$x = \{x_1, x_2, \ldots, x_i \ldots\} \in \ell^p \quad 1 \leq p \leq \infty$

$y = \{y_1, y_2, \ldots, y_i \ldots\} \in \ell^q \quad 1 \leq q \leq \infty$

com: $\frac{1}{p} + \frac{1}{q} = 1$. Então:

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \ \|y\|_q$$

Demonstração:

a) Mostraremos inicialmente que a relação é válida para $p = 1,\infty$ e $q = \infty,1$.

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \left(\sum_{i=1}^{\infty} |x_i|\right) . \sup_i y_i$$

$$\leq \|x\|_1 . \|y\|_\infty$$

b) Seja $1 \leq p < \infty$ e $1 \leq q < \infty$. Consideremos a função $f(t)$ definida para $t \geq 0$.

$$f(t) = t^\lambda - \lambda t + \lambda - 1 \qquad 0 < \lambda < 1$$

Temos:

$$f'(t) = \lambda(t^{\lambda-1} - 1) \text{ donde: } \begin{cases} f'(t) > 0 & \text{se } 0 \leq t < 1 \\ f'(t) < 0 & \text{se } t > 1 \end{cases}$$

Logo, para $t \geq 0$

$$f(t) \leq f(1) = 0$$

donde: $t^\lambda \leq \lambda t + 1 - \lambda$

com igualdade somente para $t=1$.

Fazendo

$$t = a/b \begin{cases} a \geq 0 \\ b > 0 \end{cases} \text{ vem:}$$

$$\frac{a^\lambda}{b^\lambda} \leq \lambda \frac{a}{b} + 1 - \lambda \text{ , ou seja}$$

$$(1) \qquad \boxed{a^\lambda b^{(1-\lambda)} \leq \lambda a + (1-\lambda) b}$$

com igualdade só para $a=b$. Se $b=0$ em (1), a desigualdade ainda continuará válida. Logo podemos considerar, $a \geq 0$, $b \geq 0$.

Fazendo agora:

$$a = \left(\frac{|x_i|}{\|x\|_p}\right)^p \qquad b = \left(\frac{|y_i|}{\|y\|_q}\right)^q$$

$$\lambda = \frac{1}{p} \quad e \quad 1-\lambda = \frac{1}{q}$$

e substituindo em (1) teremos:

$$\frac{|x_i \cdot y_i|}{\|x\|_p \|y\|_q} \leq \frac{1}{p}\left(\frac{x_i}{\|x\|_p}\right)^p + \frac{1}{q}\left(\frac{|y_i|}{\|y\|_q}\right)^q$$

Somando membro a membro os n primeiros termos vem:

$$\frac{\sum_{i=1}^{\infty} |x_i \cdot y_i|}{\|x\|_p \ \|y\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

e finalmente $\displaystyle\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \cdot \|y\|_q$

De acordo com (1) a igualdade é válida só para a=b, ou seja:

$$\left(\frac{|x_i|}{\|x\|_p}\right)^p = \left(\frac{|y_i|}{\|y\|_q}\right)^q \quad \forall i$$

Elevando-se ambos os termos a $\frac{1}{pq}$ vemos que a igualdade vale se e somente se:

$$\left(\frac{|x_i|}{\|x\|_p}\right)^{1/q} = \left(\frac{|y_i|}{\|y\|_q}\right)^{1/q} \quad \forall i \qquad \text{c.q.d.}$$

Caso particular: (Desigualdade de Cauchy - Schwarz).

Se $p = q = 2$ a desigualdade de Hölder fica

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{1/2} \left(\sum |y_i|^2\right)^{1/2}$$

Proposição 16: (Desigualdade de Minkowski)

Seja $x, y \in \ell^p \quad 1 \leq p \leq \infty$

Então vale:

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p$$

Demonstração:

Para $p = 1$ e $p = \infty$ o resultado é imediato. Mostraremos então que a desigualdade é válida para $1<p<\infty$.

Seja

$$x = \{x_1, x_2, \ldots, x_i, \ldots\ldots\} \in \ell^p \qquad 1<p<\infty$$
$$y = \{y_1, y_2, \ldots, y_i, \ldots\ldots\} \in \ell^p \qquad 1<p<\infty$$

Temos:

$$\sum_{i=1}^{n} |x_i+y_i|^p = \sum_{i=1}^{n} |x_i+y_i|^{p-1} \cdot |x_i+y_i|$$

$$\leq \sum_{i=1}^{n} |x_i+y_i|^{p-1} \cdot |x_i| + \sum_{i=1}^{n} |x_i+y_i|^{p-1} \cdot |y_i| \quad \forall n$$

Aplicando a desigualdade de Hölder para cada somatório à direita da desigualdade temos:

$$\sum_{i=1}^{n} |x_i+y_i|^p \leq \left[\sum_{i=1}^{n} |x_i+y_i|^{q(p-1)}\right]^{1/q} \cdot$$

$$\cdot \left[\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}\right]$$

Como $\frac{1}{p} + \frac{1}{q} = 1$, temos que $q(p-1) = p$, o que dá:

$$\sum_{i=1}^{n} |x_i+y_i|^p \leq \left(\sum_{i=1}^{n} |x_i+y_i|^p\right)^{1/q} \left[\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \right.$$

$$\left. + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}\right] \quad \forall n$$

Dividindo os dois membros da desigualdade por

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/q}$$

e considerando-se que $1 - \frac{1}{q} = \frac{1}{p}$ temos

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p} \quad \forall n$$

Podemos agora passar ao limite no 2º membro pois estaremos majorando ainda mais o lado maior.

$$\left(\sum_{i=1}^{n} |x_i+y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{1/p} \quad \forall n$$

Enfim, como a desigualdade acima vale $\forall n$ temos:

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p \qquad \text{c.q.d.}$$

Observações:

1) Se $p=2$ a desigualdade de Minkowski prova o segundo axioma da norma Euclidiana (n finito).

2) Está agora provado que ℓ^p é um espaço vetorial.

II.6. DENSIDADE E SEPARABILIDADE

Com a introdução da norma podemos considerar casos particulares de inclusão de conjuntos (e/ou espaços) uns nos outros, nos quais elementos de um conjunto podem ser vistos como limites de seqüências de um seu subconjunto.

Definição 12: Um sub-conjunto D de um conjunto E de um espaço vetorial normado X é dito "denso em E" se para cada elemento $x \in E$ e cada $\varepsilon > 0$ existe $d \in D$, tal que, $d \in B(x,\varepsilon)$ (ou seja $\|d-x\| < \varepsilon$).

Pode-se dizer em outras palavras que se D é denso em E, existem pontos de D arbitrariamente próximos de cada $x \in E$. Pode-se dizer ainda que, dado um $x \in E$ é possível construir uma seqüência de elementos de D que converge para x.

Como se poderá deduzir facilmente da Definição 12, temos que $E \subseteq \bar{D}$ e que, portanto, se E é fechado $E = \bar{D}$. Assim, no caso particular importante em que E é o próprio espaço X, temos que X é o fecho de D ($\bar{D}^X = X$).

Uma propriedade de conjuntos densos que será útil posteriormente é dada na

Proposição 17:

Se D é denso em Y e Y é denso em X então D é denso em X.

Demonstração:

Se Y é denso em X, então para $\forall x \in X$ e $\forall\ \varepsilon/2 > 0$, existe $y \in Y$, tal que

$$(2) \qquad \| y-x \| < \varepsilon/2$$

Se D é denso em Y, então para $\forall y \in Y$ e $\forall\ \varepsilon/2 > 0$, existe $d \in D$, tal que

$$(3) \qquad \| d-y \| < \varepsilon/2$$

De (2) e (3) temos

$$\| y-x \| + \| d-y \| < \varepsilon$$

$$\| d-x \| < \varepsilon$$

Portanto para $\forall x \in X$ e $\forall \varepsilon > 0$ existe $d \in D$, tal que

$$\| d-x \| < \varepsilon$$

Logo D é denso em X. c.q.d.

Exemplo 7:

O conjunto dos racionais $\mathbb{Q}$ é denso no espaço dos reais $\mathbb{R}$, pois como é sabido:

$\forall x \in \mathbb{R}$ e $\forall \varepsilon > 0$

$\exists q \in \mathbb{Q}$ / $|x-q| < \varepsilon$

Exemplo 8: (Teorema da Aproximação de Weierstrass)

O espaço dos polinômios $\mathcal{P}[a,b]$ é denso no espaço $C^0[a,b]$. Isso significa que para um dado $\varepsilon>0$, qualquer que seja a função contínua f é possível encontrar um polinômio p tal que:

$$\max_{a\le X\le b} |p(x) - f(x)| < \varepsilon$$

A demonstração do Teorema da Aproximação de Weierstrass pode ser encontrada em [1].

Definição 13: Um espaço vetorial normado é separável se contém um conjunto denso que seja enumerável.

Exemplo 9: O espaço $\mathbb{R}$ dos números reais é separável pois contém o conjunto $\mathbb{Q}$ dos racionais que é denso e enumerável.

$\forall \varepsilon>0$ e $\forall r\in \mathbb{R}$, $\exists q\in \mathbb{Q}$ tal que $|r-q| < \varepsilon$

OBS: Admitimos conhecido que o conjunto $\mathbb{Q}$ é enumerável.

Exemplo 10:

O espaço $\mathbb{R}^n$ é separável, pois contém o conjunto $\mathbb{Q}^n$ de vetores com componentes racionais que é denso (provar) e enumerável porque $\mathbb{Q}^n$ é a união enumerável de conjuntos enumeráveis [9]. Isto se pode provar por indução finita.

1º) Sabemos que $\mathbb{Q} \equiv \mathbb{Q}^1$ é enumerável.

2º) Admitindo-se que $\mathbb{Q}^{n-1}$ é enumerável, $\mathbb{Q}^n$ pode ser visto como a união conjuntos $\mathbb{Q}_i$ de vetores de $\mathbb{R}^n$ da forma $\{q_i,q\}$, onde $q\in\mathbb{Q}^{n-1}$ e q_i é o i-ésimo racional da numeração de $\mathbb{Q}$.

Logo, como supomos $\mathbb{Q}^{n-1}$ enumerável $\mathbb{Q}^n = \bigcup_i \mathbb{Q}_i$ é enumerável.

<u>Exemplo 11:</u>

Os espaços ℓ^p, $1 \le p < \infty$ são separáveis. De fato, seja D o conjunto de todas as seqüências com número finito de componentes racionais diferentes de zero. D é enumerável pois pode ser visto como a união enumerável dos conjuntos enumeráveis de seqüências de i termos racionais não nulos, $i = 1,2,3,\ldots$ (verificar).

Provemos agora que D é denso em ℓ^p.

Seja $x = (x_1, x_2, \ldots, x_i, \ldots) \in P$.

Para um dado $\varepsilon > 0$, existe um N tal que

$$\sum_{i=N+1}^{\infty} |x_i|^p < \varepsilon^p/2$$

Para cada k, $1 \le k \le N$, seja o racional q_k, tal que

$|x_k - q_k|^p < \varepsilon^p/2N$ e se $d = \{q_1, q_2, \ldots, q_N, 0, 0, \ldots\ldots\}$ então

$$\|x-d\|^p = \sum_{i=1}^{\infty} |x_i - q_i|^p$$

$$= \sum_{k=1}^{N} |x_k - q_k|^p + \sum_{i=N+1}^{\infty} |x_i|^p$$

$$< N.\ \frac{\varepsilon^p}{2N} + \frac{\varepsilon^p}{2}$$

donde:

$$\|x-d\| < \varepsilon$$

D é portanto denso em ℓ^p.

<u>Exemplo 12:</u>

O espaço ℓ^∞ não é separável. Para mostrar este fato basta exibir um contraexemplo.

Consideremos a família de seqüências de ℓ^∞ da forma

$$x_S = \{x_1, x_2, x_3, \ldots, x_i, \ldots\ldots\}$$

onde $x_i = 0$ ou 1.

Aqui S é índice que, provaremos, percorre um conjunto indexador não enumerável Σ.

Seja $x_{S'} = \{x'_1, x'_2, x'_3, \ldots, x'_i, \ldots\}$ onde $S' \neq S$. Então $\exists i$, $1 \leq i < \infty$ tal que $x_i \neq x'_i$. Logo

$$\|x_S - x_{S'}\|_\infty = 1 \quad \text{se} \quad S \neq S'$$

Portanto qualquer elemento da família, não tem nenhum outro que esteja arbitrariamente próximo dele. Consideremos $\varepsilon < 1/2$ e um conjunto D denso em ℓ^∞. Para cada S deve existir $d_S \in D$ tal que

$$\|d_S - x_S\|_\infty < \varepsilon$$

Evidentemente $d_S \neq d_{S'}$, se $S \neq S'$.

Assim se ficar provado que a família $\{x_S\}_{S\in\Sigma}$ não é enumerável, o conjunto $\{d_S\}_{S\in\Sigma} \subset D$ não é enumerável e por conseguinte a enumerabilidade de qualquer conjunto D denso em ℓ^∞ é absurda.

O leitor familiarizado com a representação binária dos reais não terá dificuldade em constatar que a cada elemento de Σ corresponde um e somente um real s tal que $0 \leq s < 1$.

Como $[0,1)$ não é enumerável a família $\{x_S\}_{S\in\Sigma}$ não pode ser enumerável.

Exemplo 13:

O espaço $\mathcal{C}[0,1]$ é separável. Pelo Teorema da aproximação de Weierstrass o espaço dos Polinômios $\mathcal{P}[0,1]$ é denso em $C^0[0,1]$.

O conjunto de todos os polinômios de coeficientes racionais $\mathcal{R}[0,1]$ é denso em $\mathcal{P}[0,1]$.

Seja $p \in \mathcal{P}[0,1]$, $p = \sum_{i=0}^{n} a_i x^i$. Para qualquer que seja $\varepsilon > 0$, pode-se encontrar racionais q_i, $i = 0,1,\ldots,n$ tais que

$$|q_i - a_i| < \frac{\varepsilon}{n+1} \qquad \forall i = 0,1,\ldots,n$$

Logo $\quad |q_i x^i - a_i x^i| < \frac{\varepsilon}{n+1} \quad \forall x \in [0,1]$, donde

(4) $$\boxed{\left|\sum_{i=0}^{n} q_i x^i - \sum_{i=0}^{n} a_i x^i\right| < \varepsilon} \qquad \forall x \in [0,1]$$

Como (4) é válida para qualquer x, é válida também para o máximo, logo

(5) $$\boxed{\| r-p \|_{0,\infty} < \varepsilon}$$

onde $r = \sum_{i=0}^{n} q_i x^i$,

Ou seja, para qualquer que seja $p \in \mathcal{P}[0,1]$ e $\varepsilon > 0$, existe $r \in \mathcal{R}$ tal que (5) se verifique, ficando demonstrado que $\mathcal{R}[0,1]$ é denso em $\mathcal{P}[0,1]$.

Como $\mathcal{P}[0,1]$ é denso em $C^0[0,1]$ e $\mathcal{R}[0,1]$ é denso em $\mathcal{P}[0,1]$, pela Proposição 17 concluimos que $\mathcal{R}[0,1]$ é denso em $C^0[0,1]$.

Enfim $\mathcal{R}[0,1]$ é enumerável pois podemos identificar cada polinômio de $\mathcal{R}[0,1]$ a uma e uma só seqüência de finitas componentes racionais não nulas, conjunto esse que é enumerável (ver Exemplo 11). $C^0[0,1]$ é portanto separável.

A importância de um espaço vetorial normado X ser separável é que qualquer $x \in X$ pode ser gerado por uma base enumerável a menos de um $\varepsilon > 0$ arbitrariamente pequeno. De fato, seja D um conjunto denso e enumerável em X, i.e., $\bar{D}^X = X$, e seja [D] o espaço gerado por D. Evidentemente $[D] \subset X$. Então como D é denso em X, $D \subset [D]$ e $[D] \subset X$, temos que [D] é denso em X (verificar!).

Seja agora $B = \{e_1, e_2, \ldots, e_n, \ldots\} \subset D$ um conjunto livre que gere [D].

B é uma base enumerável de [D]. Qualquer que seja $y \in [D]$ temos que

$$y = \sum_{i=1}^{n} \alpha_i e_i \quad \text{para certos escalares } \alpha_i \text{ e } 1 \le n < \infty$$

Logo se $x \in X$, pela densidade de [D] em X podemos escrever que

$$\left\| \sum_{i=1}^{n} \alpha_i e_i - x \right\| < \varepsilon \quad \forall\varepsilon>0$$

Quando tal ocorre, diz-se freqüente e impropriamente que

$\{e_1, e_2, \ldots, e_n, \ldots\}$ é uma "base" de X (exato a menos de um $\varepsilon>0$).

II.7. COMPACIDADE

Assim como abertos e fechados, uma classe importante de conjuntos de espaços vetoriais normados, embora mais restrita, é a dos compactos. Na literatura um compacto é em geral definido equivalentemente de maneiras diferentes. Devido ao nosso interesse específico julgamos conveniente fazê-lo por meio de subseqüências assim definidas:

Dada uma seqüência $\{x_i\}$, uma sua subseqüência, é qualquer seqüência constituída de elementos de $\{x_i\}$, que tenha sido obtida por eliminação de elementos de $\{x_i\}$. Representaremos uma subseqüência de $\{x_i\}$ por $\{x_{i_n}\}$ sendo n o novo indexador.

<u>Definição 14</u>: Um conjunto K de um espaço vetorial X é dito um "compacto" se, dada uma seqüência arbitrária $\{x_i\}$ de K, existe uma sua subseqüência $\{x_{i_n}\}$ que converge em K.

Podemos também interpretar os termos da Definição 14 com: "Se K é um conjunto compacto, uma seqüência arbitrária de K possui pelo menos um ponto de acumulação".

De fato, sendo x o limite da subseqüência $\{x_{i_n}\}$, x é ponto de acumulação de $\{x_i\}$ pois $\forall\varepsilon>0$, $\exists N(\varepsilon)$ tal que,

$$\| x_{i_n} - x \| < \varepsilon , \qquad \forall n>N(\varepsilon)$$

ou seja

$$\forall\varepsilon>0,\ \exists x_{i_n} \in \{x_i\} \quad \text{tal que} \quad x_{i_n} \in B(x,\varepsilon)$$

<u>Exemplo 14</u>:

O intervalo $[-1,1]$ é um compacto de $\mathbb{R}$. (Por exemplo da seqüência $\{1,-1,1,-1,\ldots\}$ podemos extrair a subseqüência $\{1,1,1,\ldots\}$ que converge para 1, e isto é válido para qualquer seqüência do conjunto). (Ver Exemplo 17).

Exemplo 15:

Devemos observar que $\mathbb{R}$ não é compacto. Tomemos por exemplo a seqüência

$$\{1,2,3,4,\ldots\ldots\}$$

da qual não podermos extrair nenhuma subseqüência convergente.

Exemplo 16:

O intervalo (0,1) não é compacto. Se tomarmos a seqüência

$$\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\}$$

não se tem nenhuma subseqüência que convirja nesse intervalo.

Proposição 18:

Se um conjunto K de um espaço vetorial normado X é um compacto então ele é fechado e limitado.

Demonstração:

a) K é fechado

Se $\bar{K} \neq K$, $\exists\, x \in \bar{K} / x \notin K$.

Sendo x um ponto de acumulação de K, podemos construir uma seqüência $\{x_n\} \subset K$ / $x_n \to x$ e que só tem um ponto de acumulação x que não está em K, o que é absurdo.

b) K é limitado.

Se K não é limitado podemos construir uma seqüência

$\{x_n\} \subset K$ / $\|x_n\| = C + n$

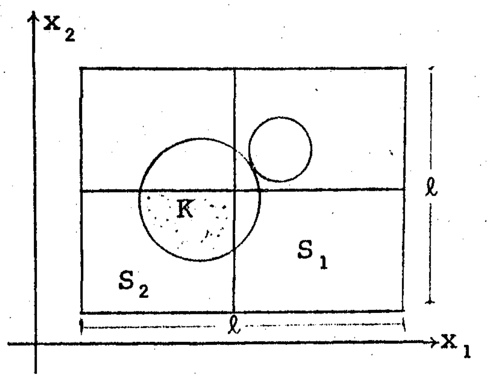
onde $C = \inf\limits_{x \in K} \|x\|$

Dessa forma $\{x_n\}$ não tem nenhum ponto de acumulação, o que é absurdo.

c.q.d.

No caso de espaços de dimensão finita pode-se provar que a recíproca da Proposição 17 é verdadeira, o que significa que um conjunto K ser fechado e limitado equivale a K ser compacto. Podemos ilustrar esse fato como

<u>Exemplo 17</u>: Consideremos um fechado e limitado K arbitrário de $\mathbb{R}^2$. De acordo com a hipótese feita podemos dizer que existe um quadrado S_1 de lados iguais a ℓ paralelos aos eixos de coordenadas tal que $S_1 \supset K$. Seja agora $\{x_i\} \subset K$. Vamos construir uma subseqüência $\{x_{i_n}\}$ de $\{x_i\}$ que converge em K.



Seja $x_{i_1} = x_1$

Agora dividimos S_1 em quatro quadrados abertos iguais. O fecho de um deles conterá uma infinidade de termos de $\{x_i\}$. Seja S_2 esse fecho e escolhemos x_{i_2} um ponto qualquer de $\{x_i\} \cap S_2$, com $i_2 > i_1$

Agora dividimos S_2 em quatro quadrados abertos iguais dos quais certo S_3 terá seu fecho contendo uma quantidade infinita de termos de $\{x_i\}$. Seja $x_{i_3} \in \{x_i\} \cap S_3$, com $i_3 > i_2$.

Aplicando o mesmo processo sucessivamente, teremos construído uma subseqüência $\{x_{i_n}\}$ de $\{x_i\}$ cujos elementos se aproximam gradualmente uns dos outros tanto quanto se queira, no sentido que:

Dado $\varepsilon > 0$ $\exists N(\varepsilon)$ inteiro tal que $\forall m,n > N(\varepsilon)$ temos:

$$\|x_{i_n} - x_{i_m}\| < \varepsilon$$

De fato, basta tomar N tal que $\varepsilon > \ell\sqrt{2}/2^N$, que a condição acima estará satisfeita por construção.

Ora, como veremos logo a seguir, tal propriedade, dita convergência no sentido de Cauchy, basta para caracterizar a convergência usual de seqüências para um limite em certa categoria de espaços, entre os quais se incluem os de dimensão finita (no caso $\mathbb{R}^2$).

Enfim, como K é fechado, o limite de $\{x_{i_n}\}$ deve estar em K que é portanto um compacto.

Em geral, entretanto, dado um espaço vetorial normado, é muito difícil caracterizar seus compactos de maneira simples.

CAPÍTULO III - ESPAÇOS COMPLETOS

III.1. SEQUÊNCIA DE CAUCHY

No Capítulo II a convergência de uma seqüência $\{x_n\}$ em um espaço vetorial normado X foi definida referindo-se a um elemento $x \in X$, limite da sequência $\{x_n\}$.

Entretanto pode-se idealizar a noção de convergência sem necessidade de se referir a esse elemento limite. Podemos imaginar a convergência como sendo uma aproximação dos elementos da seqüência, tal que a partir de um certo têrmo eles se acham tão próximos um dos outros quanto se queira.

O fato de tais seqüências convergirem efetivamente para um limite em X significará que X associado à sua norma é em certo sentido estável por passagem ao limite. Tais espaços constituem uma classe importante nas aplicações, sendo denominados espaços completos. Para formalizarmos essas noções precisamos inicialmente da:

<u>Definição 1</u>: Uma seqüência $\{x_n\}$ de um espaço vetorial normado é dita ser de Cauchy, se dado um $\varepsilon>0$, existe um inteiro $N(\varepsilon)$ tal que

$$\| x_n - x_m \| < \varepsilon \qquad \forall n,m > N(\varepsilon)$$

A título de simplificação utilizaremos também a seguinte notação:

Se $\{x_n\}$ é de Cauchy

$$\| x_n - x_m \| \to 0$$

quando n e m tendem para infinito.

<u>Proposição 1</u>:

Em um espaço vetorial normado X qualquer seqüência $\{x_n\}$ que converge para um elemento $x \in X$ é uma seqüência de Cauchy.

