



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Sistemas dinâmicos que triangularizam matrizes

por
Joaby de Souza Jucá

Brasília
2012

Agradecimentos

Além do apoio financeiro da CAPES, muitas pessoas foram essenciais para que este trabalho fosse concluído com o devido sucesso e quero registrar aqui meus agradecimentos.

Aos meus pais, pelos princípios agregados ao meu caráter e amor incondicional.

Ao amigo Luiz Fernando Frozza, sempre presente, um irmão.

Ao meu orientador, Lucas Seco, pelo cuidado por meu aprendizado em matemática, sem poupar sua experiência ou seu tempo. Também por sua contribuição a minha organização e pragmatismo.

Ao Professor Adán Corcho, que me orientou na iniciação científica na UFAL (Universidade Federal de Alagoas), cujos exemplos de profissionalismo e conselhos motivam meus estudos ainda hoje.

Ao Professor Adriano Aguiar, que orientou meu trabalho de conclusão de curso na UFAL, cujo rigor matemático aprimorou meus estudos.

A Thaynara, minha flor, que traz suavidade e alegria para minha vida.

Aos Professores Mauro Patrão e Alexandre Santana por lerem esta dissertação e por suas significativas sugestões de melhoria, além de participarem da banca examinadora de minha defesa da mesma.

Aos demais que fazem parte do meu cotidiano, cuja representação é tão importante quanto a da família. Em especial ao Igor, que abriu as portas para que eu viesse à UnB.

Aos demais funcionários do Departamento de Matemática da UnB, por todo o suporte para minha formação.

Agradeceria ainda a Deus, não fosse o senso comum à Verdade que existe em cada um.

Resumo

Nesta dissertação procuramos entender por meio de sistemas dinâmicos a convergência de um método iterativo que triangulariza e encontra os autovalores de uma matriz complexa, a saber, o método QR da análise numérica. Utilizamos apenas ferramentas de álgebra linear, cálculo em várias variáveis e ações de grupos topológicos.

Vemos também que o método QR tem um análogo de tempo contínuo, dado por uma EDO matricial que triangulariza a condição inicial. Quando a condição inicial é hermitiana, verificamos que a altura da solução com respeito a certas matrizes diagonais é uma função de Lyapunov do fluxo.

Palavras-chave: Método QR, autovalores, grupos de matrizes, ação de grupos topológicos, decomposição de Bruhat, variedade flag, sistemas dinâmicos.

Abstract

In this work we seek to understand through dynamical systems the convergence of an iterative method that triangularizes and finds the eigenvalues of a complex matrix, the so called QR method of numerical analysis. We only use tools from linear algebra, multivariate calculus and topological group actions.

We also see that the QR method has an continuous time analog, given by a matrix differential equation which triangularizes the initial condition. When the initial condition is hermitian, we check that the height of the solution with respect to certain diagonal matrices is a Lyapunov function for the flow.

Keywords: QR method, eigenvalues, matrix groups, topological group actions, Bruhat decomposition, flag manifold, dynamical systems.

Sumário

1	Introdução	1
1.1	Triangularização de matrizes e dinâmica	1
1.2	Base de Schur via ponto fixo	2
1.3	Base de Schur ortogonal via iteração	3
1.4	Forma de Schur via método QR	4
1.5	Forma de Schur via equações diferenciais	7
1.6	Perspectivas	8
1.7	Contribuições da dissertação	8
1.8	Estrutura da dissertação	9
2	Base de Schur via iterações	11
2.1	Base de Schur e ponto fixo	11
2.2	Decomposição de Bruhat (LPU)	12
2.3	Topologia da variedade flag maximal	20
2.4	Iterações em \mathbb{F}	25
2.5	Iterações em Q	27
3	Forma de Schur via método QR	33
3.1	Método QR	33
3.2	Equação diferencial	36
A	Tópicos Matriciais	49
A.1	Operações elementares e matrizes de permutação	49
A.2	Conjugações que fixam a diagonal	51
A.3	Decomposição de Gauss (PLU)	53
A.4	Decomposição de Iwasawa (QR)	55
A.5	Forma de Schur	57
A.6	Algumas funções matriciais	58
B	Topologia	63
B.1	Espaços Topológicos	63
B.2	Aplicações contínuas	64
B.3	Topologia dos Espaços Homogêneos	65
B.4	Fluxos e sistemas dinâmicos	70

Lista de notações

Nesta dissertação, todos os espaços vetoriais são complexos e de dimensão finita.