Demonstração:

Seja $\{x_n\} \to x$

Temos $\|x_n - x_m\| = \|x_n - x + x - x_m\|$

$$\leq \|x_n - x\| + \|x - x_m\|$$

Como $\|x_n - x\| \to 0$ e

$\|x - x_m\| \to 0$

então

$\|x_n - x_m\| \to 0$

e $\{x_n\}$ é de Cauchy. c.q.d.

A recíproca da Proposição 1 não é verdadeira, isto é, nem toda seqüência de Cauchy converge no espaço. O fato das seqüências de Cauchy de um espaço convergirem ou não nesse espaço é justamente o que servirá para classificar os espaços vetoriais normados em espaços completos e espaços não completos.

Proposição 2:

Toda seqüência de Cauchy de um espaço vetorial normado é limitada.

Demonstração:

Seja $\{x_n\}$ uma seqüência de Cauchy. Então para $\forall \varepsilon > 0, \exists N(\varepsilon)$ tal que

$$\|x_n - x_m\| < \varepsilon \quad , \quad \forall n,m > N(\varepsilon)$$

Para $n > N(\varepsilon)$ e $m = N(\varepsilon)$ temos

$$\|x_n\| = \|x_n - x_m + x_m\|$$

$$\leq \|x_n - x_m\| + \|x_m\|$$

Como m é fixo, existe uma constante $C(\varepsilon)$ tal que

$$\| x_m \| = C(\varepsilon)$$

Logo:

$$\| x_n \| \leq \varepsilon + C(\varepsilon) \qquad \forall n > N(\varepsilon)$$

Enfim $\{x_n\}$ é limitada, pois

$$\| x_n \| \leq M \qquad \forall n \quad \text{com}$$

$$M = \max \{ \max_{1 \leq n \leq N} \| x_n \|, \quad \varepsilon + C(\varepsilon) \}$$

III.2. ESPAÇOS DE BANACH

Definição 2: Um espaço vetorial normado X é um espaço "completo" se qualquer seqüência de Cauchy do espaço convergir para um elemento que pertença a X.

Definição 3: Um espaço vetorial normado é um espaço de Banach se ele for completo.

Exemplo 1:

O espaço dos números reais $\mathbb{R}$ com $\| x \| = |x|$ é um espaço completo. Tal afirmativa é fundamentada no chamado axioma de completividade dos reais [12] .

Análogamente $\mathbb{R}^n$ é um espaço de Banach com qualquer norma, o que provaremos mais adiante (ver Teorema 2).

Exemplo 2:

O espaço das funções contínuas $C^0[a,b]$ é um espaço de Banach com a norma

$$\| f \| = \max_{a \leq x \leq b} |f(x)|$$

Devemos mostrar que qualquer seqüência $\{f_n\}$ de Cauchy de $C^0[a,b]$ converge para $f \in C^0[a,b]$.

Mostraremos inicialmente que $f_n \to f$ ponto a ponto em $[a,b]$, depois que essa convergência é uniforme e finalmente mostra remos que o limite f é uma função contínua.

a) Se $\{f_n\}$ é uma seqüência de Cauchy então dado um $\varepsilon>0$, $\exists$ $N(\varepsilon)$ tal que

$$\| f_n-f_m\| < \varepsilon, \quad \forall n,m > N(\varepsilon)$$

De

$$\| f_n-f_m\| = \max_{a\leq x\leq b} |f_n(x) - f_m(x)| < \varepsilon$$

concluimos que

$$|f_n(x) - f_m(x)| < \varepsilon$$

para cada $a\leq x\leq b$. Isto significa que para cada ponto do intervalo $[a,b]$, $\{f_n(x)\}$ é uma seqüência de Cauchy de números reais. Como o espaço dos reais é completo, $\{f_n(x)\}$ converge para um valor real $f(x)$. Logo dada uma seqüência de Cauchy $\{f_n\}$, ela converge ponto a ponto para uma função f.

b) Mostrar que a convergência é uniforme, significa mostrar que, dado um $\varepsilon>0$ existe um $N(\varepsilon)$ tal que, para todo $a\leq x\leq b$

$$\| f_n(x)-f(x)\| < \varepsilon \quad \text{para } n\geq N(\varepsilon)$$

Isto equivale a mostrar que

$$\sup_{a\leq x\leq b} |f_n(x) - f(x)| \leq \varepsilon \quad \forall n>N(\varepsilon)$$

Se a igualdade acima não fosse verdadeira existiria um $\varepsilon>0$ tal que para todo N teríamos

$$\sup_{a\leq x\leq b} |f_n(x) - f(x)| > \varepsilon$$

bastando que $n>N$. Nesse caso então, existiria um ponto $u\in[a,b]$, tal que

$$\varepsilon < |f_n(u) - f(u)| \leq \sup_{a\leq x\leq b} |f_n(x) - f(x)|$$

para $n>N$.

Isto não está correto, pois estamos afirmando que $\{f_n(u)\}$ não converge para $f(u)$, o que contraria a convergência ponto a ponto provada no item a).

c) Para mostrar que f é uma função contínua basta mostrar que para qualquer $\varepsilon>0$, $\exists\,\delta>0$, tal que

$$|x-\bar{x}| < \delta \Rightarrow |f(x) - f(\bar{x})| < \varepsilon$$

Seja dado um $\varepsilon > 0$:

$$|f(x)-f(\bar{x})| = |f(x)-f_n(x)+f_n(x)-f_n(\bar{x})+f_n(\bar{x})-f(\bar{x})|$$

$$\leq |f(x)-f_n(x)| + |f_n(x)-f_n(\bar{x})| + |f_n(\bar{x})-f(\bar{x})|$$

Como $\{f_n(x)\}$ converge uniformemente para $f(x)$ poderemos escolher um n suficientemente grande, tal que

$$|f(x) - f_n(x)| < \varepsilon/3 \quad \text{e}$$

$$|f_n(\bar{x}) - f(\bar{x})| < \varepsilon/3$$

Como f_n é contínua $\forall n$, dado $\varepsilon/3 > 0$ podemos escolher o ponto $\bar{x}$ tal que

$$|\bar{x} - x| < \delta$$

o que implica em

$$|f_n(x) - f_n(\bar{x})| < \varepsilon/3 \quad \text{para certo } n.$$

Logo, para dado $\varepsilon>0$ podemos tomar tal $\delta>0$ e temos

$$|f(x) - f(\bar{x})| < \varepsilon \quad \text{se}$$

$$|x - \bar{x}| < \delta$$

Por conseguinte f é uma função contínua.

Exemplo 3:

O espaço ℓ^p, $1 \leq p \leq \infty$ é um espaço de Banach.

Seja $\{x_n\}$ uma seqüência de Cauchy em ℓ^p. Se dispusermos os termos de $\{x_n\} = \{x_1^n, x_2^n, \ldots\ldots\}$, $n = 1,2,\ldots$ segundo o arranjo

$$\begin{bmatrix} x_1^1 & x_2^1 & x_3^1 & \cdots \\ x_1^2 & x_2^2 & x_3^2 & \cdots \\ x_1^3 & x_2^3 & x_3^3 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

o conjunto das componentes de uma coluna k qualquer, constitui uma seqüência de Cauchy de números reais $\{x_k^n\}_{n=1,2,\ldots}$ pois

$$|x_k^n - x_k^m| = (|x_k^n - x_k^m|^p)^{1/p} \leq \| x_n - x_m \|$$

Como o espaço dos reais é completo, $\{x_k^n\}_{n=1,2,\ldots}$ converge para um limite $x_k, \forall k$.

Vamos mostrar agora que $\{x_n\}$ converge para $x = \{x_1, x_2, \ldots\}$ na norma de ℓ^p.

Dado $\varepsilon > 0$, $\exists\, N(\varepsilon)$, tal que

$$\| x_n - x_m \| < \varepsilon \quad , \forall n,m > N(\varepsilon)$$

logo

$$\left(\sum_{i=1}^{k} |x_i^n - x_i^m|^p \right)^{1/p} \leq \| x_n - x_m \| < \varepsilon$$

para $n,m > N(\varepsilon)$ e $\forall k$

Façamos $m \to \infty$ mantendo n e k fixos. Temos, pela convergência de $\{x_i^n\}$ para x_i:

$$\left(\sum_{i=1}^{k} |x_i^n - x_i|^p \right)^{1/p} \leq \varepsilon \quad , \quad \forall n > N(\varepsilon)$$

Logo como o resultado vale para todo k, fazendo este tender ao infinito obtemos:

$$\| x_n - x \| \leq \varepsilon \qquad \forall n > N(\varepsilon)$$

e portanto

$$x_n \to x$$

Resta-nos ainda provar que $x \in \ell^p$.

De acordo com a Proposição 2, se $\{x_n\}$ é de Cauchy,

$$\left(\sum_{i=1}^{k} |x_i^n|^p \right)^{1/p} \leq \| x_n \| \leq M \qquad \forall k \text{ e } \forall n$$

Como $\displaystyle \lim_{n\to\infty} \sum_{i=1}^{k} |x_i^n|^p = \sum_{i=1}^{k} \lim_{n\to\infty} |x_i^n|^p = \sum_{i=1}^{k} |x_i|^p$

temos

$$\sum_{i=1}^{k} |x_i|^p \leq M^p \qquad \forall k$$

Como a desigualdade é válida para todo k, segue-se por passagem ao limite que:

$$\sum_{i=1}^{\infty} |x_i|^p \leq M^p$$

e portanto $x \in \ell^p$.

<u>Proposição 3</u>:

Se X e Y são espaços de Banach então o espaço produto $X \times Y$ é também um espaço de Banach com a norma do produto

$$\| (x,y) \| = \| x \| + \| y \|$$

onde: $(x,y) \in X \times Y$

$x \in X$, $y \in Y$

<u>Demonstração:</u>

Seja $\{z_n\}$ uma seqüência de Cauchy de $X \times Y$, onde $z_n = (x_n, y_n)$, $x_n \in X$ e $y_n \in Y$.

$$\| z_n - z_m \| = \| (x_n, y_n) - (x_m, y_m) \|$$

$$= \| (x_n - x_m, y_n - y_m) \|$$

$$= \| x_n - x_m \| + \| y_n - y_m \|$$

Então, dado $\varepsilon > 0$, $\exists$ $N(\varepsilon)$ tal que

$$\| x_n - x_m \| + \| y_n - y_m \| < \varepsilon \ ; \ \forall n, m > N(\varepsilon)$$

logo

$$\| x_n - x_m \| < \varepsilon \ , \ \forall n, m > N(\varepsilon)$$

e

$$\| y_n - y_m \| < \varepsilon \ , \ \forall n, m > N(\varepsilon)$$

Portanto $\{x_n\}$ e $\{y_n\}$ são seqüências de Cauchy e como X e Y são espaços de Banach

$$\{x_n\} \to x \in X$$

$$\{y_n\} \to y \in Y$$

Seja agora $z \in X \times Y$, tal que, $z = (x, y)$.

Temos:

$$\| (x_n, y_n) - (x, y) \| = \| x_n - x \| + \| y_n - y \|$$

Como $\| x_n - x \| \to 0$ e $\| y_n - y \| \to 0$ então

$$\| (x_n, y_n) - (x, y) \| \to 0$$

portanto

$$\{z_n\} \to z \in X \times Y.$$

c.q.d.

Exemplo 4:

Como consequência direta da Proposição 3 podemos concluir que $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^n$ são espaços de Banach com a norma do produto. (Na verdade, como já afirmamos, o são também para qualquer outra norma, o que será provado no Teorema 2).

Assim como a seqüência de racionais $\{1;1,4;1,41;1,414,...\}$ tem como limite o irracional $\sqrt{2}$, seqüências de Cauchy de alguns espaços não convergem para um elemento pertencente ao espaço em questão.

Os dois exemplos seguintes ilustram este fato.

Exemplo 5:

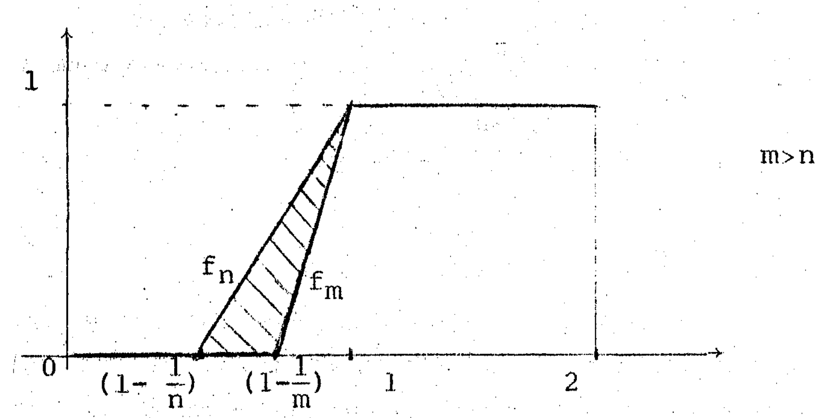
O espaço das funções contínuas $C^0[a,b]$ com a norma

$$\|f\| = \int_a^b |f(x)|dx \quad \text{não é completo}$$

Mostraremos que a seqüência $\{f_n\}$ de elementos de $C^0[0,2]$ definida por

$$f_n(x) = \begin{cases} 0 & 0 \leq x \leq 1 - \frac{1}{n} \\ nx-n+1 & 1 - \frac{1}{n} < x < 1 \\ 1 & 1 \leq x \leq 2 \end{cases}$$

converge para $f \notin C^0[0,2]$ com a norma acima definida.



$\{f_n\}$ é uma seqüência de Cauchy, senão vejamos:

$$\| f_n - f_m \| = \int_0^2 |f_n(t) - f_m(t)|dt$$

Essa integral é igual à área hachureada na figura, ou seja:

$$\| f_n - f_m \| = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{m}\right).$$

Agora $\| f_n - f_m \| \to 0$ quando n e m tendem para infinito. Logo $\{f_n\}$ é de Cauchy.

Mostremos agora que o limite de $\{f_n\}$ é uma função f não contínua.

Nesse caso temos que

$$\lim_{n\to\infty} \| f_n - f \| = 0$$

Porém

$$\| f_n - f \| = \int_0^2 |f_n(x) - f(x)|dx$$

$$= \int_0^{1-\frac{1}{n}} |f(x)|dx + \int_{1-\frac{1}{n}}^{1} |f_n(x)-f(x)|dx + \int_0^2 |1-f(x)|dx$$

Logo:

$$\lim_{n\to\infty} \| f_n - f \| = \int_0^1 |f(x)|dx + \int_1^2 |1-f(x)|dx = 0$$

Então:

$$\begin{cases} |f(x)| = 0 & 0 \le x < 1 \\ |f(x)| = 1 & 1 < x \le 2 \end{cases}$$

Portanto f é uma função não contínua e o espaço das funções contínuas $C^0[a,b]$ não é completo com a norma

$$\| f \| = \int_a^b |f(x)|dx$$

<u>Exemplo 6</u>:

O espaço das seqüências infinitas com um número finito de componentes diferentes de zero, $x = \{x_1, x_2, \ldots, x_n, 0, 0 \ldots\}$ não é um espaço completo com a norma

$$\| x \| = \max_i |x_i|$$

A seqüência

$$\{x_n\} = \{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n-1}, 0, 0, \ldots\}$$

por exemplo, é uma seqüência de Cauchy pois

$$\| x_n - x_m \| = \max\{ \frac{1}{n}, \frac{1}{m}\} \to 0 \quad \text{se } n,m \to \infty$$

Por outro lado não existe nenhuma seqüência x do espaço para a qual $\{x_n\}$ converge, pois ela converge para uma seqüência x com infinitos componentes diferentes de zero.

Exemplo 7:

Pode-se demonstrar que o espaço quociente X/M, munido da norma definida no Capítulo II, é um espaço de Banach se X é um espaço de Banach e M um subespaço completo de X.

Por exemplo considerando a norma do máximo para $X = C^0[0,1]$, o espaço de classes de funções contínuas em $[0,1]$ que coincidem nos pontos 0 e 1 é um espaço de Banach. Nesse caso $M = C^0_0[0,1]$ onde

$$C^0_0[0,1] = \{x / x \in C^0[0,1],\ x(0) = x(1) = 0\}$$

$C^0_0[0,1]$ é um subespaço completo como se poderá verificar.

Na verdade o mesmo resultado vale se M é fechado, pois como se provará a seguir M ser fechado implica em M completo, quando X é completo.

III.3. SUBESPAÇOS COMPLETOS.

Definição 4: Um subespaço deum espaço vetorial normado é um subespaço completo se toda seqüência de Cauchy no subespaço converge para um limite que pertence ao subespaço.

Teorema 1: Num espaço de Banach, um subespaço é completo se e somente se ele for fechado.

Demonstração:

Seja M um subespaço de um espaço de Banach X. Provamos inicialmente que se M é completo então ele é fechado. Se M é completo, toda seqüência de Cauchy de elementos de M converge para um elemento que pertence a M. Em particular toda seqüência de M que seja convergente, também converge em M porque toda seqüência convergente é de Cauchy.

Agora provamos: se M é fechado então M é completo. Seja $\{x_m\}$ uma seqüência de Cauchy de M.

$\{x_n\}$ é uma seqüência de Cauchy de X e como X é um espaço de Banach então $x_n \to x \in X$.

Por outro lado M é um subespaço fechado e toda seqüência convergente de M tem seu limite em M, logo $x \in M$ e M é completo. c.q.d.

Teorema 2: Em um espaço vetorial normado todo subespaço de dimensão finita é completo.

Demonstração:

Seja M um subespaço de dimensão N de um espaço vetorial X. Demonstraremos que M é completo por indução finita.

a) Para N = 1

Se $x \in M$ então $x = \alpha e$ onde e = base de M e α = escalar.

Seja $\{x_n\}$ uma seqüência de Cauchy de M. Então para cada elemento x_n temos $x_n = \alpha_n e$

Logo $\| x_n - x_m \| = \| e \| \cdot |\alpha_n - \alpha_m|$

Portanto a convergência de $\{x_n\}$ se torna equivalente à convergência dos escalares $\{\alpha_n\}$. Como os reais formam um espaço completo, decorre diretamente que M é completo quando tem dimensão N=1.

b) Vamos supor que o resultado é válido para subespaços de dimensão N-1 e demonstrar para a dimensão N.

Seja $\{e_1, e_2, \ldots, e_N\}$ uma base de M.

Se $x \in M$ então $x = \sum_{i=1}^{N} \alpha_i e_i$

ou ainda $x = \alpha_k e_k + \sum_{\substack{i=1 \\ i \neq k}}^{N} \alpha_i e_i$

Vamos definir para cada k o subespaço M_k de dimensão N-1 gerado por $\{e_i\}_{i=1,\ldots,N}^{i \neq k}$

Seja agora

$$\delta_k = \inf_{y \in M_k} \| e_k - y \|$$

Podemos interpretar δ_k como sendo a "distância" do vetor e_k ao subespaço M_k.

Agora temos $\delta_k > 0$, $\forall k$, pois se $\delta_k = 0$, poder-se-ia construir uma seqüência de vetores de M_k convergindo para $e_k \notin M_k$. Tal seqüência não pode existir porque M_k, de dimensão N-1, é completo por hipótese.

Seja

$$\delta = \min_{1 \leq k \leq N} \delta_k > 0 \quad \text{e}$$

$\{x_n\}$ uma seqüência de Cauchy em M onde cada têrmo é da forma:

$$x_n = \sum_{i=1}^{N} \alpha_i^n e_i .$$

Temos que:

$$\| x_n - x_m \| = \| \sum_{i=1}^{N} (\alpha_i^n - \alpha_i^m) e_i \| , \quad \forall n,m$$

Supondo-se que $\alpha_k^n \neq \alpha_k^m$ vem

$$\| x_n - x_m \| = \| e_k - \sum_{\substack{i=1 \\ i \neq k}}^{N} \frac{\alpha_i^n - \alpha_i^m}{\alpha_k^n - \alpha_k^m} \cdot e_i \| \cdot | \alpha_k^n - \alpha_k^m |$$

$$\geq \delta_k | \alpha_k^n - \alpha_k^m |$$

ou seja

$$(1) \qquad \boxed{\| x_n - x_m \| \geq \delta . |\alpha_k^n - \alpha_k^m|} \qquad \forall k$$

Se $\alpha_k^n = \alpha_k^m$ então

$$\| x_n - x_m \| \geq 0 \quad \text{e}$$

$$\delta |\alpha_k^n - \alpha_k^m| = 0$$

logo a relação (1) ainda continuará válida.

Portando para cada $1 \leq k \leq N$

$$\| x_n - x_m \| \geq \delta |\alpha_k^n - \alpha_k^m| \quad \text{e}$$

$$\| x_n - x_m \| \to 0 \implies |\alpha_k^n - \alpha_k^m| \to 0, \quad \forall k$$

Assim para cada k, $\{\alpha_k^n\}$ é uma seqüência de Cauchy de escalares e portanto converge para um escalar α_k.

Agora se

$$x = \sum_{k=1}^{N} \alpha_k e_k \in M, \text{ temos que}$$

$$\| x_n - x \| = \| \sum_{k=1}^{N} (\alpha_k^n - \alpha_k) e_k \|$$

$$\leq N . \max_{1 \leq k \leq N} \{ |\alpha_k^n - \alpha_k| . \| e_k \| \}$$

Como $|\alpha_k^n - \alpha_k| \to 0$ para todo k, concluimos que $\| x_n - x \| \to 0$; então $\{x_n\} \to x \in M$.

c.q.d.

III.4. ESPAÇO PRÉ-HILBERTIANO

Antes de introduzirmos uma classe importante de espaços completos - os de Hilbert - faremos um parênteses no estudo de tais espaços para introduzir alguns conceitos fundamentais relacionados com o produto escalar ou produto interno, que nada mais é do que um produto convenientemente definido de dois vetores, cujo valor é um escalar.

Denominaremos um espaço vetorial, Pré-Hilbertiano se tiver uma estrutura de espaço vetorial normado com a norma oriunda de um produto escalar.

Tal estrutura nos permitirá não só guardar o conceito de distância normalmente associado a uma norma, como também servirá para introduzir outros conceitos geométricos, possibilitando-nos pensar muitas vezes em termos menos abstratos, com o uso de nossa intuição geométrica habitual. Por exemplo, um dos conceitos importantes presentes nos espaços Pré-Hilbertianos é o da ortogonalidade que é a base de algumas operações fundamentais nesses espaços, tal como a projeção ortogonal.

<u>Definição 5</u>: Uma aplicação de $X \times X$ em $\mathbb{R}$, isto é, que associa a cada par ordenado (x,y) de elementos de X, um escalar $(x|y)$, é um produto escalar ou produto interno se ela satisfaz aos seguintes axiomas:

(I) $(x|y) = (y|x)$ (simetria)

(II) $(x+z|y) = (x|y) + (z|y)$ (linearidade)*

(III) $(\alpha x|y) = \alpha(x|y)$, α escalar

(IV) $(x|x) \geq 0$ e $(x|x) = 0 \iff x = \vec{0}$ (positividade)

(*) (II)e(III) podem ser resumidos em

$$(\alpha x+\beta z|y) = \alpha(x|y) + \beta(z|y) \quad , \quad \alpha,\beta \text{ escalares}$$

<u>Definição 6</u>: Chamamos de espaço Pré-Hilbertiano a um espaço vetorial X munido de um produto escalar.

Vamos ver mais adiante (Teorema 3) que um espaço Pré-Hilbertiano X se torna um espaço vetorial normado com a norma definida por

$$||x|| = \sqrt{(x|x)}$$

<u>Exemplo 8</u>

O espaço $\mathbb{R}^n$ constituido de n-uplas de números reais é um espaço Pré-Hilbertiano com o produto escalar do vetor

$$x = (x_1, x_2, \ldots\ldots, x_n)$$

com o vetor

$$y = (y_1, y_2, \ldots\ldots, y_n)$$

definido da seguinte forma:

$$(x|y) = \sum_{i=1}^{n} x_i y_i$$

De fato temos:

<u>Axioma I</u>:

$$(y|x) = \sum_{i=1}^{n} y_i x_i = \sum_{i=1}^{n} x_i y_i = (x|y)$$

<u>Axioma II e III</u>

$$(\alpha x+\beta z| y) = \sum_{i=1}^{n} (\alpha x_i+\beta z_i) y_i = \sum_{i=1}^{n} (\alpha x_i y_i+\beta z_i y_i)$$

$$= \sum_{i=1}^{n} \alpha x_i y_i + \sum_{i=1}^{n} \beta z_i y_i, \text{ ou seja}$$

$$(\alpha x+\beta z|y) = \alpha(x|y) + \beta(z|y)$$

<u>Axioma IV</u>

Se $x = \vec{0}$, $x_i = 0 \quad i = 1,2,\ldots,n$ e logo $(x|x) = 0$

Se $(x|x) = 0$ então $\sum_{i=1}^{n} x_i^2 = 0 \Rightarrow x_i = 0$, $i = 1,\ldots,n$

ou seja $x = \vec{0}$

Neste espaço a norma definida como $\sqrt{(x|x)}$ é denominada norma Euclidiana

$$||x|| = \left[\sum_{i=1}^{n} | x_i |^2 \right]^{1/2}$$

<u>Exemplo 9</u>

Sejam $x = \{x_1, x_2, \ldots\ldots\}$ e

$y = \{y_1, y_2, \ldots\ldots\}$

dois vetores arbitrários do espaço ℓ^2. Se definirmos um produto escalar da forma

$$(x|y) = \sum_{i=1}^{\infty} x_i y_i$$

o espaço ℓ^2 se torna um espaço Pré-Hilbertiano.

Devemos notar que o produto interno acima definido só assume valores finitos pois a desigualdade de Cauchy-Scwartz (Capítulo II) garante que

$$|(x|y)| \leq || x || . || y || ,$$

sendo que nesse caso

$|| x|| = \sqrt{(x|x)}$ é a norma usual do espaço ℓ^2.

Exemplo 10

O espaço das funções de valores reais contínuas definidas no intervalo $[a,b]$ é um espaço Pré-Hilbertiano com o produto escalar da forma

$$(x|y) = \int_a^b x(t)y(z)dt$$

Exemplo 11

O espaço $\mathcal{P}[a,b]$ dos polinômios definidos no intervalo $[a,b]$ é também um espaço Pré-Hilbertiano com o produto escalar:

$$(x|y) = \int_a^b x(t)y(t)dt$$

Vamos agora provar, como já foi anunciado, que um espaço pré-Hilbertiano está automaticamente munido de uma norma, oriunda do produto escalar. Inicialmente precisamos do

Lema 1

(Desigualdade de Cauchy-Schwarz)

Quaisquer que sejam os vetores x e y de um espaço Pré-Hilbertiano X vale

$$|(x|y)| \leq ||x|| \cdot ||y||$$

Sendo:

$$||x|| = \sqrt{(x|x)} \quad \text{e}$$

$$||y|| = \sqrt{(y|y)}$$

Demonstração

Se $y = \vec{0}$ a desigualdade é satisfeita trivialmente.

Consideremos $y \neq \vec{0}$ e a expressão

$$\phi(\alpha) = (x-\alpha y \mid x-\alpha y)$$

onde α é um escalar arbitrário.

Devido à linearidade do produto escalar temos:

$$\phi(\alpha) = (x|x) - 2\alpha(x|y) + \alpha^2 (y|y)$$

Agora, pela positividade do produto escalar (axioma IV) devemos ter $\phi(\alpha) \geq 0$, $\forall \alpha$. Mas como $\phi(\alpha)$ é um polinômio do 2º grau, isso só ocorrerá se seu discriminante for não positivo, isto é:

$$(x|y)^2 - (x|x)(y|y) \leq 0$$

e portanto:

$$|(x|y)| \leq \|x\| \cdot \|y\|$$

c.q.d.

Teorema 3

Em um espaço Pré-Hilbertiano X, a expressão

$$\| x\| = \sqrt{(x|x)}$$

define uma norma.

Demonstração

Sejam x e y dois vetores arbitrários de X. Temos por definição:

$$\| x + y\|^2 = (x+y|x+y)$$

$$= (x|x) + 2(x|y) + (y|y)$$

$$\leq \| x\|^2 + 2|(x|y)| + \| y\|^2 .$$

Aplicando o Lema 1 temos:

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

$$\leq (\|x\| + \|y\|)^2.$$

portanto

$$\|x+y\| \leq \|x\| + \|y\|$$

Os axiomas de positividade e homogeneidade da norma decorrem diretamente dos axiomas de positividade e linearidade do produto escalar, respectivamente.

c.q.d.

Proposição 4

Se em um espaço Pré-Hilbetiano X

$$(x|y) = 0$$

para todo $y \in X$ então $x = \vec{0}$

Demonstração

Se $(x|y) = 0$ para todo $y \in X$, podemos tomar, por exemplo, $y = x$. Temos então:

$$(x|x) = 0$$

o que implica que $x = \vec{0}$

c.q.d.

Proposição 5

(Lei do Paralelogramo)

Se x e y são dois vetores arbitrários de um espaço Pré-Hilbertiano X, então:

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

Demonstração:

Temos:

$$||x+y||^2 + ||x-y||^2 = (x+y|x+y) + (x-y|x-y)$$

Desenvolvendo vem:

$$||x+y||^2 + ||x-y||^2 = 2(x|x) + 2(y|y)$$
$$= 2||x||^2 + 2||y||^2$$

c.q.d.

Definição 7: O ângulo entre dois vetores x e y de um espaço Pré-Hilbertiano é um ângulo ϕ tal que, $0 \leq \phi \leq \pi$ e

$$\cos \phi = \frac{(x|y)}{||x|| \cdot ||y||}$$

Definição 8: Dois vetores x e y de um espaço Pré-Hilbertiano são dito serem ortogonais se

$$(x|y) = 0$$

Esse fato é representado por $x \perp y$.

Observe-se que se $x \perp y$ então $\cos \phi = 0 \Rightarrow \phi = \pi/2$.

Definição 9: No espaço Pré-Hilbertiano um vetor x é dito ser ortogonal a um conjunto S de vetores se $x \perp s$ para cada $s \in S$, o que será representado por $x \perp S$. Note-se que o vetor nulo é ortogonal a qualquer vetor.

Teorema 4

Em um espaço Pré-Hilbertiano, se $x \perp y$, então

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

Demonstração

$$\|x+y\|^2 = (x+y|x+y)$$

$$= (x|x) + 2(x|y) + (y|y)$$

$$= (x|x) + (y|y)$$

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

c.q.d.

Observação

O Teorema 4 é conhecido como Teorema de Pitágoras e pode ser generalizado para o caso de um número finito de vetores, de um espaço Pré-Hilbertiano, isto é, se $x_1, x_2, \ldots, x_n$ são ortogonais dois a dois e $y = x_1 + x_2 + \ldots + x_n$, então

$$\|y\|^2 = \|x_1\|^2 + \|x_2\|^2 + \ldots + \|x_n\|^2$$

Definição 10: Em um espaço Pré-Hilbertiano um conjunto S de vetores é denominado um conjunto ortogonal se $x \perp y$ para cada par $(x,y) \in S$ com $x \neq y$.

O conjunto S é dito ser ortonormal se além de ser ortogonal, cada vetor do conjunto tem norma igual a um.

Proposição 6

Um conjunto ortogonal de vetores não nulos é um conjunto de vetores linearmente independentes.

Demonstração

Seja $\{x_1, x_2, \ldots, x_n\}$ um conjunto ortogonal de vetores não nulos e seja $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ um conjunto de escalares tal que

$$\sum_{i=1}^{n} \alpha_i x_i = \vec{0}$$

Seja x_r um vetor arbitrário do conjunto. Temos:

$$\left(\sum_{i=1}^{n} \alpha_i x_i \middle| x_r\right) = \sum_{i=1}^{n} \alpha_i (x_i | x_r)$$

Logo $$\left(\sum_{i=1}^{n} \alpha_i x_i \middle| x_r\right) = \alpha_r (x_r | x_r)$$

Porém $$\left(\sum_{i=1}^{n} \alpha_i x_i \middle| x_r\right) = (\vec{0} | x_r) = 0$$

Daí, como por hipóteses $x_r \neq \vec{0}$ concluimos que $\alpha_r = 0$.

Como isto é válido para todo r, $1 \leq r \leq n$, temos que

$$\sum_{i=1}^{n} \alpha_i x_i = \vec{0} \Rightarrow \alpha_1 = \alpha_2 = \ldots \alpha_n = 0$$

Portanto $x_1, x_2, \ldots, x_n$ são linearmente independente

c.q.d.

Consideremos agora o problema de construir conjuntos ortogonais de vetores num dado espaço vetorial. Em particular essa construção pode ser realizada a partir de um conjunto de vetores linearmente independentes de acordo com o

Teorema 5

Seja $\{x_1, x_2, \ldots, x_n\}$ um conjunto finito* de vetores linearmente independentes de um espaço pré-Hibertiano X. Então existem coeficientes a_{jk}, tais que o conjunto de vetores $\{z_1, z_2, \ldots, z_n\}$ assim definido:

* Ou, mais geralmente, um conjunto livre enumerável.

$$z_1 = x_1$$

$$z_2 = a_{21} z_1 + x_2$$

$$z_3 = a_{31} z_1 + a_{32} z_2 + x_3$$

$$\vdots \qquad \vdots$$

$$z_n = a_{n1} z_1 + \ldots + a_{n,n-1} z_{n-1} + x_n$$

seja ortogonal.

Demonstração

Demonstraremos o teorema por indução.

Para n=2 basta tomarmos

$$a_{12} = \frac{-(x_2|z_1)}{\|x_1\|^2}$$

para se ter $z_1 \perp z_2$.

Agora supondo que n-1 vetores ortogonais $\{z_1, z_2, \ldots, z_{n-1}\}$ tenham sido encontrados calcularemos $z_n \neq 0$ satisfazendo a n-ésima equação acima, com z_n ortogonal aos vetores $z_1, z_2, \ldots, z_{n-1}$

Seja

$$\boxed{z_n = a_{n1} z_1 + a_{n2} z_2 + \ldots + a_{n,n-1} z_{n-1} + x_n} \tag{2}$$

onde:

$a_{n1}, a_{n2}, \ldots, a_{n,n-1}$ são os coeficientes que desejamos determinar.

Para dado r, r<n, fazemos o produto escalar $(z_n|z_r)$. Considerando a ortogonalidade de $\{z_1, z_2, \ldots, z_{n-1}\}$ temos:

$$(z_n|z_r) = a_{nr}(z_r|z_r) + (x_n|z_r)$$

Como desejamos $z_n \perp z_r$, $\forall r < n$, temos que:

$$(3) \qquad \boxed{a_{nr} = - \frac{(x_n|z_r)}{(z_r|z_r)}} \qquad r = 1,2,\ldots,n-1$$

substituindo (3) em (2) encontramos

$$z_n = a_{n1}z_1 + a_{n2}z_2 + \ldots + a_{n,n-1}z_{n-1} + x_n$$

sendo z_n não nulo porque, como se poderá verificar, é uma combinação linear não trivial dos vetores $x_1, x_2, \ldots, x_n$, que são linearmente independentes. Além disso, z_n é ortogonal a cada um dos vetores $z_1, z_2, \ldots, z_{n-1}$.

c.q.d.

O resultado do Teorema 5 dá margem à instituição de um processo de ortogonalização de vetores conhecido como a <u>ortogonalização de Gram-Schmidt.</u> .

Seja $\{x_1, x_2, \ldots, x_n, \ldots\}$ um conjunto livre de vetores , finito ou enumerável, de um espaço Pré-Hilbertiano X.

O processo de ortogonalização de Gram Schmidt consiste em encontrar uma seqüência ortonormal de vetores $\{e_i\}$, tal que, para cada inteiro n, o espaço gerado pelos n primeiros e_i's é o mesmo espaço gerado pelos n primeiros x_i's. Em suma, para cada n temos:

$$[e_1, e_2, \ldots, e_n] = [x_1, x_2, \ldots, x_n]$$

O processo é a aplicação direta do <u>Teorema 5</u> e consiste das seguintes etapas:

Façamos inicialmente

$$1. \quad e_1 = \frac{x_1}{\|x_1\|}$$

Agora, para cada $r = 2,3,\ldots,n$, calculamos

$$2. \quad z_r = x_r - \sum_{i=1}^{r-1} (x_r|e_i)e_i$$

$$3. \quad e_r = \frac{z_r}{\|z_r\|}$$

Podemos notar que

$$e_r, z_r \perp e_i, \ \forall i, \ i<r, \quad r = 2,3,\ldots,n,\ldots$$

Como os e_i's, para dado n, são linearmente independentes por serem vetores ortogonais dois a dois e como são, em última análise, combinações lineares não triviais de

$\{x_1, x_2, \ldots, x_n\}$, o subespaço por eles gerado

$[e_1, e_2, \ldots, e_n]$ coincide com $[x_1, x_2, \ldots, x_n]$.

O processo de ortogonalização de Gram-Schmidt é freqüentemente usado para a construção de bases ortogonais de espaços de dimensão finita. Por exemplo, podemos ortonormalizar a base canônica $\{1, x, x^2\}$ de $\mathcal{P}^2[-1,1]$ pelo processo de Gram- Schimidt obtendo a base $\{e_1, e_2, e_3\}$ com:

$$e_1(x) = \frac{\sqrt{2}}{2}$$

$$e_2(x) = \frac{\sqrt{6}}{2}\, x$$

$$e_3(x) = \frac{\sqrt{10}}{4}\,(3x^2-1)$$

III.5. ESPAÇO DE HILBERT

Como anunciado no parágrafo anterior, trataremos agora os espaços munidos de produto escalar que sejam completos para a norma a ele associada. Consideraremos daqui para a frente como um espaço Pré-Hilbertiano, um espaço vetorial normado com tal norma.

Definição 11: Um espaço Pré-Hilbertiano completo é chamado um espaço de Hilbert.

Um problema clássico que consiste numa primeira aplicação dos conceitos apresentados é a procura da distância mínima de certo elemento x de um espaço vetorial X a um subconjunto M de X. Em outras palavras, procuramos um vetor $m_0 \in M$ que minimize

$$\| x-m_0 \| \leq \| x-m \| \quad , \ \forall m \in M$$

A distância mínima de x a M é então $\| x-m_0 \|$.

Quando M é um subconjunto de um espaço Pré-Hilbertiano X, se existir a solução m_0, veremos que é possível caracterizá-la de modo bem simples, com a introdução do conceito de projeção. Entretanto a existência e a unicidade de tal vetor não estará em geral garantida, a menos que formulemos hipóteses mais restritivas sôbre M e X .

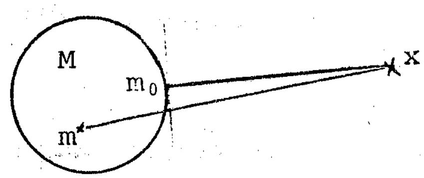
Por exemplo, um resultado importante nesse quadro, é o Teorema Clássico da Projeção, que postula que se M é um subespaço fechado de um espaço de Hilbert X, poderemos não só afirmar a existência do vetor minimizante m_0 como também sua unicidade.

No entanto, por razões de ordem didática preferimos dar inicialmente um resultado correlato englobando o Teorema Clássico da Projeção que poderá assim ser deduzido como uma conseqüência imediata.

Teorema 6: (Teorema da Projeção Generalizada)

Seja M um subconjunto convexo fechado de um espaço de Hilbert X. Se x é um vetor arbitrário de X, existe um único vetor $m_0 \in M$, dito projeção de x sôbre M, tal que:

$$\| x - m_0 \| \leq \| x - m \| \quad ; \forall m \in M$$



Além disso, a condição necessária e suficiente para que m_0 seja o único vetor minimizante é que

$$(x - m_0 | m - m_0) \leq 0 \quad , \forall m \in M$$

Demonstração:

Para provar a existência de m , vamos considerar uma seqüência $\{m_i\}$ de M, tal que:

$$\| x - m_i \| \to \delta = \inf_{m \in M} \| x - m \|$$

Pela lei do paralelograma temos

$$\| (m_j - x) + (x - m_i) \|^2 + \| (m_j - x) - (x - m_i) \|^2 =$$

$$= 2 \| m_j - x \|^2 + 2 \| x - m_i \|^2$$

Logo:

$$\| m_j - m_i \|^2 = 2\| m_j - x\|^2 + 2\| x - m_i \|^2 - 4\| x - \frac{m_i + m_j}{2} \|^2$$

Sendo M um convexo, $(m_i+m_j)/2 \in M$ e portanto

$$\| x - \frac{m_i + m_j}{2} \| \geq \delta \ .$$

Temos então que

$$\| m_j - m_i \|^2 \leq 2\| m_j - x\|^2 + 2 \| x - m_i \|^2 - 4\delta^2 .$$

Como $\| m_i - x\|^2 \to \delta^2$ concluimos que

? , e a demonstração fecha isto!

$$\| m_j - m_i \| \to 0$$

ou seja, a seqüência $\{m_i\}$ é de Cauchy em X e portanto converge para $m_0 \in X$, dado que X é de Hilbert. Agora como M é por hipótese fechado, então $m_0 \in M$.

Enfim, se $\{m_i\} \to m_0$ então

$$\| x - m_i \| \to \| x - m_0 \| = \delta$$

Para provar a unicidade de m_0, vamos supor que exista $m_1 \in M$ com

$$\| x - m_1 \| = \delta$$

A seqüência $\{m_n\}$ definida por

$$m_n = m_0 \quad \text{para n par}$$
$$m_n = m_1 \quad \text{para n ímpar}$$

é tal que $\| x - m_n \| \to \delta$

Entretanto, pela análise de existência acima, $\{m_n\}$ é de Cauchy e convergente. Ora, isto só é possível se

$$m_1 = m_0$$

Enfim tratamos do problema de caracterizar m_0, o único vetor minimizante, por

$$(x-m \;|m-m_0) \leq 0, \; \forall m \in M$$

Contrariando essa hipótese, vamos supor que exista um vetor $m_1 \in M$ tal que

$$(x-m_0 \,|m_1 - m_0) = \varepsilon > 0$$

Como M é convexo, para $0 \leq \alpha \leq 1$, temos:

$$m_\alpha = (1-\alpha) m_0 + \alpha m_1 \in M.$$

Consideremos a função:

$$\phi(\alpha) = \| x - m_\alpha \|^2 = \| (1-\alpha)(x-m_0) + \alpha (x-m_1) \|^2$$

$$= (1-\alpha)^2 \| x-m_0 \|^2 + 2\alpha(1-\alpha)(x-m_0 \,|x-m_1) + \alpha^2 \| x-m_1 \|^2$$

Tomemos agora a derivada de $\phi(\alpha)$ no ponto $\alpha = 0$:

$$\frac{d\phi}{d\alpha}\Big/_{\alpha=0} = -2 \| x-m_0 \|^2 + 2(x-m_0 \,|x-m_1)$$

$$= -2(x-m_0 \,|m_1-m_0)$$

$$= -2\,\varepsilon < 0$$

Portanto, nesse caso existiria um α tal que

$$\| x - m_\alpha \| < \| x - m_0 \|$$

o que é absurdo, por hipótese.

Agora suponhamos que $m_0 \in M$ é tal que

$$(x - m_0 \,| m - m_0) \leq 0 \;, \; \forall m \in M$$

Teríamos então $\forall m \in M$

$$\| x-m \|^2 = \| x-m_0 + m_0 - m \|^2$$

$$= \| x-m_0 \|^2 + 2(x-m_0 | m_0 - m) + \| m_0 - m \|^2$$

Supondo agora que $m \neq m_0$, vem:

$$\| x-m \|^2 > \| x-m_0 \|^2 .$$

m_0 é portanto vetor minimizante.

c.q.d.

Uma observação importante que é possivel ser feita com base no último argumento usado no Teorema 6 é a seguinte:

Mesmo que não tivéssemos garantia de existência do vetor que minimizasse a distância de $x \in X$, espaço Pré-Hilbertiano, a um seu subconjunto M (que não satisfaça a nenhuma condição particular), se encontrássemos $m_0 \in M$, tal que

$$(x-m_0 | m-m_0) \leq 0 \quad , \quad \forall m \in M$$

então m_0 seria vetor minimizante da distância de x a qualquer vetor de M.

<u>Teorema 7</u>: (Teorema Clássico de Projeção)

Seja M um subespaço fechado de um espaço de Hilbert X. Se x é um vetor arbitrário de X, então existe um único vetor $m_0 \in M$ dito <u>projeção ortogonal de x sobre M</u>, tal que:

$$\| x-m_0 \| \leq \| x-m \| \quad , \quad \forall m \in M$$

Além disso a condição necessária e suficiente para que m_0 seja esse único vetor é que $x-m_0$ seja ortogonal a M.

<u>Demonstração</u>:

A existência e a unicidade de m_0 decorre diretamente do teorema anterior tendo-se em conta que um subespaço fechado é um convexo fechado.

Vamos mostrar somente que m_0 é único, se e somente se, $x-m_0$ for ortogonal a M.

a) A condição é necessária:

De fato pelo Teorema 6 se m_0 é o vetor minimizante

$$(x-m_0|m-m_0) \leq 0, \quad \forall m \in M$$

Como M é subespaço, então podemos tomar

$$m = n + m_0$$

para certo n arbitrariamente escolhido em M.

Temos:

$$(x-m_0|n) \leq 0, \quad \forall n \in M$$

Ainda por ser M subespaço teríamos:

$(x-m_0|-n) \leq 0, \ \forall n \in M$, o que dá enfim:

$(x-m_0|n) = 0, \quad \forall n \in M$, isto é,

$$x - m_0 \perp M$$

b) Se $x-m_0 \perp M$ então m_0 é o vetor minimizante.

Seja $m_1 \in M$ tal que $m_1 \neq m_0$.

Temos pelo Teorema de Pitágoras:

$$\| x- m_1 \|^2 = \| x-m_0 + m_0-m_1 \|^2$$

$$= \| x-m_0 \|^2 + \| m_0-m_1 \|^2$$

Logo:

$$\| x-m_1 \| > \| x-m_0 \| \qquad \forall m_1 \neq m_0$$

c.q.d.

Observação:

Se no Teorema da Projeção o espaço considerado é um espaço Pré-Hilbertiano (não completo), a existência do vetor minimizante m_0 não pode ser afirmada. Porém se ele existe, então ele é a projeção ortogonal do vetor x sôbre o subespaço fechado M. A demonstração desse fato se processa de forma análogo à utilizada no Teorema 7.

Vejamos agora algumas propriedades adicionais dos espaços de Hilbert:

Definição 12: Seja M um subconjunto de um espaço de Hilbert X. Chamamos de complemento ortogonal de M ao conjunto de todos os vetores de X que sejam ortogonais a M e o denotaremos por $M^{\perp}$.

Se o conjunto M consiste apenas do vetor nulo, o complemento ortogonal $M^{\perp}$ de M é todo o espaço X.

Proposição 7:

Sejam M e N subconjuntos de um espaço de Hilbert. Então valem as propriedades:

1º) $M^{\perp}$ é um subespaço fechado

2º) $M \subset M^{\perp\perp}$

3º) Se $M \subset N$, então $N^{\perp} \subset M^{\perp}$

4º) $M^{\perp\perp\perp} = M^{\perp}$

Demonstração:

1º) $M^{\perp}$ é um subespaço, pois a combinação linear de vetores ortogonais a M é também ortogonal a M.

Seja $x_n \to x$ uma seqüência convergente de $M^{\perp}$. Temos:

$$(x_n|m) = 0, \ \forall m \in M \text{ e } \forall n, \text{ dondé,}$$

$$(x_n - x + x|m) = (x_n - x|m) + (x|m) = 0$$

Logo, como $x_n \to x$ e $(x_n - x|m) \leq ||x_n - x|| \, ||m||$ temos $(x|m) = 0$, $\forall m \in M$ e então $x \in M^{\perp}$, isto é, $M^{\perp}$ é um subespaço fechado.

2º) Se $x \in M$ então $x \perp M^{\perp}$. Mas se $x \perp M^{\perp}$ então $x \in M^{\perp\perp}$.
Logo, se $x \in M$, $x \in M^{\perp\perp}$, ou seja, $M \subset M^{\perp\perp}$

3º) Se $y \in N^{\perp}$ então $y \perp N$. Mas como $M \subset N$, então $y \perp M$. Logo, se $y \in N^{\perp}$ então $y \in M^{\perp}$, ou seja $M \subset N \Rightarrow N^{\perp} \subset M^{\perp}$.

4º) Da 2º propriedade temos que $M \subset M^{\perp\perp}$ e também que $M^{\perp} \subset M^{\perp\perp\perp}$. Porém a 3º propriedade implica que $M^{\perp\perp\perp} \subset M^{\perp}$.
Logo $M^{\perp\perp\perp} = M^{\perp}$

c.q.d.