- $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$ espaço vetorial das matrizes $n \times n$ sobre \mathbb{C} .
- $G = \text{Gl}(n, \mathbb{C})$ *grupo linear geral*, constituído pelas matrizes invertíveis de \mathfrak{g} .
- I matriz identidade.
- E_{ij} matriz de \mathfrak{g} cuja ij -ésima entrada é 1 e as demais são nulas.
- X, Y elementos de \mathfrak{g} .
- g, h elementos de G .
- $(v_1 | \cdots | v_n)$ matriz em \mathfrak{g} cujas colunas são os vetores $v_1, \dots, v_n \in \mathbb{C}^n$.
- $X[k] \in \mathfrak{gl}(k, \mathbb{C})$ *k-menor principal* de $X \in \mathfrak{g}$, constituída pelas k primeiras colunas de X truncadas até a linha k .
- $\text{diag}(X)$ matriz diagonal cujas entradas são as respectivas entradas da diagonal de X .
- $\text{diag}(\lambda_1, \dots, \lambda_n)$ matriz diagonal com entrada λ_i na i^{a} linha e i^{a} coluna.
- $\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}$ exponencial de X .
- $X^\dagger = (X_{ji})$ *transposta* de $X = (X_{ij})$.
- $X^* = \overline{X}^\dagger$ *transposta conjugada* ou *adjunta* de $X = (X_{ij})$, onde $\overline{X} = (\overline{X_{ij}})$.
- $U = \{(g_{ij}) \in G : g_{ij} = 0 \text{ para } i > j\}$ subgrupo das matrizes triangulares superiores invertíveis.
- $L = \{(g_{ij}) \in G : g_{ij} = 0 \text{ para } i < j\}$ subgrupo das matrizes triangulares inferiores invertíveis.
- $Q = \{g \in G : g^*g = I\}$ subgrupo *unitário* de G .
- $R = \{g \in U : g_{ii} > 0 \forall i\}$ subgrupo das matrizes triangulares superiores invertíveis com entradas reais e positivas na diagonal.
- $\mathfrak{q} = \{X \in \mathfrak{g} : X^* = -X\}$ subespaço das matrizes *anti-hermitianas*.
- $\mathfrak{t} = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ se } i > j \text{ e } X_{ii} \in \mathbb{R} \text{ para todo } i\}$ subespaço das matrizes triangulares superiores com entradas reais na diagonal.
- $\mathfrak{u} = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ para } i > j\}$ subespaço das matrizes triangulares superiores.
- $\mathfrak{l} = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ para } i < j\}$ subespaço das matrizes triangulares inferiores.

- $\mathfrak{d} = \{X \in \mathfrak{g} : X_{ij} = 0, \text{ se } i \neq j\}$ conjunto das matrizes diagonais.
- $\mathfrak{n}^- = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ se } i \leq j\}$ conjunto das matrizes triangulares estritamente inferiores.
- $\mathfrak{n}^+ = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ se } i \geq j\}$ conjunto das matrizes triangulares estritamente superiores.
- $S_n = \{\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\} : \sigma \text{ é bijeção}\}$ conjunto das permutações de n elementos.
- I_σ , com $\sigma \in S_n$, *matriz de permutação*, elemento em g definido por $I_\sigma e_i = e_{\sigma(i)}$ para todo $i = 1, \dots, n$.
- Ad *adjunta* (grande), ou *conjugação*, é uma aplicação de G em $\text{Gl}(\mathfrak{g})$ definida por $\text{Ad}(g)X = gXg^{-1}$, para todo $g \in G$ e $X \in \mathfrak{g}$.
- $[X, Y] = XY - YX$ *colchete* ou *comutador* de X por Y .
- ad *adjunta* (pequena), ou *colchete*, é uma aplicação de \mathfrak{g} em $\mathfrak{gl}(\mathfrak{g})$ definida por $\text{ad}(X)Y = [X, Y]$, para todo $X, Y \in \mathfrak{g}$.
- $\mathbb{F} = \{(V_1, \dots, V_n) : V_1 \subset \dots \subset V_n \text{ são subespaços de } \mathbb{C}^n \text{ com } \dim V_i = i, \forall i\}$ *variedade flag maximal* de \mathbb{C}^n .
- $[u_1, \dots, u_k]$ subespaço gerado pelos vetores $u_1, \dots, u_k \in \mathbb{C}^n$.