<u>Definição 11</u>: Um espaço vetorial X é a <u>soma direta</u> de dois subespaços M e N se cada vetor $x \in X$ tem uma <u>única</u> representação $x = m+n$ onde $m \in M$ e $n \in N$. Denotaremos isto por $X = M \oplus N$. (Devemos notar que a diferença entre soma direta e a soma usual definida no Capítulo I é a unicidade da representação agora imposta)(ver Exemplo I.19).

<u>Teorema 8</u>:

Se M é um subespaço fechado deum espaço de Hilbert X então $X = M \oplus M^{\perp}$.

<u>Demonstração</u>:

Seja x um vetor arbitrário de X. Pelo teorema da projeção, existe um único vetor $m_0 \in M$, tal que,

$$\| x-m_0 \| \leq \| x-m \| , \ \forall m \in M \quad e$$

$$n_0 = x - m_0 \in M^{\perp}.$$

Temos então que

(4) $\boxed{x = n_0 + m_0}$, com $m_0 \in M$ e $n_0 \in M^{\perp}$.

Por outro lado, a representação (4) é única, pois caso contrário teríamos

(5) $\boxed{x = m_1 + n_1}$

com $m_1 \in M$ e $n_1 \in M^{\perp}$.

Subtraindo (5) de (4) vem

$\vec{0} = m_1 - m_0 + n_1 - n_0$.

Como M e $M^{\perp}$ são subespaços,

$m_1 - m_0 \in M$ e
$n_1 - n_0 \in M^{\perp}$

Agora pelo teorema de Pitágoras temos:

$|| \vec{0} ||^2 = || m_1 - m_0 ||^2 + || n_1 - n_0 ||^2 = 0$

Portanto $m_1 = m_0$ e $n_1 = n_0$

c.q.d.

Teorema 9:

Se M é um subespaço fechado de um espaço de Hilbert X, então $M = M^{\perp\perp}$

Demonstração:

Seja x um vetor qualquer de $M^{\perp\perp}$. Pelo Teorema 8 temos:

$x = m + n$ com $m \in M$ e $n \in M^{\perp}$

Porém, pela Proposição 7, 2º propriedade, temos que:

$m \in M$ então $m \in M^{\perp\perp}$

Logo, $x - m \in M^{\perp\perp}$

Porém, como $x-m = n$, então, $n \in M^{\perp\perp}$. Mas como $n \in M^{\perp}$, então $n \perp n$, o que implica que $n = \vec{0}$.

Dessa forma $x=m$ e $M^{\perp\perp} \subset M$ o que pela Proposição 7, 2ª propriedade, implica em:

$M = M^{\perp\perp}$ c.q.d.

III.6. APROXIMAÇÃO

Vamos agora tratar com mais detalhes o problema de determinar o(s) vetor(es) minimizante(s) da distância de $x \in X$, sendo X um espaço vetorial normado, a um subespaço M de X.

A importância desse assunto pode ser bem ilustrada, por exemplo, pelo clássico problema de aproximar uma função f do espaço de dimensão infinita $C^0[a,b]$ por um polinômio m_0 de um espaço de dimensão finita $\mathcal{P}^n[a,b]$ que representará f para efeitos práticos e fazendo o erro $e = \| f-m_0 \|$ se tornar tão pequeno quanto possível. Nesse caso o que ganhamos é a simplicidade inerente aos espaços de dimensão finita.

De acordo com a natureza do espaço X e do subespaço M, teremos diferentes resultados, o que mostraremos a seguir usando, quando necessário, exemplos ilustrativos:

1º) X é de Hilbert e M é fechado
2º) X é de Hilbert e M não é fechado
3º) X é Pré-Hilbertiano e M é fechado
4º) X é Pré-Hilbertiano e M é completo
5º) X é de Banach (sem produto escalar) e M é fechado
6º) X é de Banach e M de dimensão finita

1º) X é de Hilbert e M é fechado

O Teorema Clássico da Projeção Ortogonal, afirma a existência e a unicidade do vetor minimizante m_0. Nesse caso m_0 é a projeção ortogonal do vetor x sobre o subespaço fechado M.

2º) X é de Hilbert e M não é fechado

Na demonstração do teorema clássico da projeção, construimos uma seqüência $\{m_i\}$ de elementos de M e mostramos que se esta seqüência faz com que $\| x-m_i \|$ tenda para um mínimo, então ela é de Cauchy. Como X é de Hilbert $\{m_i\}$ tem seu limite $m_0 \in X$.

No caso em questão poderemos chegar até essa situação. Porém como M não é fechado, não poderemos afirmar que m_0 pertencerá a M. Logo nem sempre a solução existe.

Típico dentro dessa categoria é o problema de calcular a distância mínima de um vetor de um espaço de Hilbert X a um subespaço M denso em X, $M \neq X$. Por definição M não é fechado, logo, se tomarmos $x \in X$, com $x \notin M$, a distância de x a qualquer vetor de m pode ser feita arbitrariamente pequena, isto é,

$$\inf_{m \in M} \| x-m \| = 0 \quad .$$

No entanto tal ínfimo não é atingido por nenhum elemento de M, ou seja, o problema determinar o vetor minimizante m_0 não tem solução.

3º) X é Pré-Hilbertiano e M é fechado

O vetor minimizante nesse caso nem sempre existe. Mostraremos isso através de um contra-exemplo.

Exemplo 12:

a) Consideremos o espaço Pré-Hilbertiano X das funções contínuas em [-1,1] com o produto escalar

$$(f|g) = \int_{-1}^{1} fg dx$$

X não é completo para a norma correspondente o que se pode provar de forma análoga à do Exemplo 5.

b) Consideremos M o subespaço de X para o qual

$$\int_{-1}^{0} f dx = \int_{0}^{1} f dx \qquad \forall f \in M$$

Mostraremos inicialmente que M é fechado com a norma

$$\| f \| = [\int_{-1}^{1} f^2 dx]^{1/2}.$$

Seja $\{f_n\} \subset M$ e f uma função tal que

$$\| f_n - f \| = \int_{-1}^{1} |f_n - f|^2 dx \to 0$$

Temos então que

$$\left[\int_{-1}^{1} |f_n - f|^2 dx\right]^{1/2} \to 0$$

Logo
$$\int_{0}^{1} |f_n - f|^2 \, dx \to 0 \quad \text{e} \quad \int_{-1}^{0} |f_n - f|^2 \, dx \to 0$$

Por outro lado sabemos que

$$\left[\int_{0}^{1} (f_n - f)\, dx\right]^2 \leq \left[\int_{0}^{1} |f_n - f|\, dx\right]^2$$

e pela desigualdade de Schwartz:

$$\leq \int_{0}^{1} |f_n - f|^2 dx$$

Logo, como o 2º membro da desigualdade acima tende a zero temos:

$$\int_{0}^{1} (f_n - f)\, dx \to 0, \text{ ou seja}$$

$$\int_{0}^{1} f_n dx \to \int_{0}^{1} f dx$$

Da mesma forma poderíamos provar que

$$\int_{-1}^{0} f_n dx \to \int_{-1}^{0} f dx$$

Como
$$\int_{0}^{1} f_n dx = \int_{-1}^{0} f_n dx \quad \forall n,$$

pela unicidade do limite vem:

$$\int_{-1}^{0} f dx = \int_{0}^{1} f dx$$

e portanto $f \in M$.

c) Tomaremos agora a função $f(x) = x$ que pertence a X, mas não a M_1.

Temos $\left(\int_{0}^{1} x dx \neq \int_{-1}^{0} x dx\right)$ e mostraremos que a função minimizante f_0 não existe.

Vamos mostrar inicialmente que

$$|f(x)-x| \geq 1/2 \qquad \forall f \in M \quad e \ -1 \leq x \leq 1$$

ou, equivalentemente, que $|f(x)-x| < 1/2$ implica em absurdo.

Temos:

$$-1/2 < f(x)-x < 1/2$$

$$x-1/2 < f(x) < x+1/2$$

$$\int_0^1 (x-1/2)\,dx < \int_0^1 f(x)\,dx < \int_0^1 (x+1/2)\,dx$$

$$0 < \int_0^1 f(x)\,dx < 1$$

Analogamente obteríamos:

$$-1 < \int_{-1}^0 f(x)\,dx < 0$$

Como $f \in M$ temos por hipótese que

$$\int_0^1 f(x)\,dx = \int_{-1}^0 f(x)\,dx$$

donde concluimos o absurdo.

Portanto devemos ter:

$$|f(x) - x| \geq 1/2 \qquad \forall f(x) \in M$$

d) Se $f_0 \in M$ é o vetor minimizante, vimos que ele é a projeção ortogonal de $f \in X$ sobre M, ou seja

$$(f - f_0|g) = 0 \qquad \forall g \in M$$

No caso em pauta onde $f(x) = x$ temos

$$\int_{-1}^1 [x-f_0(x)]g(x)\,dx = 0 \qquad \forall g \in M.$$

Tomemos então $g(x) = 1$, $\forall x$. Temos claramente $g \in M$, donde concluimos que

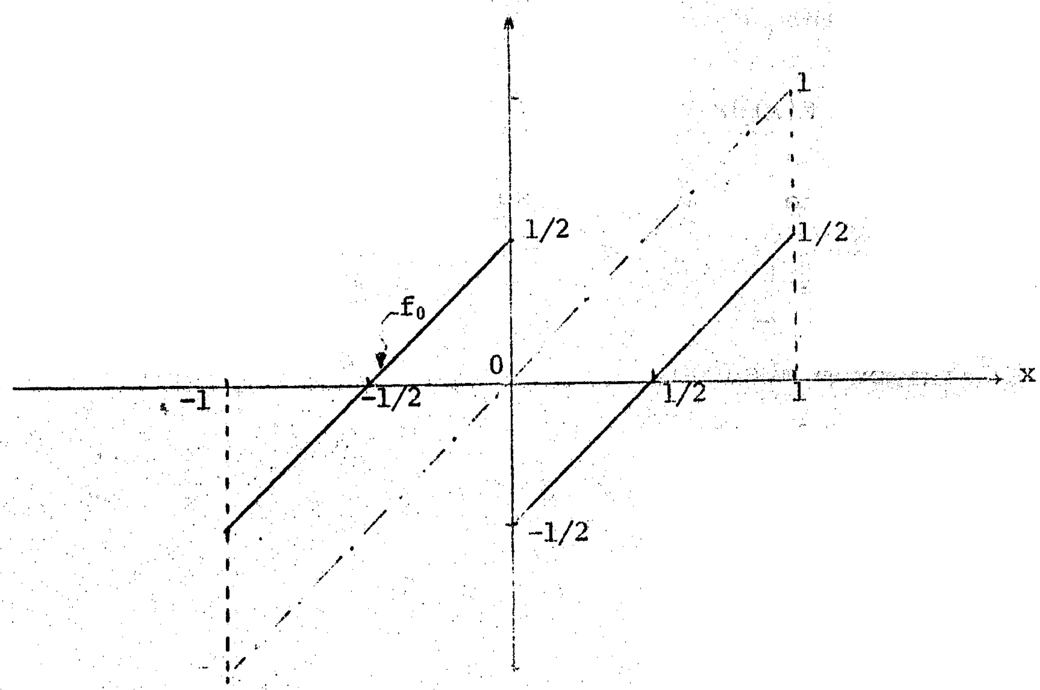
$$\int_{-1}^{1} f_0(x)\,dx = \int_{-1}^{1} x\,dx = 0$$

Como
$$\int_{-1}^{1} f_0(x)\,dx = 2\int_{0}^{1} f_0(x)\,dx = 2\int_{-1}^{0} f_0(x)\,dx$$

concluimos que:

$$\int_{-1}^{0} f_0(x)\,dx = 0 \quad \text{e} \quad \int_{0}^{1} f_0(x)\,dx = 0$$

Ora, a única função f_0 que satisfaz às condições acima e tal que
$|f_0(x)-x| \geq \frac{1}{2}$ $\forall x$ $[-1,1]$ é a função ilustrada abaixo,



que não pertence a M por não ser contínua.

Logo o problema não tem solução.

4º) X é Pré-Hilbertiano e M é completo

Nesse caso o vetor minimizante $m_0 \in M$ existe e é único. A demonstração é quase a mesma do Teorema clássico da Projeção, com a única diferença de que a existência de m, limite da seqüência de Cauchy $\{m_i\}$, pode ser confirmada pelo fato de M ser completo nesse caso.

Por exemplo para qualquer função contínua $f \in C^0[a,b]$ existe um polinômio de grau máximo n $p_0 \in \mathcal{P}^n[a,b]$ que minimiza, $\forall p \in \mathcal{P}^n[a,b]$ a distância:

$$\| f-p \| = \left[\int_a^b |f(x) - p(x)|^2 dx \right]^{1/2}$$

De fato como qualquer subespaço de dimensão finita é completo temos o resultado, apesar de $C^0[a,b]$ não ser completo para essa norma.

5º) X é de Banach (sem produto escalar) e M é fechado

O vetor minimizante nesse caso nem sempre existe. Para mostrar isto, vamos exibir um exemplo:

Exemplo 13:

a) Consideremos o espaço $C^0[-1,1]$ que é (completo) de Banach com a norma do máximo.

b) Consideremos, como no 3º caso, o subespaço M como sendo o conjunto das funções f de $C^0[-1,1]$ tais que

$$\int_{-1}^{0} f(x)dx = \int_0^1 f(x)dx$$

Mostraremos que M definido dessa forma é fechado.

Seja $\{f_n\}$ uma seqüência de funções de M e vamos supor que ela converge para uma função $f \in X$. Temos

$$\left| \int_{-1}^0 f_n(x)dx - \int_{-1}^0 f(x)dx \right| \leq \int_{-1}^0 |f_n(x) - f(x)| dx$$

$$\leq \max_{x \in [-1,1]} |f_n(x) - f(x)| \cdot 1$$

$$\leq \| f_n(x) - f(x) \| \to 0$$

logo: $\displaystyle\int_{-1}^{0} f_n(x)dx \to \int_{-1}^{0} f\ (x)dx.$

Analogamente provaríamos que:

$$\int_{0}^{1} f_n(x)dx \to \int_{0}^{1} f\ (x)dx$$

Porém, como $\forall n$, f_n pertence a M, temos:

$$\int_{-1}^{0} f_n(x)dx = \int_{0}^{1} f_n(x)dx\ ,\ \forall n$$

Então pela unicidade do limite temos que:

$$\int_{-1}^{0} f\ (x)dx = \int_{0}^{1} f\ (x)dx$$

donde concluimos que $f \in M$. Logo M é fechado.

c) Tomemos agora a função $f(x) = x$, $f \notin M$, e mostraremos que a função f_0 que minimiza

$$\| f-g \|\ ,\quad \forall g \in M$$

não pertence a M.

d) No 3º caso, já provamos que

$$|g(x)-x| \geq 1/2\ ,\quad \forall g \in M \text{ e } -1 \leq x \leq 1$$

Procuremos pois f_0 tal que

$$\| f_0 - f \| = \max_{x \in [-1,1]} |f_0(x)-x| = 1/2$$

e) Para tal função f_0 valeria também

$$\int_{-1}^{0} f_0(x)dx = \int_{0}^{1} f_0(x)dx$$

ou seja, não é possível ser a distância mínima igual a 1/2, pois nesse caso, f_0 seria a mesma função não contínua ilustrada no 3º caso.

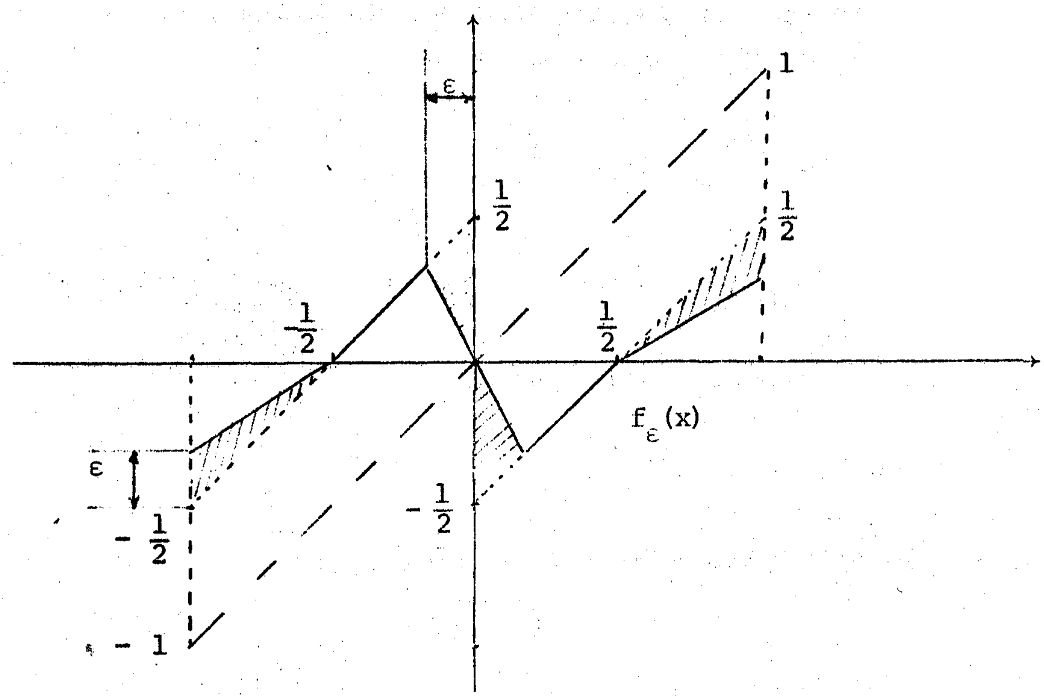
Por outro lado, dado $\varepsilon > 0$, é sempre possível encontrar uma função f_ε, contínua, com

$$\| f_\varepsilon - f \| = \max_{x \in [-1,1]} |f_\varepsilon(x) - x| = \frac{1}{2} + \varepsilon$$

e obedecendo à relação

$$\int_{-1}^{0} f_\varepsilon(x)\,dx = \int_{0}^{1} f_\varepsilon(x)\,dx$$

como mostrado na figura abaixo



Podemos pois construir uma seqüência de funções $\{f_\varepsilon\} \subset M$ com valores de ε cada vez menores ($\varepsilon \to 0$) ou seja, a distancia mínima de $f(x) = x$ a M pode ser feita tão próxima de 1/2 quanto quisermos. Porém esse valor ínfimo não é atingido e o problema de minimização não tem solução neste caso.

6º) <u>X é de Banach e M de dimensão finita</u>

Provaremos que nesse caso existe o vetor minimizante $m_0 \in M$, mas que ele não é necessariamente único.

Seja $\{e_1, e_2, \ldots, e_n\}$ base de M. Se $x \in M$ a solução é trivial; se $x \notin M$, o problema pode ser enunciado da seguinte forma:

Seja

$$\delta = \inf \| x - \sum_{j=1}^{n} \alpha_j e_j \| , \quad \alpha_j \in \mathbb{R}$$

ou seja:

$$\delta = \inf_{m \in M} \| x-m \| \leq \| x-m \| , \ \forall m \in M.$$

Existirá um vetor $m_0 \in M$ tal que

$$\delta = \| x - m_0 \| \ ?$$

De acordo com as propriedades do ínfimo, existe uma seqüência $\{y_i\}$ de M tal que

$$\delta_i = \| x-y_i \| \to \delta$$

com

$$\delta_i \geq \delta_{i+1} , \ \forall i.$$

Consideremos o subespaço

$$W = [e_1, e_2, \ldots, e_n, x]$$

de dimensão n+1 (pois $x \notin M$) e a bola $\overline{B(0,\delta_1)}$ de X.

Seja $w_i = y_i - x \in \overline{B(0,\delta_1)}$, $\forall i$

$\overline{B(0,\delta_1)}$ é fechado limitado de W espaço de dimensão finita, logo é um compacto. Podemos extrair de $\{w_i\}$ uma subseqüência $\{w_{i_k}\}$ convergente em W, à qual corresponde uma sub-seqüência $\{y_{i_k}\}$ com $w_{i_k} = x-y_{i_k}$ que terá a mesma propriedade.

De fato sendo w o limite de $\{w_{i_k}\}$ temos:

$$y = \lim_{k\to\infty} y_{i_k} = w-x$$

Temos então $y \in M$, dado que M é fechado.

Agora provaremos que

$$\lim_{k\to\infty} \| y_{i_k} - x \| = \delta$$

De fato, caso contrário, existiria $\varepsilon>0$, tal que, $\forall K$, teríamos:

$$\| y_{i_k} - x \| > \delta+\varepsilon \ , \ \forall k > K$$

Mas, se isso ocorre para tal ε, não seria possível encontrar N, tal que

$$\| y_i - x \| - \delta < \varepsilon \ , \ \forall i>N$$

pois, para todo N, poderíamos sempre encontrar $i_k>N$, $k>K$, com

$$\| y_{i_k} - x \| - \delta > \varepsilon$$

o que contraria a hipótese $\| y_i - x \| \to \delta$

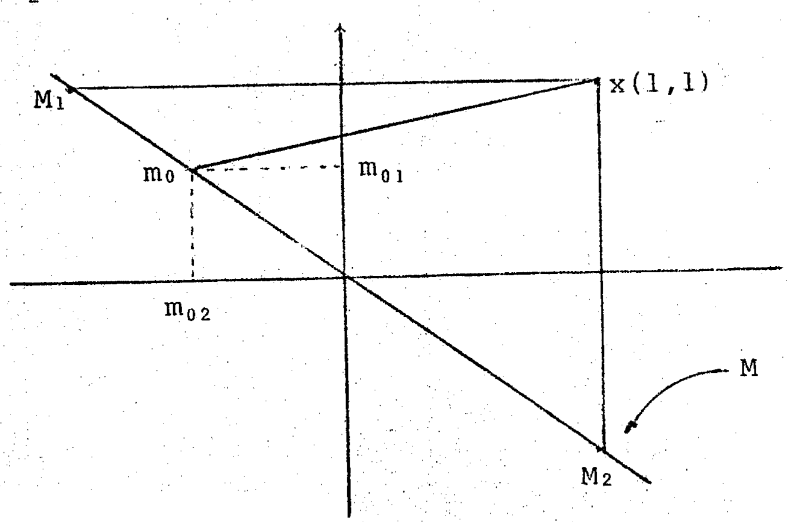
Portanto y é o vetor minimizante, ou seja, $m_0 = y$.

<u>Exemplo 14</u>:

Consideremos o espaço de Banach $X = \mathbb{R}^2$ com a norma de $x = (x_1, x_2) \in X$ definida como

$$\| x \| = |x_1| + |x_2|$$

Tomemos como subespaço M, a reta tal que $y_1 = -y_2$, $\forall y = (y_1, y_2) \in M$



Nesse caso X é de Banach e M é um subespaço de dimensão finita.

Tomemos o ponto $x = (1,1)$, $x \notin M$.

Vamos agora procurar um ponto $m_0 \in M$ tal que

$$\| x-m_0 \| \leq \| x-m \| \quad , \forall m \in M$$

Apenas observando a figura acima poderemos concluir que todos os pontos de M compreendidos entre os pontos M_1 e M_2, satisfazem à condição acima. Além disso, nesse caso particular, a distância mínima é igual a dois.

$$\| x-m_0 \| = |1-m_{01}| + |1-m_{02}|$$

Como $\quad m_{01} = -m_{02} \quad$ e

$$|m_{01}| = |m_{02}| \leq 1$$

então

$$\| x-m_0 \| = 2 \qquad \forall m_0 \in [M_1, M_2]$$

Portanto existe uma infinidade de soluções para esse problema de minimização.

Para terminar observamos que todos os resultados válidos para o caso em que o subespaço M é fechado, que foram aqui apresentados, podem ser estendidos analogamente ao caso em que M é um convexo fechado.

CAPÍTULO IV

APLICAÇÕES LINEARES CONTÍNUAS

IV.1 - INTRODUÇÃO

O conceito de continuidade de funções de variáveis reais ou complexas introduzidas no Calculo Elementar, pode ser estendido aos espaços vetoriais munidos de estruturas adequadas. No nosso caso, a topologia associada a uma norma permitirá essa extensão de modo bastante direto, o que nos conduzirá a uma classe importante de espaços vetoriais: a das aplicações lineares contínuas, que constituem, como outros tópicos previamente aqui tratados, um poderoso instrumento para as aplicações.

Inicialmente julgamos oportuno introduzir alguns conceitos básicos relativos a aplicações em geral:

Definição 1: Sejam dois conjuntos X e Y. Uma aplicação A de X em Y é uma lei de associação que a cada elemento de X faz corresponder um único elemento de Y. Denotamos:

$$\begin{aligned} A:\ & X \to Y \\ & x \mapsto y \end{aligned}$$

sendo $y=A(x)$ o elemento de Y ao qual corresponde $x \in X$ através da lei de associação A.

Por exemplo, uma função de uma variável real nada mais é do que uma aplicação de $\mathbb{R}$ em $\mathbb{R}$ ou de um subconjunto de $\mathbb{R}$ em outro.

$$\begin{aligned} f:\ & \mathbb{R} \to \mathbb{R} \\ & x \mapsto f(x) \end{aligned}$$

O conjunto X sobre o qual a aplicação A é definida é denominado o domínio de A. Podemos ainda considerar aplicações A de X em Y, mas que só são definidas para elementos de um subconjunto de X que será nesse caso o domínio denotado $D(A)$.

O conjunto dos elementos $y \in Y$ para os quais existe $x \in X$ tal que $y=A(x)$ é dito a imagem de X sob a aplicação A, (também dito contradomínio de A) representada por $A(X)$.

Por exemplo, para a aplicação

$$\begin{aligned} A:\ & \mathbb{R}^2 \to \mathbb{R} \\ & (x,y) \to (x^2+y^2+1)^{-1} \end{aligned} \quad \text{temos:}$$

$$A(\mathbb{R}^2)=(0,1]$$

Já a função de uma variável definida por $f(x)=\sqrt{1-x}$ pode ser considerada como uma aplicação

$$f: \mathbb{R} \to \mathbb{R}$$
$$x \mapsto \sqrt{1-x}$$

tendo por domínio o intervalo não limitado $(-\infty,1]$ e por imagem $[0,\infty)$.

Uma aplicação pode ser classificada segundo a maneira global em que a correspondência se estabelece entre os elementos de X e Y.

Assim, se a imagem de uma aplicação é todo Y, ela é dita <u>sobrejetiva</u>. Sucintamente se uma aplicação A é sobrejetiva então:

$$\forall y \in Y \quad \exists\, x \in D(A) \;/\; y=A(x)$$

Se por outro lado, a cada elemento y da imagem A(X) existir um único elemento $x \in X$ tal que $y=A(x)$ então A é dita <u>injetiva</u>.

Sucintamente, se A é injetiva então:

$$x_1, x_2 \in D(A) / A(x_1)=A(x_2) \Rightarrow x_1=x_2 .$$

Por exemplo, a aplicação

$$A: \mathbb{R} \to \mathbb{R}$$
$$x \mapsto e^x$$

não é sobrejetivamas é injetiva.

Já a aplicação

$$A: \mathbb{R}^2 \to \mathbb{R}$$
$$(x,y) \mapsto xy$$

é sobrejetiva mas não é injetiva.

Uma aplicação em X de um seu subconjunto M talque:

$$I: M \to X$$
$$x \mapsto x$$

é certamente injetiva e é dita <u>injeção canônica</u> de M em X.

Quando uma aplicação é ao mesmo tempo injetiva e sobrejetiva ela é dita <u>bijetiva</u> . Por exemplo, a aplicação

$$A: \mathbb{R} \to \mathbb{R}$$
$$x \mapsto x^3$$

é bijetiva.

Se $D(A)=X$ e A é bijetiva então A é dita uma bijeção de X em Y. Nesse caso, há uma correspondência <u>um a um</u> entre os elementos de X e os de Y, ou uma <u>bijeção</u> de X em Y:

$\forall y \in Y$ existe um único $x \in X$ tal que $A(x)=y$.
$\forall x \in X$ existe um único $y \in Y$ tal que $y=A(x)$.

IV.2- <u>DEFINIÇÕES E PROPRIEDADES</u>

Estudemos agora o caso específico em que os conjuntos X e Y são espaços vetoriais.

<u>Definição 2</u> Uma aplicação A

$$A: X \to y$$

é dita ser linear se ela possuir a propriedade seguinte:

Para todo par de vetores x_1, x_2 do espaço X e par de escalares α_1, α_2, temos:

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)$$

Uma aplicação linear é ainda comumente denominada um <u>operador linear</u>.

Da Definição 2 decorre imediatamente que se A é linear

$A(\vec{0}_X) = \vec{0}_Y$ pois $\vec{0}_X = 0x \quad \forall x \in X$ e além disso

$A(0x) = 0.A(x) = \vec{0}_Y$.

O exemplo mais trivial de aplicação linear é o da <u>aplicação nula</u> (ou identicamente nula)

$$0: X \to Y$$
$$x \mapsto \vec{0}_Y$$

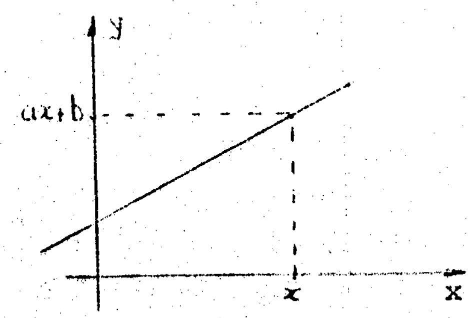
Outra aplicação linear fundamental é a <u>aplicação identidade</u>:

$$\mathcal{J}: X \to X,$$
$$x \mapsto x$$

Um exemplo simples que motiva a denominação de aplicação linear é o seguinte:

$$A: \mathbb{R} \to \mathbb{R}$$
$$x \mapsto ax+b$$

onde a e b são escalares.



<u>Exemplo 1</u> Podemos considerar uma matriz $A(m\times n)$ como um operador linear de $\mathbb{R}^n$ em $\mathbb{R}^m$

$$A: \mathbb{R}^n \to \mathbb{R}^m$$
$$x \mapsto Ax$$

De fato, de acordo com as definições clássicas de operações matriciais

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2 \quad ; \; x_1, x_2 \in \mathbb{R}^n, \; \alpha_1, \alpha_2 \text{ escalares.}$$

Inversamente qualquer aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ pode ser associada a uma matriz $A\ (m\times n)$

De fato, se A é uma aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ temos, $\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$:

$A(x) = A\left(\sum_{j=1}^{n} x_j e_j\right)$ onde $\{e_j\}_{j=1}^{n}$ é a base canônica de $\mathbb{R}^n$. Pela linearidade temos:

$$A(x) = \sum_{j=1}^{n} x_j A(e_j)$$

Por outro lado se considerarmos $\{f_i\}_{i=1}^{m}$ a base canônica de $\mathbb{R}^m$ podemos dizer que as componentes de $A(e_j) \in \mathbb{R}^m$ são escalares, a_{ij}

$$i = 1, 2, \ldots, m, \quad \forall j, \text{ ou seja}$$

$$A(e_j) = \sum_{i=1}^{m} a_{ij} \cdot f_i$$

Por conseguinte $\forall x \in X$ temos

$$A(x) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} f_i \sum_{j=1}^{n} a_{ij} x_j$$

ou seja, as componentes de $y=A(x)$ são as componentes de $\mathcal{A}x$ onde $\mathcal{A}$ é a matriz $m \times n$(*)

$$\mathcal{A} = [a_{ij}].$$

(*) Nota: Apesar de $\mathcal{A}$ matriz ser intrinsicamente de natureza distinta de A operador linear, na prática confundiremos esses dois entes matemáticos. A rigor o que temos é um isomorfismo entre o conjunto de operadores lineares de $\mathbb{R}^n$ em $\mathbb{R}^m$ e o conjunto de matrizes $m \times n$ de acordo com o que veremos a seguir.

Podemos considerar o produto AB de duas aplicações lineares A e B

$$B: X \to Y$$
$$A: Y \to Z$$

onde X,Y e Z são espaços vetoriais.

Nesse caso

$$AB: X \to Z$$

é definido por $AB(x) = A[B(x)] \quad \forall x \in X$

Por exemplo, se B é uma aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ e A de $\mathbb{R}^m$ em $\mathbb{R}^p$ então AB, aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}^p$, é a matriz $p \times n$ produto matricial de A por B (verificar).

Analogamente podemos definir produtos de n aplicações lineares $A_n, A_{n-1}, \ldots, A_2, A_1$

$A_i: X_i \to X_{i+1}$, X_i espaços vetoriais $i=1,2,\ldots,n+1$ como sendo a aplicação linear P.

$P: X_1 \to X_{n+1}$ tal que

$$P(x) = A_n[A_{n-1}[\ldots[A_2[A_1(x)]\ldots] \quad \forall x \in X_1$$

No caso particular em que $X_1 = X_2 = \ldots = X_{n+1} = X$ e $A_1 = A_2 = \ldots = A_n$ temos a aplicação potência inteira de A.

$$A^n: X \to X$$

Uma aplicação linear A bijetiva com D(A)= , tem importância particular nas aplicações sendo por isso denominada um isomorfismo dos espaços vetoriais X e Y. Assim dois espaços vetoriais serão ditos isomorfos se existir uma aplicação linear que seja uma bijeção de um no outro.

Um exemplo marcante de isomorfismo é o seguinte:

Todo espaço de dimensão n é isomorfo ao $\mathbb{R}^n$.

De fato, se X é um espaço vetorial de dimensão n, então existe uma base $\{e_i\}_{i=1}^{n}$ e cada vetor $x \in X$ pode ser escrito sob a forma

$$x = \sum_{i=1}^{n} x_i e_i$$

sendo x_i as componentes de x com respeito à base $\{e_i\}_{i=1}^{n}$. Portanto, se considerarmos a aplicação definida por

$$A: X \to \mathbb{R}^n$$
$$x \to (x_1, x_2, \ldots, x_n)$$

vemos que ela é linear e um a um

Consideremos de agora em diante que X e Y são espaços vetoriais normados.

Definição 3: Uma aplicação linear A de um espaço vetorial normado X em outro Y é dita isométrica se conserva as normas, isto é sendo $\| \cdot \|_X$ e $\| \cdot \|_Y$ as normas em X e Y, respectivamente, temos:

$$\| A(x) \|_Y = \| x \|_X \quad \forall x \in X.$$

Quando X e Y são espaços vetoriais normados isomorfos através de uma aplicação linear isométrica A, dizemos que A é um isomorfismo de espaços vetoriais normados.

Quando uma aplicação A é bijetiva, podemos definir sua inversa A^{-1} com

$$A^{-1}: Y \to X$$
$$A(x) \to x$$

isto é, a aplicação inversa de A leva cada elemento $y \in Y$ ao seu único correspondente em X através da aplicação A^{-1}.

Podemos provar a seguinte

Proposição 1: A inversa de uma aplicação linear é linear

Demonstração: Seja $y_1=A(x_1)$ e $y_2 = A(x_2)$

Temos por definição:

$$\alpha_1 A^{-1}(y_1) + \alpha_2 A^{-1}(y_2) = \alpha_1 x_1 + \alpha_2 x_2$$

Porém pela linearidade de A temos:

$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y_1 + \alpha_2 y_2$

$\alpha_1 x_1 + \alpha_2 x_2 = A^{-1}(\alpha_1 y_1 + \alpha_2 y_2)$ o que dá finalmente

$\alpha_1 A^{-1}(y_1) + \alpha_2 A^{-1}(y_2) = A^{-1}(\alpha_1 y_1 + \alpha_2 y_2)$ como desejado

c.q.d.

Se uma aplicação linear A é injetiva temos

$$A(x)=\vec{0}_Y \Rightarrow x=\vec{0}_X$$

No entanto quando A é qualquer podemos ter $x \neq \vec{0}$ tais que:

$$A(x)=\vec{0}_Y$$

O conjunto de tais vetores de X constitui um subespaço de X, como se poderá verificar sem dificuldades, denominado o espaço nulo de A e denotado por $\mathcal{N}(A)$. Por exemplo, se A é uma matriz (n×n) singular de rank n-1 aplicando $\mathbb{R}^n$ em $\mathbb{R}^n$, seu espaço nulo é o subespaço gerado pelo autovetor unitário associado ao seu autovalor nulo. Nesse caso A não é bijetiva.

Inúmeros exemplos de aplicações lineares serão considerados no decorrer deste Capítulo. Vamos agora definir uma classe importante de aplicações:

Definição 4: Uma aplicação de um espaço vetorial X no corpo de escalares é dita um funcional.

Um exemplo simples de funcional é a aplicação norma que definimos no Capítulo II.

$$\begin{aligned} \|\cdot\| : X &\to \mathbb{R} \\ x &\mapsto \|x\| \end{aligned}$$

A aplicação norma não é um funcional linear como se pode facilmente verificar.

Os funcionais lineares representam importante papel em Análise Funcional, especialmente se tiverem a propriedade de continuidade que enunciaremos a seguir para aplicações quaisquer:

Definição 5: Sejam dois espaços vetoriais normados X e Y. Uma aplicação A:

$$A: X \to Y$$
$$x \mapsto y$$

será dita contínua num ponto $x_o \in X$ se, dado um $\varepsilon > 0$ arbitrário, existir $\delta > 0$ dependendo de ε, eventualmente de x_o, tal que:

$$\| x - x_o \|_X < \delta \Rightarrow \| A(x) - A(x_o) \|_Y < \varepsilon$$

Mais sucintamente, a continuidade de A em x_o significa que para toda bola $B(y_o, \varepsilon)$ de Y com $y_o = A(x_o)$, existe uma bola $B(x_o, \delta)$ de X tal que $x \in B(x_o, \delta) \Rightarrow A(x) \in B(y_o, \varepsilon)$.

Uma aplicação A será dita contínua se for contínua em todos os pontos de seu domínio.

Um exemplo dos mais triviais de aplicação contínua é a norma:

$$\| \ \|: X \to \mathbb{R}$$
$$x \mapsto \| x \|$$

De fato, como a norma de $\mathbb{R}$ é o módulo, dado $\varepsilon > 0$, podemos tomar $\delta = \varepsilon$ e temos:

$$\| x - x_o \| < \delta \Rightarrow | \ \| x \| - \| x_o \| \ | < \varepsilon \quad \forall x_o \in X,$$ uma vez que

$$\| x \| - \| x_o \| \leq \| x - x_o \|.$$

Exemplo 2: A aplicação

$$\int : C^o[a,b] \to \mathbb{R}$$
$$f \mapsto \int_a^b f(x)dx$$

é um funcional linear contínuo sobre $C^o[a,b]$.

De fato admitindo conhecida a linearidade da aplicação, temos:

$$\left| \int_b^a f(x)dx - \int_a^b f_o(x)dx \right| = \left| \int_a^b [f(x) - f_o(x)]dx \right|$$

$$\leq \int_a^b \max_{x\in[a,b]} |f(x) - f_o(x)|dx$$

$$= |b-a| \; \| f-f_o \|$$

Logo $\forall\varepsilon>0$ basta tomar $\delta=\varepsilon/(b-a)$ para termos:

$$\| f-f_o \|<\delta \Rightarrow \left|\int_a^b f(x)dx - \int_a^b f_o(x)dx\right|< \varepsilon, \quad \forall f,f_o\in C^o[a,b]$$

Podemos caracterizar aplicações contínuas de uma maneira por vezes bastante útil que é a seguinte:

<u>Proposição 2</u>: A é uma aplicação de X em Y contínua em $x_o\in X$ se e somente se, para toda sequência $\{x_n\}\subset X$ convergindo para x_o, tivermos $\{A(x_n)\}\subset Y$ convergindo para $A(x_o)\in Y$.

<u>Demonstração</u>: Suponhamos inicialmente que A seja contínua em x_o.

Se existisse uma seqüência $\{x_n\}\subset X$ convergindo para $x_o\in X$ para a qual $A(x_n) \not\to A(x_o)$ teríamos:

$$\exists\varepsilon>0 \text{ tal que } \forall N \quad \| A(x_n)-A(x_o) \|_Y\geq\varepsilon \quad \forall n>N.$$

Porém, como $x_n \to x$, $\forall\ \delta>0$ temos um N tal que

$$\| x_n-x \|_X < \delta \qquad \forall n>N$$

Ora, nesse caso para tal ε não seria possível encontrar $\delta>0$ tal que $\| x-x_o \|_X < \delta \Rightarrow \| A(x)-A(x_o) \|_Y <\varepsilon$, pois para qualquer δ escolhido teríamos uma infinidade de termos da sequência $\{x_n\}$ que não satisfariam a essa condição, o que está em contradição com a continuidade de A.

Suponhamos agora que para toda seqüência $\{x_n\}$ com $x_n \to x_o$ em X tenhamos $A(x_n) \to A(x_o)$ em Y.

Imaginemos que nessa circunstância A não seja contínua.

Nesse caso, dado $\varepsilon>0$ teremos:

Para a bola de Y de centro em $y_o=A(x_o)$ e raio ε, temos que:

$$\forall\delta>0 \quad \exists x\in B(x_o,\delta) \text{ tal que } A(x)\notin B(y_o,\varepsilon)$$

Como δ é arbitrário, podemos construir para δ sucessivamente igual a 1, 1/2, 1/3,..., 1/n,... uma seqüência $\{x_n\}\subset X$ tal que:

$$x_n\in B(x_o,\frac{1}{n}) \quad \text{com} \quad A(x_n)\notin B(y_o,\varepsilon)$$

Claramente temos $x_n \to x_o$ em X mas $A(x_n) \not\to A(x_o)$ em Y, o que contraria a hipótese.

c.q.d.

No Exemplo **2** vimos uma aplicação linear (funcional) contínua em todos os pontos. Na verdade podemos agora estabelecer o resultado mais geral seguinte, para operadores lineares:

<u>Proposição 3</u>: Se uma aplicação linear A é contínua em dado ponto $x_o\in X$ então ela é contínua sobre todo X.

<u>Demonstração</u>: Seja $x\in X$ e uma seqüência arbitrária $\{x_n\}\subset X$ tal que $x_n \to x$.

$$\text{Temos:} \quad A(x_n) - A(x) = A(x_n) - A(x) + A(x_o) - A(x_o)$$
$$= A(x_n-x+x_o) - A(x_o)$$

Claramente temos $x_n-x+x_o \to x_o$. Logo como A é contínua em x_o então $A(x_n-x+x_o) \to A(x_o)$, donde se conclui que

$$A(x_n) \to A(x) \quad \text{quando } x_n \to x.$$

Pela Proposição **2** fica então estabelecida a continuidade de A sobre todo X.

c.q.d.

Vamos agora introduzir o conceito de limitação para aplicações que nos permitirá caracterizar de modo marcante as aplicações lineares contínuas.

Definição 6: Uma aplicação de um espaço vetorial normado X em outro Y é dita limitada se existir uma constante C tal que

$$\| A(x) \|_Y \leq C \| x \|_X \quad \forall x \in X.$$

Por exemplo, a aplicação do Exemplo 2 é limitada com $C=b-a$.

Agora vamos provar o resultado de caracterização:

Teorema 1: Uma aplicação linear é contínua se e somente se for limitada.

Demonstração: Suponhamos inicialmente que A seja limitada. Temos para uma seqüência $\{x_n\}$ arbitrária de X convergindo para $x \in X$:

$$\| A(x_n)-A(x) \|_Y = \| A(x_n-x) \|_Y \leq C \| x_n-x \|_X$$

Logo $A(x_n) \to A(x)$ pois $\| x_n-x \| \to 0$ e A é contínua.

Suponhamos agora que A seja contínua. Em particular o será em $x_o=\vec{0}$.

Então, dado $\varepsilon>0$, $\exists \delta>0$ tal que

$$\| x \|_X < \delta \Rightarrow \| A(x) \|_Y < \varepsilon,$$

Porém, pela linearidade de A e com $0<\delta'<\delta$, temos:

$$\| A(x) \|_Y = \frac{\| x \|_X}{\delta'} \| A(\frac{\delta' x}{\| x \|_X}) \|_Y$$

Como $\| \frac{\delta' x}{\| x \|_X} \|_X < \delta$ temos:

$$\| A(x) \|_Y \leq \frac{\varepsilon}{\delta'} \| x \|_X \quad \forall x \in X, \text{ ou seja}$$

A é limitada com $C = \frac{\varepsilon}{\delta'}$.

c.q.d.

Um caso particular importante de aplicações lineares contínuas é o da injeção canônica de um subespaço M de um espaço vetorial normado X, munido de norma diferente da de X. Nes se caso, se existir uma constante C tal que

$$\| x \|_X \leq C \| x \|_M \qquad \forall x \in M$$

então a <u>injeção de M em X será contínua</u>.

<u>Exemplo 3</u>: Seja X o espaço c_o com a norma

$$\| x \| = \max_i | x_i |$$

Seja M o subespaço das seqüências com um número finito de componentes não nulas normado por

$$\| x \|_M = \sum_{i=1}^{N} |x_i|$$

onde N é o maior índice de componente não nula de x.

Então a injeção de M em X é contínua pois é limitada com $C=1$.

Como o estudo das aplicações lineares contínuas será a tônica do restante deste capítulo daremos exemplos de aplicações lineares não contínuas:

<u>Exemplo 4</u>: Seja X o espaço das seqüências com um número finito de componentes não nulas com a norma:

$$\| x \| = \max_i |x_i| \qquad x=(x_1,x_2,\ldots,x_N,0,\ldots)$$

O funcional linear definido por

$$\mathfrak{F}: X \to \mathbb{R}$$
$$x \to \sum_{i=1}^{N} x_i$$

não é limitado. De fato, considerando-se a seqüência $\{e_n\}$ definida por

$$e_1 = (1,0,0,\ldots,0,\ldots,0,\ldots)$$
$$e_2 = (1,1,0,\ldots,0,\ldots,0,\ldots)$$
$$\vdots$$
$$e_n = (1,1,1,\ldots,1,\ldots,0,\ldots)$$

temos: $\| e_n \| = 1 \ \forall n$. No entanto $\mathcal{F}(e_n)$ pode assumir valores arbitrariamente grandes, o que implica na não existência de C tal que

$$|\mathcal{F}(x)| \leq C \ \| x \| \ \forall x \in X.$$

Observe-se que se X fosse normado com

$$\| x \| = \sum_{i=1}^{N} |x_i|$$

o funcional seria limitado - e portanto contínuo - com C=1.

Exemplo 5: Seja a aplicação definida por

$$D: C^0[a,b] \to C^0[a,b]$$
$$f \mapsto \frac{df}{dx}$$

O domínio de D é o subespaço de $C^0[a,b]$ das funções continuamente diferenciáveis em $[a,b]$, $C^1[a,b]$.

D não é contínua pois derivadas de funções f continuamente diferenciáveis pertencentes a um conjunto limitado, isto é, tais que

$$\| f \|_{0,\infty} = \max_{x[a,b]} |f(x)| < M$$

para certo M finito, podem assumir valores arbitrariamente grandes. Considere, por exemplo, a seqüência $\{f_n\} \subset C^0[0,1]$ definida por

$$f_n = \sqrt{x + 1/n}$$

Temos $\| f_n \|_{0,\infty} \leq \sqrt{2} \quad \forall n$. No entanto $\| \frac{df_n}{dx} \|_{0,\infty}$ não tem limitação.

Agora, se considerarmos D como aplicação de $C^1[a,b]$ em $C^0[a,b]$, sendo $C^1[a,b]$ normado por:

$$\| f \|_{1,\infty} = \max_{x \in [a,b]} |f(x)| + \max_{x \in [a,b]} |f'(x)|$$ temos uma

aplicação contínua pois:

$$\| f' \|_{0,\infty} \leq C \ \| f \|_{1,\infty} \qquad \forall \ f \in C^1[a,b] \text{ com } C=1.$$

Nos Exemplos 4 e 5 vimos que uma mesma aplicação linear sobre um espaço vetorial, pode ser contínua ou não quando consideramos diferentes normas para esse espaço. Muitas vezes, no entanto, estamos interessados em trabalhar com normas diferentes $\| \cdot \|_1$ e $\| \cdot \|_2$ que sejam equivalentes no seguinte sentido:

Existem duas constantes m e M estritamente positivas tais que $m\| x \|_1 \leq \| x \|_2 \leq M\| x \|_1 \quad \forall x \in X.$

Nesse caso temos o seguinte resultado:

Proposição 4: Se uma aplicação linear A sobre um espaço vetorial X for contínua para X munido de certa norma $\| \cdot \|_1$, o será também sobre X normado com qualquer outra norma $\| \quad \|_2$ que lhe seja equivalente, valendo também a recíproca.

Demonstração: Seja $m\| x \|_2 \leq \| x \|_1 \leq M\| x \|_2 \; \forall x \in X$ e A tal que

$$\| A(x) \| \leq C\| x \|_1 \quad \forall x \in X$$

Como $\| x \|_1 \leq M\| x \|_2$ temos

$$\| A(x) \| \leq C' \| x \|_2 \qquad \forall x \in X, \text{ com } C' = MC .$$

Por outro lado se

$$\| A(x) \| \leq C\| x \|_2 \quad \forall x \in X \quad \text{temos:}$$

$$\| A(x) \| \leq C'\| x \|_1 \; \forall \; x \in X, \text{ com } C' = \frac{C}{m}. \qquad \text{c.q.d.}$$

Pode-se provar que em espaços de dimensão finita todas as normas são equivalentes [1], o que não vale absolutamente em espaços vetoriais quaisquer. Por exemplo, sobre $C^0[a,b]$, as normas

$$\| f \|_\infty = \max_{x \in [a,b]} |f(x)|$$

e

$$\| f \|_1 = \int_a^b |f(x)| dx$$

não são equivalentes pois, se por um lado temos

$$\| f \|_1 \leq M \| f \|_\infty \qquad \forall f \in C^0[a,b] \quad \text{com } M = b-a$$

(ver Exemplo 2), não existe a constante m da desigualdade no sentido inverso. Pode-se constatar isso considerando-se, por exemplo, a seqüência $\{f_n\}$ de funções de $C^0[0,2]$ definida por

$$f_n = \frac{(n+1)x^n}{2^{n+1}}$$

Vimos que quando uma aplicação linear é um a um podemos definir a aplicação inversa A^{-1} igualmente linear e um a um. Agora se A é também contínua temos o seguinte resultado que enunciamos sem demonstração, para a qual nos referimos a [6], por exemplo.

Teorema 2 (Teorema do inverso de Banach): Se A é uma aplicação linear contínua bijeção de um espaço de Banach X em outro Y, então A^{-1} é uma aplicação linear contínua bijeção do espaço de Banach Y em X.

IV.3 - ESPAÇOS DE OPERADORES CONTÍNUOS

Vimos que $\overset{o}{C}[a,b]$, o espaço de aplicações funções f:

$$\begin{aligned} f: [a,b] &\to \mathbb{R} \\ x &\mapsto f(x) \end{aligned}$$

que sejam contínuas, é um espaço vetorial normado e completo para uma norma adequada.

Analogamente o conjunto $\mathcal{L}(X,Y)$ dos operadores (lineares) contínuos de um espaço vetorial normado X em outro Y também é espaço vetorial com as operações de adição e multiplicação escalar definidas, respectivamente, de modo análogo ao caso de funções, isto é:

Se $A,B \in \mathcal{L}(X,Y)$ então:

A+B é o operador tal que

$$(A+B)(x) = A(x)+B(x) \qquad \forall x \in X.$$

αA é o operador tal que

$$(\alpha A)(x) = \alpha A(x) \qquad \forall x \in X, \quad \alpha \text{ escalar}.$$

Claramente, se A e B são contínuos então A+B e αA também o serão pois se $x_n \to x$ em X então:

$$(A+B)(x_n) = A(x_n) + B(x_n) \to A(x) + B(x) = (A+B)(x)$$

e

$$(\alpha A)(x_n) = \alpha A(x_n) \to \alpha A(x) = (\alpha A)(x).$$

Vamos agora nos interessar pelo problema de normar $\mathcal{L}(X,Y)$. Vimos que uma aplicação linear é contínua se e somente se for limitada. É claro que a constante de limitação de um dado operador não é única, isto é, existe uma infinidade de constantes C tais que:

$$\| A(x) \| \leq C \| x \| \quad \forall x \in X$$

Agora, para $x \neq \vec{0}$ temos:

$$\frac{\| A(x) \|}{\| x \|} \leq C \quad \text{e vamos definir}$$

$$(1) \quad \boxed{\| A \| = \sup_{x \neq \vec{0}} \frac{\| A(x) \|}{\| x \|}}$$

para em seguida provar que com isso temos efetivamente uma norma para $\mathcal{L}(X,Y)$. Antes porém observamos que se C é uma constante de limitação de A temos pela própria definição de supremo:

$$\| A \| \leq C$$

Por outro lado $\| A \|$ é uma constante de limitação de A pois por definição:

$$\forall x \in X,\ x \neq \vec{0} \quad \frac{\| A(x) \|}{\| x \|} \leq \sup_{x \neq \vec{0}} \frac{\| A(x) \|}{\| x \|} \quad \Rightarrow$$

$$(2) \quad \boxed{\| A(x) \| \leq \| A \| \; \| x \|} \qquad \forall x \in X$$

A conseqüência disso é que $\| A \|$ é a menor possível dentre as constantes de limitação de A.

Uma outra observação importante é que, se encontrarmos um vetor $x_o \in X$ tal que, para uma dada constante de limitação C tenhamos:

$$\| A(x_o) \| = C \| x_o \|$$

então

$$\| A \| = \frac{\| A(x_o) \|}{\| x_o \|} = C$$

De fato, se não fosse assim, teríamos

$$\| A \| < C$$

e logo $\| A(x_o) \| > \| A \| \; \| x_o \|$, o que seria um absurdo.

Proposição 5: A expressão (1) define uma norma sobre $\mathcal{L}(X,Y)$.

Demonstração:

1º axioma: Inicialmente se $A=0$ então $\| A \| = 0$.

Por outro lado, se $\| A \|=0$ então $A=\vec{0}$ porque nesse caso $\| A(x) \| \leq 0 \quad \forall x \in X$, ou seja $A(x)=\vec{0} \quad \forall x \in X$.

2º axioma: Temos:

$$\| A+B \| = \sup_{x\neq\vec{0}} \frac{\| (A+B)(x) \|}{\| x \|} \leq \sup_{x\neq\vec{0}} \frac{\| A(x) \| + \| B(x) \|}{\| x \|}$$

$$\leq \sup_{x\neq\vec{0}} \frac{\| A(x) \|}{\| x \|} + \sup_{x\neq\vec{0}} \frac{\| B(x) \|}{\| x \|} = \| A \| + \| B \| .$$

3º axioma: Temos:

$$\| \alpha A \| = \sup_{x\neq\vec{0}} \frac{\| \alpha A(x) \|}{\| x \|} = |\alpha| \sup_{x\neq\vec{0}} \frac{\| A(x) \|}{\| x \|} = |\alpha| \; \| A \| .$$

c.q.d.

Uma conseqüência imediata de (2) é que o produto de aplicações lineares contínuas é contínuo pois $\|AB(x)\| \leq \|A\| \; \|B(x)\| \leq \|A\| \; \|B\| \; \|x\|$. Muitas vezes é interessante usar expressões equivalentes a (1) tais como:

$$(3) \quad \boxed{\| A \| = \sup_{\| x \|=1} \| A(x) \|}$$

$$(4) \quad \boxed{\| A \| = \sup_{\| x \|\leq 1} \| A(x) \|}$$

Proposição 6: (1), (3) e (4) definem a mesma norma para $\mathcal{L}(X,Y)$.

Demonstração: Provemos que (1) <=> (3)

Em primeiro lugar, por se tratar de conjunto mais restrito:

$$\sup_{\|x\|=1} \frac{\| A(x) \|}{1} \le \sup_{x\ne 0} \frac{\| A(x) \|}{\| x \|}.$$

Por outro lado, $\sup_{x\ne 0} \frac{\| A(x) \|}{\| x \|} = \sup_{x\ne 0} \| A(\frac{x}{\| x \|}) \|$

Chamando z todos os vetores da forma $\frac{x}{\| x \|}$ temos:

$$\sup_{x\ne 0} \| A(\frac{x}{\| x \|}) \| = \sup_{z} \| A(z) \|.$$

Como $\| z \|=1$ devemos ter:

$$\sup_{z} \| A(z) \| \le \sup_{\| x \|=1} \| A(x) \| \Rightarrow \sup_{x\ne 0} \frac{\| A(x) \|}{\| x \|} = \sup_{\| x \|=1} \| A(x) \|$$

Para completar a demonstração basta provar (3) <=> (4). Claramente temos:

$$\sup_{\| x \|=1} \| A(x) \| \le \sup_{\| x \|\le 1} \| A(x) \|$$

Por outro lado $\forall z\in X$ tal que $\| z \|\le 1$, $\exists\, x\in X$ tal que $\| x \|=1$ dado por $x = \frac{z}{\| z \|}$. Portanto:

$$\| A(\frac{z}{\| z \|}) \| = \| A(x) \| \Rightarrow \| A(z) \| = \| z \| \, \| A(x) \| \Rightarrow \| A(z) \| \le \| A(x) \|$$

Logo $\| A(z) \| \le \sup_{\| x \|=1} \| A(x) \| \quad \forall z\in X$ tal que $\| z \| \le 1$.

Por conseguinte

$\sup_{\| x \|\le 1} \| A(x) \| \le \sup_{\| x \|=1} \| A(x) \|$ o que dá o resultado desejado.

c.q.d.

Vejamos agora alguns exemplos:

<u>Exemplo 6</u>: Seja $X=\mathbb{R}^n$ e $Y=\mathbb{R}^m$ munidos das normas $\| \cdot \|_\infty$ respectivas. Vamos tentar obter uma expressão prática para a norma de elementos de $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ que, como sabemos, estão associados a matrizes $m\times n$.

Temos então:

$$\| A \|_\infty = \sup_{x\neq 0} \frac{\| A(x) \|_\infty}{\| x \|_\infty} = \sup_{x\neq 0} \frac{\max\limits_{1\le i\le m} | \sum_{j=1}^{n} a_{ij}x_j |}{\max\limits_{1\le i\le n} |x_i|}$$

Porém:

$$| \sum_{j=1}^{n} a_{ij}x_j | \le \sum_{j=1}^{n} |a_{ij}||x_j| \le \max_{1\le j\le n} |x_j| \sum_{j=1}^{n} |a_{ij}|$$

Logo

$$\| A \|_\infty = \sup_{x\neq 0} \frac{\max\limits_{1\le i\le m} [\max\limits_{1\le j\le n} |x_j| \sum_{j=1}^{n} |a_{ij}|]}{\max\limits_{1\le j\le n} |x_i|} \le \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|$$

Já temos uma constante de limitação de A. Agora verificamos que para um vetor particular $x_o \in \mathbb{R}^n$ temos:

$$\| Ax_o \| = [\sum_{j=1}^{n} |a_{Mj}|] \| x_o \|$$

onde M é o índice de uma linha de A cuja soma dos módulos dos elementos tem valor máximo.

Escolhendo-se pois x_o tal que

$x_{o_j} = \text{sinal } (a_{Mj}) \quad j=1,2,\ldots,n$ temos o resultado

$$\| A \|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|,$$

O leitor poderá agora verificar que, se a norma considerada para $\mathbb{R}^n$ e $\mathbb{R}^m$ fosse $\| \cdot \|_1$ teríamos:

$$\| A \|_1 = \sup_{x\neq 0} \frac{\| Ax \|_1}{\| x \|_1} = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|$$

<u>Exemplo 7</u>: Seja ainda $X=\mathbb{R}^n$ e $Y=\mathbb{R}^m$ munidos da norma euclidiana.

Nesse caso temos:

$$\| A \|_2 = \sup_{x\neq 0} \frac{\| Ax \|_2}{\| x \|_2} \quad \text{onde}$$

$$\| x \|_2 = (x,x)^{1/2} = [\sum_{i=1} x_i^2]^{1/2}$$

Sabemos que $\| Ax \|^2 = (Ax,Ax) = x^T A^T Ax = (A^T Ax,x)$ onde x^T e A^T são os transpostos de x e de A, respectivamente.

Como A^TA é uma matriz simétrica $n \times n$, positiva definida, tem um conjunto de n autovetores ortonormais [5] $\{u_i\}_{i=1}^n$ que formam uma base de $\mathbb{R}^n$. Logo $\forall x \in \mathbb{R}^n$ temos:

$x = \sum_{j=1}^n c_j u_j$ onde c_j são escalares componentes de x.

Seja agora μ_i o autovalor de A^TA associado a u_i, $i=1,2,\ldots,n$.

Temos:

$$A^TAx = \sum_{i=1}^n c_i A^TAu_i = \sum_{i=1}^n c_i \mu_i u_i$$

Logo:

$$\| Ax \|_2^2 = \left(\sum_{i=1}^n c_i \mu_i u_i, \sum_{j=1}^n c_j u_j\right) = \sum_{i=1}^n \mu_i \sum_{j=1}^n (c_i u_i, c_j u_j)$$

$$= \sum_{i=1}^n \mu_i c_i^2 \leq \max_{1 \leq i \leq n} \{\mu_i\} \sum_{i=1}^n c_i^2 .$$

Analogamente obtemos

$$\| x \|_2^2 = \sum_{i=1}^n c_i^2$$

Logo $\| A \|_2 \leq \max_{1 \leq i \leq n} \sqrt{\mu_i} = \sqrt{\mu_{Max}}$

Novamente, obtida uma constante de limitação para o operador A, basta encontrar um vetor $x_o \in \mathbb{R}^n$ para o qual

$$\| Ax_o \|_2 = \sqrt{\mu_{Max}} \| x_o \|_2$$

O x_o em questão é qualquer autovetor associado ao maior autovalor de A^TA pois nesse caso

$$\| Ax_o \|_2 = (A^TAx_o,x_o)^{1/2} = (\mu_{Max}x_o,x_o)^{1/2} = \sqrt{\mu_{Max}} \| x_o \|_2 .$$

Logo a norma matricial associada aos operadores lineares contínuos A do espaço euclidiano $\mathbb{R}^n$ ao $\mathbb{R}^m$ é a raiz quadrada do maior autovalor de $A^T A$. Essa norma é dita <u>norma espectral</u>.

<u>Exemplo 8</u>: Sejam $C^0[0,1]$ e $C^1[0,1]$ associados às normas $\|\cdot\|_{0,\infty}$ e $\|\cdot\|_{1,\infty}$ respectivamente e o operador D:

$$D: C^1[0,1] \to C^0[0,1]$$
$$f \mapsto f' = \frac{df}{dx}$$

Vimos que $D \in \mathcal{L}(C^1(0,1), C^0[0,1])$

Temos

$$\| D \| = \sup_{f \neq 0} \frac{\| f' \|_{0,\infty}}{\| f \|_{1,\infty}} = \sup_{f \neq 0} \frac{\max\limits_{x \in [0,1]} |f'(x)|}{\max\limits_{x \in [0,1]} |f'(x)| + \max\limits_{x \in [0,1]} |f(x)|}$$

Logo temos $\| D \| \leq 1$

Tomando a seqüência de funções de $C^1[0,1]$ definida por

$$f_n(x) = \begin{cases} (x - \frac{1}{n})^2 & \text{se } 0 \leq x \leq \frac{1}{n} \\ 0 & \text{se } \frac{1}{n} < x \leq 1 \end{cases}$$

vemos que

$$\sup_n \frac{\| f_n' \|_{0,\infty}}{\| f_n \|_{1,\infty}} = 1$$

Logo $\| D \| = 1$.

Podemos agora provar o resultado seguinte:

<u>Teorema 3</u>: $\mathcal{L}(X,Y)$ é um espaço de Banach com a norma (1) se Y é um espaço de Banach.

<u>Demonstração</u>: Seja $\{A_n\}$ uma seqüência de Cauchy de $\mathcal{L}(X;Y)$

Temos: $$\lim_{m,n \to \infty} \| A_m - A_n \| = 0$$

Como $\| A_m(x) - A_n(x) \| \le \| A_m - A_n \| \; \| x \| \quad \forall x \in X$ temos:

$$\lim_{m,n\to\infty} \| A_m(x) - A_n(x) \| = 0 \quad \forall x \in X$$

o que mostra que $\{A_n(x)\}$ é uma seqüência de Cauchy de Y.

Como por hipótese Y é um espaço de Banach, existe um elemento de Y dependendo de x, digamos A(x), para o qual converge $\{A_n(x)\}$.

Provemos agora que o operador A assim definido, isto é:

$$A: X \to Y$$
$$x \to \lim_{n\to\infty} A_n(x) \quad \text{é linear.}$$

De fato, com $x_1, x_2 \in X$ e α_1, α_2 escalares,

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \lim_{n\to\infty} A_n(\alpha_1 x_1 + \alpha_2 x_2).$$

Como A_n é linear para todo n,

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \lim_{n\to\infty} [\alpha_1 A_n(x_1) + \alpha_2 A_n(x_2)] =$$

$$= \alpha_1 \lim_{n\to\infty} A_n(x_1) + \alpha_2 \lim_{n\to\infty} A_n(x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2).$$

Provemos agora que A é limitado.

Como $\{A_n\}$ é seqüência de Cauchy, dado $\varepsilon > 0$, existe N tal que

$$\| A_m(x) - A_n(x) \| \le \| A_m - A_n \| \; \| x \| < \varepsilon \| x \| \quad \forall x \in X, \; n,m > N.$$

Como

$$\lim_{m\to\infty} A_m(x) = A(x) \text{ temos:}$$

$$\| A(x) - A_n(x) \| \le \varepsilon \| x \| \quad \forall n > N, \; \forall x \in X.$$

Logo

$$\| A(x) \| \le \| A_n(x) \| + \varepsilon \| x \|$$

isto é

$$\| A(x) \| \le (\| A_n \| + \varepsilon) \| x \| \quad \forall n > N, \; \forall x \in X.$$

Sendo $\{A_n\}$ de Cauchy, existe uma constante M tal que

$\| A_n \| \le M \quad \forall n.$ Logo

$$\| A(x) \| \leq (M+\varepsilon) \| x \| \qquad \forall x \in X$$ e portanto A é limitado.

Resta provar que $A_n \to A$ em $\mathcal{L}(X,Y)$.

De fato, $\forall x \in X$ e $\forall \varepsilon > 0$, $\exists N$ tal que

$$\| A_n(x) - A(x) \| < \varepsilon \| x \| \qquad \forall n > N.$$

Logo para $x \neq \vec{0}$ temos:

$$\frac{\| A_n(x) - A(x) \|}{\| x \|} < \varepsilon \qquad \forall n > N, \text{ ou seja:}$$

$$\| A_n - A \| = \sup_{x \neq \vec{0}} \frac{\| A_n(x) - A(x) \|}{\| x \|} \leq \varepsilon \qquad \forall n > N, \text{ donde se conclui}$$

$$\| A_n - A \| \to 0$$

c.q.d.

IV.4 - ESPAÇOS DUAIS

Em virtude do importante papel que representa em Análise Funcional, o espaço $\mathcal{L}(X,Y)$ quando $Y = \mathbb{R}$ recebe a denominação especial de espaço dual de X denotado X^*.

X^* vem então a ser o espaço vetorial normado constituido por todos os funcionais lineares contínuos definidos sobre o espaço vetorial normado X.

A norma de X^* é definida por

$$\| \mathcal{F} \| = \sup_{x \neq \vec{0}} \frac{|\mathcal{F}(x)|}{\| x \|}$$

para a qual, de acordo com o Teorema 3 é um espaço de Banach, seja X completo ou não.

Conforme poderemos observar de alguns exemplos, o dual de X, apesar de ser constituido por elementos de natureza fundamentalmente distinta dos de X, freqüentemente se identificará, via isomorfismo de espaços vetoriais normados, senão ao próprio X, a um espaço vetorial de mesma natureza do que X.

<u>Exemplo 9</u>: O dual do espaço euclidiano $\mathbb{R}^n$ se identifica ao próprio $\mathbb{R}^n$ no sentido seguinte:

$$\forall \mathfrak{F} \in [\mathbb{R}^n]^* \qquad \exists y \in \mathbb{R}^n \text{ único tal que}$$

$$\mathfrak{F}(x) = \sum_{i=1}^{n} x_i y_i = (x|y) \qquad \forall x \in \mathbb{R}^n, \quad \text{com } \|\mathfrak{F}\| = \|y\|.$$

e, inversamente, cada $y \in \mathbb{R}^n$ define um único funcional $\mathfrak{F} \in [\mathbb{R}^n]^*$ com as propriedades acima.

De fato, em primeiro lugar, constatamos que, se $y \in \mathbb{R}^n$ podemos definir um funcional linear contínuo $\mathfrak{F}$ sobre $\mathbb{R}^n$ por:

$$\mathfrak{F}(x) = (y|x) \qquad \forall x \in \mathbb{R}^n$$

A linearidade tida como evidente temos:

$$|\mathfrak{F}(x)| = |(y|x)| \leq \|y\| \; \|x\| \qquad \forall x \in \mathbb{R}^n \quad \text{donde}$$

$\mathfrak{F}$ é limitado com $C = \|y\|$. Além disso podemos constatar que na verdade

$$\|\mathfrak{F}\| = \|y\| \quad \text{tomando } x=y.$$

Agora, seja $\mathfrak{F}$ um funcional linear contínuo sobre $\mathbb{R}^n$.

Sendo $\{e_i\}_{i=1}^{n}$ a base canônica de $\mathbb{R}^n$ temos:

$$\mathfrak{F}(x) = \mathfrak{F}\left[\sum_{i=1}^{n} x_i e_i\right] = \sum_{i=1}^{n} x_i \, \mathfrak{F}(e_i)$$

Então, dado $\mathfrak{F}$ podemos definir um vetor fixo $y = (y_1, y_2, \ldots, y_n)$ com $y_i = \mathfrak{F}(e_i)$ e vemos que $\mathfrak{F}(x) = (x|y) \quad \forall x \in X$.

O mesmo argumento usado anteriormente serve para mostrar que $\|\mathfrak{F}\| = \|y\|$.

<u>Exemplo 10</u>: Seja $X = \ell^1$ com a norma usual. Vamos mostrar que há um isomorfismo entre $[\ell^1]^*$ e ℓ^∞ no sentido seguinte: a cada $\mathfrak{F} \in [\ell^1]^*$ podemos associar uma única seqüência $y \in \ell^\infty$ com

$$\mathfrak{F}(x) = \sum_{i=1}^{\infty} x_i y_i \qquad \forall x = (x_1, x_2, \ldots, x_i, \ldots) \in \ell^1 \quad \text{e}$$

$$\|\mathfrak{F}\| = \|y\| = \sup_i |y_i|$$

e, inversamente, cada $y \in \ell^\infty$ define um único funcional $\mathfrak{F} \in [\ell^1]^*$ com as propriedades acima.

Inicialmente cada $y \in \ell^\infty$ define dessa forma um funcional linear contínuo sobre ℓ^1 pois:

$$|\mathcal{F}(x)| = |\sum_{i=1}^{\infty} x_i y_i| \leq \sum_{i=1}^{\infty} |x_i|\ |y_i| \leq \sup_i |y_i| \sum_{i=1}^{\infty} |x_i|, \text{ isto é:}$$

$|\mathcal{F}(x)| \leq C \, \| x \|_1$ com $C = \| y \|_\infty$, ou seja

$$\| \mathcal{F} \| \leq \| y \|_\infty$$

Agora, provamos que também $\| \mathcal{F} \| \geq \| y \|_\infty$. De fato seja $\{x_n\}$ a seqüência de ℓ^1, $x_n = (x_1^n, x_2^n, \ldots, x_i^n, \ldots)$, com

$$\begin{cases} x_i^n = \text{sinal } y_n & \text{se } i=n \\ x_i^n = 0 & \text{se } i \neq n \end{cases}$$

Nesse caso $|y_n| = \mathcal{F}(x_n) \leq \| \mathcal{F} \|\ \| x_n \|_1 = \| \mathcal{F} \| \quad \forall n.$

Logo $\sup_n |y_n| = \| y \|_\infty \leq \| \mathcal{F} \|$ donde se conclui que

$$\| \mathcal{F} \| = \| y \|_\infty .$$

Enfim, provamos que, dado um funcional $\mathcal{F} \in [\ell^1]^*$ existe $y \in \ell^\infty$ tal que

$$\mathcal{F}(x) = \sum_{i=1}^{\infty} x_i y_i$$

Seja $\{e_1, e_2, \ldots, e_i, \ldots\}$ com $\begin{cases} e_j^i = 1 & \text{se } i=j \\ e_j^i = 0 & \text{se } i \neq j \end{cases}$

Considerando-se que $\forall x \in \ell^p$, $1 \leq p < \infty$, tem-se $z_n \to x$ com $z_n = \sum_{i=1}^{n} x_i e_i$ (verificar), conclui- que $x = \sum_{i=1}^{\infty} x_i e_i$. (*)

Logo

$$\mathcal{F}(x) = \mathcal{F}\left(\sum_{i=1}^{\infty} x_i e_i\right) = \mathcal{F}\left(\lim_{n\to\infty} \sum_{i=1}^{n} x_i e_i\right)$$

Por outro lado, $\mathcal{F}\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i\, \mathcal{F}(e_i) \quad \forall n.$

Então, pela continuidade de $\mathcal{F}$, como $x = \lim_{n\to\infty} \sum_{i=1}^{n} x_i e_i$, temos:

*Portanto (a menos de um ε arbitrariamente pequeno) $\{e_1, e_2, \ldots, e_j, \ldots\}$ é uma "base" de ℓ^p, $1 \leq p < \infty$ (ver sub-capítulo II.6).

$$\mathfrak{F}(x) = \lim_{n\to\infty} \sum_{i=1}^{n} x_i\, \mathfrak{F}(e_i) = \sum_{i=1}^{\infty} x_i\, \mathfrak{F}(e_i)$$

Resta provar que $y = (\mathfrak{F}(e_1), \mathfrak{F}(e_2), \ldots, \mathfrak{F}(e_i), \ldots) \in \ell^\infty$.

De fato, se assim não fosse teríamos:

$$\sup_i |\mathfrak{F}(e_i)| = \sup_i \frac{|\mathfrak{F}(e_i)|}{\| e_i \|_1} = \infty$$

o que está em contradição com a suposta limitação de $\mathfrak{F}$.

<u>Exemplo 11</u>: De modo análogo ao Exemplo 10, vamos provar que existe um isomorfismo entre $[\ell^p]^*$ e ℓ^q com

$$\frac{1}{p} + \frac{1}{q} = 1, \quad p>1, \text{ isto é:}$$

Dado $\mathfrak{F} \in [\ell^p]^*$, $\exists\, y \in \ell^q$ único ou dado $y \in \ell^q$, $\exists\, \mathfrak{F} \in [\ell^p]^*$ único tal que:

$$\mathfrak{F}(x) = \sum_{i=1}^{\infty} x_i y_i \quad \forall x \in \ell^p, \text{ com } \| \mathfrak{F} \| = \| y \|_q .$$

Em primeiro lugar, se $\mathfrak{F}$ se exprime dessa forma com $y \in \ell^q$, temos pela desigualdade de Hölder :

$$|\mathfrak{F}(x)| = \left|\sum_{i=1}^{\infty} x_i y_i\right| \le \left[\sum_{i=1}^{\infty} x_i^p\right]^{1/p} \left[\sum_{i=1}^{\infty} y_i^q\right]^{1/q}$$

donde $\| \mathfrak{F} \| \le \| y \|_q$.

Para provarmos que $\| \mathfrak{F} \| = \| y \|_q$, basta considerarmos a seqüência $\{x_n\} \subset \ell^p$ com

$$\begin{cases} x_i^n = |y_i|^{q/p} \operatorname{sinal}\{y_i\} & \text{se } i \le n \\ x_i^n = 0 & \text{se } i > n \end{cases}$$

Temos então:

$$\| x_n \|_p = \left[\sum_{i=1}^{n} |y_i|^q\right]^{1/p} \text{ e}$$

$$\mathfrak{F}(x_n) = \sum_{i=1}^{n} |y_i|^{q/p + 1} = \sum_{i=1}^{n} |y_i|^q$$

Como $|\mathfrak{F}(x_n)| \leq \|\mathfrak{F}\| \; \|x_n\|_p \quad \forall n$, temos:

$$\sum_{i=1}^{n} |y_i|^q \leq \|\mathfrak{F}\| \left[\sum_{i=1}^{n} |y_i|^q\right]^{1/p} \Rightarrow \left[\sum_{i=1}^{n} |y_i|^q\right]^{1/q} \leq \|\mathfrak{F}\| \quad \forall n.$$

Logo $\|y\|_q = \left[\sum_{i=1}^{\infty} |y_i|^q\right]^{1/q} \leq \|\mathfrak{F}\|$ e só podemos ter

$$\|y\|_q = \|\mathfrak{F}\|.$$

Para provar agora que a cada $\mathfrak{F}\in[\ell^p]^*$ está associado $y\in\ell^q$ procedemos como no Exemplo 10 e concluimos que

$$\mathfrak{F}(x) = \sum_{i=1}^{\infty} x_i \, \mathfrak{F}(e_i)$$

Agora verificamos que $y=(\mathfrak{F}(e_1), \mathfrak{F}(e_2), \ldots, \mathfrak{F}(e_i), \ldots)\in\ell^q$ com os mesmos argumentos usados para provar que

$$\|y\|_q \leq \|\mathfrak{F}\|.$$

<u>Exemplo 12</u>: Vimos que ℓ^q é isómorfo ao dual de ℓ^p se $\frac{1}{p} + \frac{1}{q} = 1$ com $1\leq p<\infty$.

Vamos agora ver que ℓ^1 é isomorfo ao dual de c_o com a norma do supremo e mais tarde que não o é ao dual de ℓ^∞, como se poderia supor.

Em primeiro lugar, constatamos que $\forall x = (x_1,x_2,\ldots,x_i,\ldots)\in c_o$ e com $y = (y_1,y_2,\ldots,y_i,\ldots)\in \ell^1$ temos:

$$\mathfrak{F}(x) = \sum_{i=1}^{\infty} x_i y_i \leq \sup_i |x_i| \sum_{i=1}^{\infty} |y_i| = C \; \|x\|$$

com $C = \|y\|_1$, isto é: $\|\mathfrak{F}\| \leq \|y\|_1$.

Seja agora $\mathfrak{F}$ um elemento qualquer de c_o^*.

Como no Exemplo 10, temos que $\forall x\in c_0, x = \sum_{i=1}^{\infty} x_i e_i$, ou seja, $\{e_1,e_2,\ldots,e_i,\ldots\}$ é uma "base" de c_0. (a menos de um $\varepsilon>0$)

Logo temos: $\mathfrak{F}(x) = \mathfrak{F}\left(\sum_{i=1}^{\infty} x_i e_i\right)$ e como nos Exemplos 10 e 11:

$$\mathfrak{F}(x) = \sum_{i=1}^{\infty} x_i \, \mathfrak{F}(e_i).$$

Agora definindo-se $y_i = \mathfrak{F}(e_i)$, seja $y = (y_1, y_2, \ldots, y_i, \ldots)$

Provemos que $y \in \ell^1$. Para tanto definimos uma seqüência $\{x_n\} \in c_o$ com

$$\begin{cases} x_i^n = \text{sinal } y_i & \text{se } i \leq n \\ x_i^n = 0 & \text{se } i > n \end{cases}$$

É claro que $\| x_n \| = 1$ $\forall n$. Por outro lado, por construção,

$$\mathfrak{F}(x_n) = \sum_{i=1}^{n} |y_i| \leq \| \mathfrak{F} \| \; \| x_n \| = \| \mathfrak{F} \| \;\; \forall n.$$

Por conseguinte $\lim\limits_{n\to\infty} \sum\limits_{i=1}^{n} |y_i| \leq \| \mathfrak{F} \| \Rightarrow y \in \ell^1$ e $\| y \|_1 \leq \| \mathfrak{F} \|$.

Enfim como da 1ª parte da demonstração já estabelecemos que

$\| y \|_1 \geq \| \mathfrak{F} \|$ temos $\| y \|_1 = \| \mathfrak{F} \|$ como desejado.

No Exemplo 9 vimos um isomorfismo entre o espaço X e seu dual.

Também no Exemplo 11, se considerarmos p=2 vemos que ℓ^2 se identifica ao seu dual. Observa-se ainda que em ambos os casos o espaço X em questão é de Hilberte que o funcional linear contínuo se exprime pelo produto escalar de X com um elemento fixo de X. Na verdade pode-se provar o interessante e poderoso resultado seguinte:

<u>Teorema 4</u>: (Teorema de Riesz).

Um espaço de Hilbert X pode ser identificado a seu dual X^* de forma que a todo $\mathfrak{F} \in X^*$ corresponde um único $y \in X$ e a todo $y \in X$ corresponde um único $\mathfrak{F} \in X^*$ tais que:

$$\forall x \in X \qquad \mathfrak{F}(x) = (y|x) \quad \text{com} \quad \| \mathfrak{F} \| = \| y \|$$

<u>Demonstração</u>: Inicialmente cada $y \in X$ define um funcional linear contínuo sobre X com

$$\mathfrak{F}(x) = (y|x) \qquad \forall x \in X$$

sendo a linearidade óbvia e a continuidade conseqüência de

$$|\mathfrak{F}(x)| \leq \|y\| \, \|x\| \quad \forall x \in X, \text{ isto é, } C = \|y\|.$$

Na verdade $\|\mathfrak{F}\| = \|y\|$ bastando para a verificação tomar $x=y$.

Seja agora $\mathfrak{F} \in X^*$. O espaço nulo de $\mathfrak{F}$ é um subespaço fechado de X pois se $\{x_n\} \subset \mathcal{N}(\mathfrak{F})$ temos

$\mathfrak{F}(x_n)=0 \quad \forall n$. Logo, se $x_n \to x$ em X temos pela suposta continuidade de $\mathfrak{F}$ que $\mathfrak{F}(x) = \lim\limits_{n\to\infty} \mathfrak{F}(x_n) = 0$ e $x \in \mathcal{N}(\mathfrak{F})$.

Agora se $\mathcal{N}(\mathfrak{F}) = X$ temos $\mathfrak{F} = \theta$ e podemos associar a $\mathfrak{F}$ o vetor nulo de X.

Suponhamos então $\mathcal{N}(\mathfrak{F}) \neq X$. Nesse caso como $\mathcal{N}(\mathfrak{F})$ é fechado, X é a soma direta $\mathcal{N}(\mathfrak{F}) \oplus \mathcal{N}(\mathfrak{F})^{\perp}$ sendo $\mathcal{N}(\mathfrak{F})^{\perp}$ não reduzido ao vetor nulo. Tomemos pois um vetor $z \neq \vec{0}$ tal que $z \in \mathcal{N}(\mathfrak{F})^{\perp}$. Dado que nesse caso $\mathfrak{F}(z) \neq 0$. podemos encontrar um vetor $w \in \mathcal{N}(\mathfrak{F})^{\perp}$ com

$$w = \frac{z}{\mathfrak{F}(z)} \quad \text{e nesse caso} \quad \mathfrak{F}(w) = 1.$$

Agora dado $x \in X$, $x - \mathfrak{F}(x) w \in \mathcal{N}(\mathfrak{F})$ pela linearidade de $\mathfrak{F}$ e pela construção de w.

Porém $(x - \mathfrak{F}(x) w | w) = 0$ dado que $w \in \mathcal{N}(\mathfrak{F})^{\perp}$. Desenvolvendo vem:

$$(x|w) = (\mathfrak{F}(x) w | w) \Rightarrow \mathfrak{F}(x) = \frac{(x|w)}{\|w\|^2}$$

Então definindo-se $y = \dfrac{w}{\|w\|^2}$ vemos que

$$\mathfrak{F}(x) = (y|x) \quad \forall x \in X.$$

Resta provar que y é único para o que basta supor que também

$$\mathfrak{F}(x) = (y'|x) \quad \forall x \in X$$

Nesse caso porém

$$(y' - y | x) = 0 \quad \forall x \in X \Rightarrow y' = y. \qquad \text{c.q.d.}$$

Vamos agora apresentar um resultado bastante forte tanto para desenvolvimentos teóricos como para aplicações: o Teorema de Hahn-Banach. Apesar de ter um enunciado que poderá parecer, no início, algo abstrato, de imediato veremos que alguns seus corolários permitem estabelecer resultados de caracterização de duais de certos espaços, o que seria extremamente difícil pela forma mais direta dos Exemplos 9,10,11,12.

Em benefício da simplicidade vamos demonstrar esse teorema para o caso em que X é um espaço separável. O caso geral é tratado por exemplo em [1]. Antes de enunciar o Teorema julgamos conveniente introduzir o conceito seguinte:

Definição 7: Seja M subespaço de um espaço vetorial normado X e f um funcional linear contínuo sobre M. Uma extensão de f a todo X é um funcional $\mathfrak{F}$,

$$\mathfrak{F}: X \to \mathbb{R}$$

tal que

$$\mathfrak{F}(m) = f(m) \quad \forall m \in M$$

Exemplo 13: Seja X o espaço das seqüências de reais que convergem, com a norma

$$\| x \| = \sup_i |x_i|.$$

Podemos considerar c_o como um subespaço de X com a mesma norma. Seja agora o funcional nulo

$$\theta: c_o \to \mathbb{R}$$

Podemos estender θ a X de várias maneiras sendo uma delas tomando $\mathfrak{F} = \theta$.

Outra maneira de estender esse funcional é através de

$$\mathfrak{F}: X \to \mathbb{R}$$
$$x \mapsto \lim_{i\to\infty} x_i$$

$\mathfrak{F}$ é claramente linear, contínuo com C=1

$$\mathfrak{F}(x) \leq \lim_{i\to\infty} |x_i| \leq \sup_i |x_i| = \| x \|$$

Além disso $\mathfrak{F}(x) = f(x) = 0 \quad \forall x \in c_o$.

<u>Teorema 5</u>: (Teorema de Hahn-Banach): Todo funcional linear limitado f definido apenas sobre um subespaço M de um espaço vetorial normado X tem uma extensão linear $\mathcal{F}$ definida sobre todo X, limitada pela norma de f, isto é:

Se $\| f \| = \sup\limits_{m \neq 0} \dfrac{|f(m)|}{\| m \|}$ então $|\mathcal{F}(x)| \leq \| f \| \, \| x \| \quad \forall x \in X.$

<u>Demonstração</u> (supondo que X é separável):

Nesse caso existe um conjunto enumerável $D = \{x_1, x_2, \ldots, x_i, \ldots\}$ denso em X ou seja $\forall \varepsilon > 0$ e $\forall x \in X, \exists x_i \in D$ tal que

$$\| x_i - x \| < \varepsilon.$$

Seja S o subespaço de X gerado por M acrescido de vetores $y_1, y_2, \ldots, y_k, \ldots$, retirados de [D] linearmente independentes entre si e de M.

$S = [M, y_1, y_2, \ldots, y_k, \ldots]$ é denso em X, pois $S \supset D$

Vamos agora usar a indução.

Estendemos inicialmente f a X_1 gerado por M e y_1.

($y_1 \notin M$ por construção)

Assim se $x \in X_1$, x é da forma $x = m + \alpha y_1$ para certo $m \in M$ e α escalar.

Seja $f_1(x) = f(m) + \lambda\alpha$, $\lambda \in \mathbb{R}$.

Vamos mostrar que para certo λ bem escolhido f_1 é a extensão procurada:

f_1 é linear para dado λ o que se pode verificar sem dificuldades.

f_1 é extensão de f pois se $\alpha = 0$ tem :

$$|f_1(x)| = |f(m)| \leq \| f \| \, \| m \| \qquad \forall m \in M.$$

Queremos ainda que f_1 seja limitada por $\| f \|$, isto é:
$|f_1(x)| \leq \| f \| \, \| x \| \quad \forall x \in X_1$.

Se $\alpha \neq 0$ devemos ter então, dividindo-se membro a membro por $-\alpha$:

$$| f(\frac{m}{-\alpha})-\lambda | \leq \| f \| \; \| \frac{m}{-\alpha} - y_1 \| \qquad \forall m \in M$$

Como M é subespaço temos que ter na realidade:

$$| f(m) - \lambda | \leq \| f \| \; \| m-y_1 \| \qquad \forall m \in M$$

Para que tal ocorra é preciso tomar

$$f(m) - \| f \| \; \| m-y_1 \| \leq \lambda \leq f(m) + \| f \| \; \| m-y_1 \| \quad \forall m \in M$$

Logo, para que λ exista devemos ter

$$\bigcap_{m \in M} [c_m^-, c_m^+] \neq \emptyset \quad \text{onde} \quad c_m^- = f(m) - \| f \| \; \| m-y_1 \|$$

$$\text{e} \quad c_m^+ = f(m) + \| f \| \; \| m-y_1 \| .$$

Ora, $\forall \; u,v \in M$ temos:

$$f(u) - f(v) = f(u-v) \leq | f(u-v) | \leq \| f \| \; \| u-v \| \text{, ou ain-}$$

da $f(u)-f(v) \leq \| f \| \; \| u-y_1+y_1-v \| \leq \| f \| \; [\| u-y_1 \| + \| v-y_1 \|]$

ou seja: $f(u) - \| f \| \; \| u-y_1 \| \leq f(v) + \| f \| \; \| v-y_1 \| \quad \forall u,v \in M.$

Logo $\sup\limits_{u \in M} [f(u) - \| f \| \; \| u-y_1 \|] \leq f(v) + \| f \| \; \| v-y_1 \| \quad \forall v \in M$

ou ainda

$$c^- = \sup_{u \in M} [f(u) - \| f \| \; \| u-y_1 \|] \leq \inf_{v \in M} [f(v) + \| f \| \; \| v-y_1 \|] = c^+$$

como $[c^-, c^+]$ está contido em $\bigcap_{m \in M} [c_m^-, c_m^+]$ esta interseção não é vazia e podemos pois tomar $\lambda \in [c^-, c^+]$.

Agora por construções análogas estendemos f_1 a $X_2=[M, y_1, y_2]$ obtendo um funcional f_2, $\| f_2 \| \leq \| f \|$ e assim sucessivamente podemos estender f a S obtendo um funcional F com $F(m) = f(m)$ $\forall m \in M$ e $\| F \| \leq \| f \|$

Enfim, como S é denso em X para todo $x \in X$ $\exists \{s_n\} \subset S$ com $s_n \to x$.

Agora definimos o funcional $\mathfrak{F}$ por:

$$\mathfrak{F}(x) = \lim_{n\to\infty} F(s_n)$$

$\mathfrak{F}$ é linear e $\mathfrak{F}(s) = F(s)$ se $s\in S$ pois F é contínuo.

Além disso $|\mathfrak{F}(x)| = |\lim_{n\to\infty} F(s_n)| \leq \lim_{n\to\infty} \| f \| \, \| s_n \|$

Logo $|\mathfrak{F}(x)| \leq \| f \| \, \| x \| \quad \forall x\in X$ c.q.d.

<u>Observação</u>: Na verdade $\| f \|$ é a norma de $\mathfrak{F}$ pois

$$\| f \| \geq \sup_{\substack{x\neq\vec{0}\\ x\in X}} \frac{|\mathfrak{F}(x)|}{\| x \|} \geq \sup_{\substack{m\in M\\ m\neq\vec{0}}} \frac{|\mathfrak{F}(m)|}{\| m \|} = \| f \|$$

Vejamos agora algumas conseqüências importantes do Teorema de Hahn-Banach.

<u>Corolário 1</u> Seja $x_o\in X$, X sendo espaço vetorial normado. Então existe um funcional $\mathfrak{F}\in X^*$ tal que

$$\mathfrak{F}(x_o) = \| \mathfrak{F} \| \quad \| x_o \|$$

(ou seja, o supremo da norma de algum funcional $\mathfrak{F}$ do dual de X é atingido para $x=x_o$).

$$\sup_{x\neq\vec{0}} \frac{|\mathfrak{F}(x)|}{\| x \|} = \frac{\mathfrak{F}(x_o)}{\| x_o \|} = \| \mathfrak{F} \|$$

<u>Demonstração</u>: Seja $M = [x_o]$

Se $m\in M$ então $m=\alpha x_o$ para certo escalar α.

Seja $f : M \to \mathbb{R}$ tal que $f(m) = \alpha \| x_o \|$

f é claramente linear e além disso

$$\| f \| = \sup_{\alpha\in\mathbb{R}} \frac{|\alpha \| x_o \| \,|}{\| \alpha x_o \|} = 1$$

Ora, pelo Teorema de Hahn-Banach existe $\mathfrak{F}$ tal que

$$\mathfrak{F} : X \to \mathbb{R} \quad \text{com} \quad \| \mathfrak{F} \| = \| f \|$$

Esse funcional $\mathfrak{F}$ será tal que

$$\mathfrak{F}(x_o) = 1 \| x_o \| = \| f \| \; \| x_o \| . \qquad \text{c.q.d}$$

<u>Observação</u>: Na demonstração do Corolário 1 encontramos um funcional $\mathfrak{F}$ de norma unitária. Na verdade poderíamos encontrar funcionais com qualquer norma λ bastando para isso definir

$$f(m) = \lambda\alpha \; \| x_o \| \qquad \forall m \in M.$$

A recíproca do Corolário 1 não é verdadeira, isto é, pode existir $\mathfrak{F} \in X^*$ para o qual não exista $x_o \in X$ com

$$\| \mathfrak{F} \| = \sup_{x \neq 0} \frac{|\mathfrak{F}(x)|}{\| x \|} = \frac{\mathfrak{F}(x_o)}{\| x_o \|} \,, \quad \text{senão vejamos:}$$

<u>Exemplo 14</u>: Seja $X = \ell^1$. Sabemos que $X^* \equiv \ell^\infty$; tomemos o funcional $\mathfrak{F}$ associado a $y \in \ell^\infty$ $\quad y = (0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots)$

Temos $\quad \| y \|_\infty = \sup_{1 \le i \le \infty} |\frac{i-1}{i}| = 1 = \| \mathfrak{F} \|$

Agora considerando a base $\{e_i\}$ de ℓ^1 vemos que na verdade, $|\mathfrak{F}(x)| < \| x \| \quad \forall x \in X$ pois:

$$|\mathfrak{F}(e_i)| < \| e_i \| = 1 \qquad \forall i.$$

donde $\quad |\mathfrak{F}(x)| \le \sum_{i=1}^{\infty} |x_i \, \mathfrak{F}(e_i)| < \sum_{i=1}^{\infty} |x_i| \qquad \forall x \in \ell^1$

o que mostra que não é possível encontrar $x_o \in \ell^1$ com

$$\frac{\mathfrak{F}(x_o)}{\| x_o \|} = 1$$

<u>Corolário 2</u> Seja X um espaço vetorial normado e M subespaço de X. M é denso em X se e somente se todo funcional linear contínuo sobre X que se anular identicamente sobre M for também identicamente nulo sobre X.

<u>Demonstração</u>: A condição é necessária, isto é:

$\bar{M} = X \Rightarrow \{ \mathfrak{F}(m)=0 \quad \forall m \in M \Rightarrow \mathfrak{F}(x)=0 \quad \forall x \in X\}$

De fato, se M é denso em X então

$\forall x \in X \ \exists \{x_n\} \subset M$ tal que $\| x_n - x \| \to 0$.

Agora se $\mathfrak{F}(x_n)=0 \quad \forall n$ temos $\lim\limits_{n\to\infty} \mathfrak{F}(x_n)=0$ e por continuidade, $\mathfrak{F}(x)=0$.

A condição é suficiente, isto é:

$\{\mathfrak{F}(m)=0 \quad \forall m \in M \Rightarrow \mathfrak{F}(x)=0 \quad \forall x \in X\} \Rightarrow \bar{M} = X$.

Suponhamos que M não seja denso em X. Nesse caso $\exists x_o \in X$ tal que $\inf\limits_{m \in M} \| x_o - m \| = \delta > 0$

Vamos agora construir um funcional $\mathfrak{F}_o$ tal que $\mathfrak{F}_o(m)=0$ $\forall m \in M$ mas para certo $u \notin M$, $u \in X$ $\mathfrak{F}_o(u) \neq 0$.

Seja W o subespaço de X gerado por M e x_o. Se $w \in W$ então $w = m + \alpha x_o$ para certo $m \in M$ e α escalar.

Seja F_o um funcional definido sobre W tal que

$$F_o(w) = \alpha \qquad \forall w \in W$$

F_o se anula identicamente sobre M e é linear.

Por outro lado $F_o(w) \le C \| w \| \quad \forall w \in W$.

De fato, com

$$C = \sup_{\substack{w \in W \\ w \neq \vec{0}}} \frac{|F_o(w)|}{\| w \|} = \sup_{\substack{\alpha \in \mathbb{R} \\ m \in M}} \frac{|\alpha|}{\| m + \alpha x_o \|}$$

$$= \sup_{\substack{\alpha\in\mathbb{R}\\ m\in M}} \frac{1}{\left\| \frac{-m}{|\alpha|} - x_o \right\|} = \sup_{m\in M} \frac{1}{\| m-x_o \|}$$

já que M é subespaço. Logo $C = \frac{1}{\delta} > 0$.

Agora, pelo Teorema de Hahn-Banach, podemos prolongar F_o em $\mathfrak{F}_o$ sobre X com $|\mathfrak{F}_o(x)| \leq \frac{1}{\delta} \| x \| \quad \forall x\in X$ e $\mathfrak{F}_o(x) = F_o(x) \quad \forall x\in W$.

Mas se $u = m + x_o \in X$ temos $\mathfrak{F}_o(u) = 1 \neq 0$ o que é absurdo. c.q.d.

Como X^* é também um espaço vetorial normado, podemos definir seu dual $[X^*]^*$. Note-se que a qualquer $x\in X$ está associado um funcional linear contínuo $\tilde{x}$ sobre X^* definido por

$$\tilde{x}(\mathfrak{F}) = \mathfrak{F}(x) \quad \forall \mathfrak{F}\in X^*$$

De fato temos a linearidade pois

$$\tilde{x}(\alpha_1\mathfrak{F}_1+\alpha_2\mathfrak{F}_2) = (\alpha_1\mathfrak{F}_1+\alpha_2\mathfrak{F}_2)(x) = \alpha_1\mathfrak{F}_1(x)+\alpha_2\mathfrak{F}_2(x)$$

$$= \alpha_1\tilde{x}(\mathfrak{F}_1) + \alpha_2\tilde{x}(\mathfrak{F}_2)$$

e a continuidade pois

$$|\tilde{x}(\mathfrak{F})| \leq \| \mathfrak{F} \| \; \| x \| \quad \forall \mathfrak{F}\in X^* \quad \text{ou seja} \quad \| \tilde{x} \| \leq \| x \|.$$

Adicionalmente pelo Corolário 1 do teorema de Hahn-Banach $\forall x\in X$, e por conseguinte para o $\tilde{x}\in[X^*]^*$ associado, $\exists \mathfrak{F}\in X^*$ tal que:

$$\| \mathfrak{F} \| = \frac{\mathfrak{F}(x)}{\| x \|} = \frac{\tilde{x}(\mathfrak{F})}{\| x \|}, \text{ ou seja } \| x \| = \frac{\tilde{x}(\mathfrak{F})}{\| \mathfrak{F} \|} \text{ o que nos}$$

leva a concluir que na verdade $\| \tilde{x} \| = \| x \|$.

<u>Definição 8</u>: Bidual X^{**} de um espaço vetorial normado X é o espaço de funcionais lineares contínuos definidos sobre seu dual X^*.

Pela análise acima vemos que X sempre se identifica a uma parte do seu bidual. Aliás para efeito de simplificação escrevemos impropriamente

$$X \subset X^{**}$$

como feito freqüentemente na literatura, tendo-se no entanto presente, que a inclusão em X^{**} é de um espaço isomorfo a X'e não do próprio X.

<u>Definição 9</u>: Um espaço de Banach X é dito <u>reflexivo</u> se $X \equiv X^{**}$.

Exemplos de espaços reflexivos são numerosos:

$\mathbb{R}^n$ espaço euclidiano é reflexivo
Todo espaço de Hilbert é reflexivo
ℓ^p é reflexivo se $p>1$

No entanto, ℓ^1 não é reflexivo o que será uma conseqüencia do:

<u>Corolário 3</u> . Se X é um espaço de Banach reflexivo então $\forall \mathcal{F} \in X^* \quad \exists x_o \in X$ tal que

$$\| \mathcal{F} \| = \sup_{x \neq 0} \frac{|\mathcal{F}(x)|}{\| x \|} = \frac{\mathcal{F}(x_o)}{\| x_o \|}$$

<u>Demonstração</u>: De fato, pelo Corolário 1 $\forall \mathcal{F}_o \in X^* \quad \exists \tilde{x}_o \in X^{**}$ tal que

$$\frac{\tilde{x}_o(\mathcal{F}_o)}{\| \mathcal{F}_o \|} = \| \tilde{x}_o \|$$

Como todo funcional linear contínuo de X^{**} é dado por elemento de X, nesse caso sendo x_o o elemento de X associado a $\tilde{x}_o$ temos:

$$\| x_o \| = \frac{\tilde{x}_o(\mathcal{F}_o)}{\| \mathcal{F}_o \|} = \frac{\mathcal{F}_o(x_o)}{\| \mathcal{F}_o \|}$$ ou seja, para dado

$\mathcal{F}_o \in X^*$, $\exists x_o \in X$ tal que

$$\| \mathcal{F}_o \| = \frac{|\mathcal{F}_o(x_o)|}{\| x_o \|}$$ c.q.d.

Como conseqüência vemos que ℓ^1 não é reflexivo pois para o funcional $\mathcal{F}_o$ associado à seqüência $(0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots) \in \ell^\infty$ não existe $x_o \in \ell^1$ tal que

$$\| \mathcal{F}_o \| \quad \frac{|\mathcal{F}_o(x_o)|}{\| x_o \|}$$

como vimos no Exemplo 14.

Para concluir este parágrafo fazemos uma observação sobre notação freqüentemente adotada para duais. Define-se a aplica - ção dualidade:

$$< .|. >: X \times X^* \to \mathbb{R}$$
$$(x,x^*) \to x^*(x)$$

Assim se x^* é um funcional linear contínuo sobre X escrevemos alternativamente: $x^*(x) = < x|x^* >$

Vemos que a dualidade $< .|. >$ é, em certo sentido, semelhante ao produto escalar, especialmente com respeito às propriedades de linearidade e continuidade. Vejamos:

A linearidade com respeito a x valendo por definição, a linearidade com respeito a x^* decorre da própria definição da adição e multiplicação escalar em X^*.

Por outro lado o produto escalar é contínuo no sentido seguinte:

Se $x_n \to x$ e $y_n \to y$ então $(x_n|y_n) \to (x|y)$.

De fato, temos:

$$\begin{aligned} |(x|y) - (x_n|y_n)| &= |(x|y) - (x_n|y) + (x_n|y) - (x_n|y_n)| \\ &= |(x-x_n|y) + (x_n|y-y_n)| \\ &\leq \| x-x_n \| \, \| y \| + \| x_n \| \, \| y-y_n \| \end{aligned}$$

Como $\{x_n\}$ converge, $\exists M>0$ tal que $\| x_n \| \leq M \quad \forall n$.

Logo, quando $n\to\infty$ temos $(x_n|y_n) \to (x|y)$.

O mesmo argumento poderá ser usado para provar que a aplicação dualidade é contínua, isto é:

$<x_n|x_n^*> \to <x|x^*>$ se $x_n \to x$ em X e $x_n^* \to x^*$ em X^*, dado que, como para o produto escalar, também vale:

$$|<x,x^*>| \leq \| x \| \, \| x^* \|$$

Nota-se enfim que, na prática, X^* é frequentemente "substituído" por um espaço a ele isomorfo na aplicação dualidade. Nesse sentido, por exemplo, se X é um espaço de Hilbert pode-se tomar

$$< .|. > = (.|.)$$

IV-5 - OPERADORES ADJUNTOS

Vamos agora estudar uma associação de operadores através da dualidade, em particular, no caso em que são definidas sobre espaços de Hilbert.

Damos inicialmente a noção mais geral de ajuntos:

Sejam X e Y espaços vetoriais normados e $A \in \mathcal{L}(X,Y)$

Como sabemos X^* e Y^* são espaços vetoriais normados. Agora para $y^* \in Y^*$ dado, podemos definir o funcional sobre X associado a A, dependendo de y^*:

$$A^*(y^*) : X \to \mathbb{R}$$
$$x \mapsto < A(x)|y^* >$$

Claramente $A^*(y^*)$ é linear. Além disso é contínuo pois

$$|<x|A^*(y^*)>|=|<A(x)|y^*>| \leq \| y^* \| \; \| A(x) \| \leq \| y^* \| \; \| A \| \; \| x \|$$

isto é $C = \| y^* \| \; \| A \|$.

Por outro lado, a dependência de y^* é linear pois $\forall x \in X$

$$\begin{aligned} <x|A^*(\alpha_1 y_1^* + \alpha_2 y_2^*)> &= <A(x)|\alpha_1 y_1^* + \alpha_2 y_2^*> = \alpha_1 <A(x)|y_1^*> + \alpha_2 <A(x)|y_2^*> \\ &= \alpha_1 <x|A^*(y_1^*)> + \alpha_2 <x|A^*(y_2^*)> \\ &= <x|\alpha_1 A^*(y_1^*) + \alpha_2 A^*(y_2^*)> \end{aligned}$$

ou seja:

$$A^*(\alpha_1 y_1^* + \alpha_2 y_2^*) = \alpha_1 A^*(y_1^*) + \alpha_2 A^*(y_2^*)$$

Podemos então considerar a aplicação:

$$A^*: Y^* \to X^*$$
$$y^* \mapsto A(y^*)$$

como um operador linear de acordo com a

Definição 10: O operador linear A^* de Y^* em X^* definido por

$$\langle x|A^*(y^*)\rangle = \langle A(x)|y^*\rangle \quad \forall x \in X, \quad \forall y^* \in Y^*.$$

A^* é dito operador adjunto de A

Proposição 7: O operador adjunto é limitado com $\| A^* \| = \| A \|$

Demonstração: Já vimos que

$$|\langle x|A^*(y^*)\rangle| \le \| y^* \| \, \| A \| \, \| x \| \qquad \forall x \in X$$

Logo

$$\| A^*(y^*) \| \le \| A \| \, \| y^* \| \Rightarrow \| A^* \| \le \| A \|$$

Tomando agora cada $y_o \in A(X)$, o Corolário 1 do Teorema de Hahn-Banach assegura que existe $y_o^* \in Y^*$ tal que $\| y_o^* \| = 1$ e $\langle y_o|y_o^*\rangle = \| y_o \|$

Por outro lado, como $\exists x_o \in X$ tal que $y_o = A(x_o)$ temos:

$$\| A(x_o) \| = \langle A(x_o)|y_o^*\rangle = \langle x_o|A^*(y_o^*)\rangle \le \| A^*(y_o^*) \| \, \| x_o \| \le \| A^* \| \, \| x_o \|$$

Por conseguinte temos também

$\| A \| \le \| A^* \|$. o que prova o resultado desejado

c.q.d.

O leitor poderá provar como exercício as seguintes propriedades elementares dos adjuntos:

1ª) Se $\mathcal{I}$ é o operador identidade definido sobre um espaço vetorial X, $\mathcal{I}^*$ é o operador identidade definido sobre X^*

2ª) Se $A_1, A_2 \in \mathcal{L}(X,Y)$ então $(A_1+A_2)^* = A_1^* + A_2^*$

3ª) Se $A \in \mathcal{L}(X,Y)$, para α escalar temos $(\alpha A)^* = \alpha A^*$

No caso em que o operador A tem um inverso A^{-1}, podemos igualmente definir o inverso de seu adjunto de acordo com a

<u>Proposição 8</u>: Se A^{-1} existe então existe $(A^*)^{-1} = (A^{-1})^*$.

<u>Demonstração</u>: Precisamos provar que A^* é bijetivo, ou seja:

<u>A^* é injetivo</u>: Se $A^*(y_1^*) = A^*(y_2^*)$ então

$$< A(x)|y_1^*> = <A(x)|y_2^*> \quad \forall x \in X.$$

Como A é sobrejetivo $<y|y_1^*>=<y|y_2^*>$ $\forall y \in Y$ ou seja

$$y_1^* = y_2^*.$$

<u>A^* é sobrejetivo</u>: Dado $x^* \in X^*$ e $\forall x \in X$, temos:

$<x|x^*> = <A^{-1}(y)|x^*> = <y|(A^{-1})^*(x^*)>$ ou seja
$<x|x^*> = <A(x)|(A^{-1})^*(x^*)> = <x|A^*[(A^{-1})^*(x^*)]>$

Logo $x^* = A^*[(A^{-1})^*(x^*)]$, o que significa que existe $y^* \in Y^*$ tal que

$$x^* = A^*(y^*), \text{ sendo } y^*=(A^{-1})^*(x^*)$$

Enfim, da prova de sobrejetividade concluimos também que

$$A^*(A^{-1})^* = \mathcal{I} \quad \text{ou seja} \quad (A^*)^{-1} = (A^{-1})^*. \quad \text{c.q.d.}$$

Vamos agora dar exemplos simples de adjuntos

<u>Exemplo 15</u> Seja $X=\mathbb{R}^n$ e $Y=\mathbb{R}^m$

Como sabemos, nesse caso podemos identificar X^* a X e Y^* a Y.

Como convencionado no parágrafo precedente vamos escrever:

$$<x|y> = (x|y)$$

Seja agora A um operador linear contínuo de X em Y que convencionamos igualmente considerar como sendo a matriz m×n associada.

Assim, o adjunto de A será dado por

$$(Ax,y)_m = (x|A^*y)_n \quad \forall x \in \mathbb{R}^n \text{ e } \forall y \in \mathbb{R}^m, \text{ onde}$$

$(. | .)_m$ e $(. | .)_n$ representam, respectivamente, o produto escalar em $\mathbb{R}^m$ e $\mathbb{R}^n$.

Porém:

$$(Ax|y)_m = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) y_i$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}y_i = (x|z)_n$$

z é o vetor de $\mathbb{R}^n$ cuja j-ésima componente é o produto escalar da j-ésima coluna de A por y, ou seja da j-ésima linha de A^T matriz transposta de A, por y. Nesse caso portanto $z = A^T y$

Logo

$(Ax|y)_m = (x|A^T y)_n$ $\forall x,y$, donde se conclui que $A^* = A^T$

<u>Exemplo 16</u> Seja $Y = X = \ell^p$, $1 \le p < \infty$ e A o operador de acréscimo de uma componente:

$$A(x) = (0, x_1, x_2, \ldots)$$

Claramente $A \in \mathcal{L}(X,Y)$ com $\| A \| = 1$

Agora com $y^* \in \ell^q$, $\frac{1}{p} + \frac{1}{q} = 1$, temos:

$$< A(x) | y^* > = \sum_{i=2}^{\infty} x_{i-1} y_i^* = \sum_{i=1}^{\infty} x_i x_i^* = < x | x^* >$$

onde $x^* = (x_1^*, x_2^*, \ldots) \in \ell^q$ com $x_i^* = y_{i+1}^*$

Logo o operador adjunto A^* é o operador de supressão de uma componente, isto é:

$$A^*(y^*) = (y_2^*, y_3^*, y_4^*, \ldots)$$

Como no Exemplo 15, quando X e Y são espaços de Hilbert, definimos preferivelmente o adjunto por intermédio do produto escalar, identificando esses espaços aos duais respectivos pelo Teorema de Riesz. Assim se $A \in \mathcal{L}(X,Y)$ temos:

$$(A(x)|y)_X = (x|A^*(y))_Y \qquad \forall x \in X, \quad \forall y \in Y.$$

Uma conseqüência imediata dessa caracterização do adjunto, é que para espaços de Hilbert $A^{**}=A$.

No caso ainda mais particular em que X=Y, o operador A^* pertence ao mesmo espaço que A e podemos considerar a igualdade $A^*=A$.

<u>Definição 11</u>: Se X é um espaço de Hilbert e $A \in \mathcal{L}(X,X)$ é tal que $A^*=A$, então A é dito <u>auto-adjunto</u>.

Se no Exemplo 15 supomos m=n, a matriz A será um operador auto-adjunto no caso em que for simétrica.

Como provaremos a seguir, uma classe importante de operadores auto-adjuntos é a constituída pelos operadores projeção:

<u>Definição 12</u>: Seja X um espaço de Hilbert e M um seu subespaço fechado. O operador linear (sobrejetivo)

$$P: X \to M$$

que associa a $x \in X$ à sua componente m da decomposição única

$$x=m+n, \quad m \in M \quad e \quad n \in M^{\perp}$$

é dito operador projeção de X sobre M.

O operador projeção é contínuo pois, de acordo com o Teorema de Pitágoras,

$$\| x \|^2 = \| m \|^2 + \| n \|^2, \quad \text{Logo}$$

$$\| P(x) \| = \| m \| \le \| x \| \qquad \forall x \in X.$$

Um operador projeção tem ainda a interessante propriedade de ser idempotente de acordo com a

<u>Definição 13</u>: Um operador $A:X \to X$ é dito <u>idempotente</u> se $A^2=A$.

Observe-se que se um operador é idempotente $A^m = A$ para todo m inteiro, $m>0$.

Vamos agora provar o seguinte resultado.

<u>Proposição 9</u>: Um operador contínuo: $P: X \to X$ é um operador projeção se e somente se for auto-adjunto e idempotente.

<u>Demonstração</u>: Inicialmente se P é operador projeção temos:

$$P(x)=m \quad \forall x \in X \quad e \quad P[P(x)]= P(m)=m$$

Logo $P^2 = P$.

Além disso, se $x=m+n$ e $y=p+q$ $m,p \in M$, $n,q \in M^{\perp}$ sendo M subespaço fechado de X temos:

$$(P(x)|y) = (m|p+q) = (m|p) + (m|q) = (m|p)$$

Porém $(x|P(y)) = (m+n|p) = (m|p) + (n|p) = (m|p)$

Logo $P=P^*$.

Agora seja P tal que $P=P^*$ e $P^2=P$

$$P: X \to X$$

Seja P(X) a imagem de P. Provemos que se $x=m+n$ onde $m \in \overline{P(X)}$ e $n \in \overline{P(X)}^{\perp}$ então $P(x)=m$, o que mostrará não só que P é projeção sobre $\overline{P(X)}$ como também que P(X) é fechado.

Temos: $P(n) \in P(X)$. Logo $(P(n)|n)=0 \quad \forall n$.

Como $P=P^2$ então $(P^2(n)|n)=(P(n)|P(n))=0 \Rightarrow \| P(n) \|=0$, ou seja $P(n) = \vec{0} \qquad \forall n \in \overline{P(X)}^{\perp}$

Resta provar que $P(x)=m \quad \forall x \in X$

Temos $P(x)=P(m)$ pois P é linear e $P(n)=\vec{0}$.

Agora $\forall m \in \overline{P(X)}$ existe uma seqüência $\{x_n\} \subset X$ tal que

$$P(x_n) \to m$$

Como P é contínuo $P^2(x_n) \to P(m)$.

Porém $P^2=P$, donde $P(x_n) \to P(m)$.

Logo, pela unicidade dos limites $P(m)=m$ ou seja $P(x)=m$. Como qualquer $m \in \overline{P(X)}$ está na imagem $P(X)$ temos $P(X)=\overline{P(X)}$. c.q.d.

Para espaços de Hilbert podemos estabelecer relações simples e úteis entre imagens e espaços nulos de um operador e seu adjunto, respectivamente.

Teorema 6: Sejam X e Y espaços de Hilbert e $A \in \mathcal{L}(X,Y)$. Então

$$[A(X)]^{\perp} = \mathcal{N}(A^*)$$

Demonstração: Seja $y^* \in \mathcal{N}(A^*) \subset Y^*$ e $y \in A(X) \subset Y$.

Então $y=A(x)$ para algum $x \in X$ e

$$(y|y^*) = (A(x)|y^*) = (x|A^*(y^*)) = 0 \Rightarrow$$

$$\mathcal{N}(A^*) \subset [A(X)]^{\perp}$$

Suponhamos agora $y^* \in [A(X)]^{\perp}$. Temos:

$$(y^*|y) = 0 \quad \forall y \in A(X) \Rightarrow (y^*, A(x))=0 \quad \forall x \in X$$

Logo $(A^*(y^*)|x)=0 \quad \forall x \in X \Rightarrow A^*(y^*)=\vec{0}_Y$

ou seja, $y^* \in \mathcal{N}(A^*)$ e $[A(X)]^{\perp} \subset \mathcal{N}(A^*)$. c.q.d.

Exemplo 17: Seja $A: \mathbb{R}^n \to \mathbb{R}^m$. A consiste de n colunas de vetores $\vec{a}_i$, $i=1,2,\ldots,n$ de $\mathbb{R}^m$.

$A(\mathbb{R}^n)$ é constituida das combinações lineares da forma

$$\sum_{i=1}^{n} c_i \vec{a}_i$$

Logo é o subespaço de $\mathbb{R}^m$ gerado pelos $\vec{a}_i$: $A(\mathbb{R}^n)=[\vec{a}_1,\vec{a}_2,\ldots,\vec{a}_n]$. $[A(\mathbb{R}^n)]^{\perp}$ consiste dos vetores $y \in \mathbb{R}^m$ tais que $(y|\vec{a}_i)=0 \quad \forall_i$.

Por outro lado A^*, que é a transposta de A, tem os $\vec{a}_i$ como linhas. Logo um vetor $y \in \mathbb{R}^m$ será de $\mathcal{N}(A^*)$ se seu produto escalar por cada linha resultar em zero isto é: $(y|\vec{a}_i)=0 \quad \forall i$, ou seja reencontramos a mesma condição.

<u>Corolário 4</u>: Se X e Y são espaços de Hilbert vale

1. $\overline{A(X)} = [\mathcal{N}(A^*)]^\perp$
2. $[A^*(Y)]^\perp = \mathcal{N}(A)$
3. $\overline{A^*(Y)} = [\mathcal{N}(A)]^\perp$

<u>Demonstração</u>: 1. decorre de $[A(X)]^{\perp\perp} = \overline{A(X)}$

2. Aplicamos o Teorema 6 a $A^* \in \mathcal{L}(Y,X)$ e temos $[A^*(Y)]^\perp = \mathcal{N}(A^{**}) = \mathcal{N}(A)$

3. Tomamos em 2. o complemento ortogonal de ambos os membros. c.q.d.

<u>Exemplo 18</u>: Usando p=2 no Exemplo 16 temos para $y=(y_1,y_2,\ldots)$

$A^*(y)=(y_2,y_3,\ldots)$. Logo $\mathcal{N}(A^*)$ é o conjunto de seqüências que só podem ter a primeira componente não nula.

$[\mathcal{N}(A^*)]^\perp$ é pois o conjunto de seqüências cuja 1ª componente é nula.

Como A é tal que $A(x)=(0,x_1,x_2,\ldots)$ vemos que A(X) é também o conjunto de seqüências cuja primeira componente é nula.

Observe-se que esse conjunto é fechado e temos pois

$$A(X) = \overline{A(X)} = [\mathcal{N}(A^*)]^\perp .$$

BIBLIOGRAFIA

[1] Bertrandias, J.P. - Mathématiques pour l'Informatique Vol. 1 - (Analyse Fonctionnelle), Armand Colin, Paris, 1970.

[2] Fernandez, P.J. - Medida e Integração, Instituto de Matemática Pura e Aplicada, CNPq (Projeto Euclides), Rio de Janeiro, 1976.

[3] Kolmogórov, A.N., Fomín, S.V. - Elementos de la Teoria de Funciones y del Análisis Funcional, Editorial MIR, Moscú, 1972.

[4] Kreider L., Kuller R.G., Ostberg D.R., Perkins F.W. - Addison-Wesley Publishing Company INC, Reading, Massachusetts, USA, 1966.

[5] Lichnerowicz, A - Linear Algebra & Analysis, Holden - Day- Inc., San Francisco, 1967.

[6] Luenberger, D.G. - Optimization by Vector Space Methods, John Wiley & Sons, Inc., New York, 1969.

[7] Lusternik, L.A., Sobolev, V.J. - Elements of Functional Analysis, John Wiley & Sons, New York, 1974.

[8] Nowosad, P. - Introdução à Análise Funcional, Instituto de Matemática da UFPe, Recife, 1969.

[9] Ruas de B.S.V. - Introdução aos Problemas Variacionais - Editora Guanabara Koogan S/A, Rio de Janeiro (a sair) em 1978.

[10] Schechter, M. - Principles of Functional Analysis, Acad. Press., New York, 1971.

[11] Shilov, G.E. - Elementary Functional Analysis, The MIT Press, Cambridge, Mass., 1974.

[12] Weir, A.J. - Lebesque Integration and Measure, Cambridge University Press, Cambridge, 1974.

[13] Wilansky, A. - Functional Analysis, Blaisdell Publishing Company, N.Y, Toronto, London, 1964.