Capítulo 1

Introdução

Neste primeiro capítulo vamos motivar e introduzir os assuntos tratados nesta dissertação, incluindo perspectivas sobre continuação do trabalho e uma breve revisão bibliográfica. Ao final, vamos expor a estrutura dos demais capítulos.

1.1 Triangularização de matrizes e dinâmica

Encontrar todos os autovalores de uma matriz é uma questão prática de fundamental interesse. Qual é um método prático para obter todos os autovalores de uma matriz complexa $n \times n$? Como os autovalores correspondem às raízes de um polinômio de grau n (o polinômio característico da matriz), em geral não existem métodos exatos. Devemos então nos contentar em procurar métodos iterativos que encontram os autovalores por aproximações sucessivas.

Uma matriz $n \times n$ complexa X sempre possui um autovetor v_1 , com autovalor $\lambda_1 \in \mathbb{C}$,

$$Xv_1 = \lambda_1 v_1,$$

o que equivale, geometricamente, a uma direção $V_1 = [v_1] = \mathbb{C}v_1$ invariante por X ,

$$XV_1 \subset V_1.$$

Mas X nem sempre possui uma base de autovetores. A *forma de Schur* de uma matriz $n \times n$ complexa X (ver Seção A.5) nos diz que sempre existe uma base ordenada $b \in \text{Gl}(n, \mathbb{C})$ na qual a matriz de X é triangular superior, isto é,

$$b^{-1}Xb = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ & & \lambda_n \end{pmatrix},$$

onde os vetores da base são as colunas de b . Uma base b com essa propriedade será chamada de *base de Schur* para X e a matriz triangular superior $b^{-1}Xb$, uma *forma de Schur* para X . Triangularizar X é encontrar uma forma ou uma base de Schur para X .

Uma base de Schur b para X sempre pode ser ortogonalizada, via processo de Gram-Schmidt, para se obter uma base de Schur ortogonal q para X que está no grupo Q

das matrizes unitárias. Com isso, teremos que as formas de Schur para X são todas unitariamente conjugadas à X .

Os autovalores de X aparecem todos na diagonal da forma de Schur de modo que, uma maneira de encontrar simultaneamente todos os autovalores de X é encontrar uma base ou uma forma de Schur para X . Isto motiva a pergunta: Qual é um método iterativo para triangularizar X ?

Observação 1.1. *Seja X hermitiana, isto é, $X^* = X$. É fácil ver que toda forma de Schur para X é diagonal com entradas reais. Segue que X tem autovalores reais, que encontrar uma forma de Schur para X é diagonalizar X e que encontrar uma base de Schur para X é encontrar uma base de autovetores para X .*

1.2 Base de Schur via ponto fixo

Se a matriz g é invertível, uma maneira natural de procurar uma base de Schur para g é a seguinte. Geometricamente, uma base na qual g é triangular superior corresponde a um flag maximal de subespaços g -invariantes (ver Seção 2.1), isto é, uma sequência encaixante de subespaços de \mathbb{C}^n

$$(V_1, \dots, V_n), \quad \text{com } V_i \subset V_{i+1} \quad \text{e} \quad \dim V_i = i,$$

para os quais

$$gV_i = V_i, \quad i = 1, \dots, n.$$

Para procurar flags g -invariantes, consideramos o conjunto \mathbb{F} de todos os flags maximais de \mathbb{C}^n e a ação de $G = \text{Gl}(n, \mathbb{C})$ em \mathbb{F} dada por

$$g(V_1, \dots, V_n) = (gV_1, \dots, gV_n), \quad (1.1)$$

onde tomamos g invertível para que preserve a dimensão dos subespaços. Assim, uma base de Schur para $g \in G$ corresponde a um ponto fixo x' de g em \mathbb{F}

$$x' = (V_1, \dots, V_n) \in \mathbb{F}$$

$$gx' = x' \iff gV_i = V_i, \quad i = 1, \dots, n.$$

A forma de Schur garante que a ação de g em \mathbb{F} possui pontos fixos. Como encontrá-los?

Um método iterativo natural para procurar pontos fixos é dinâmico: o limite da iteração $g^t x$, quando $t \in \mathbb{Z}$ tende a infinito, é ponto fixo de g , desde que o limite exista. Isto é, sempre que

$$\lim_{t \rightarrow \infty} g^t x = x' \quad (1.2)$$

teremos

$$gx' = g \left(\lim_{t \rightarrow \infty} g^t x \right) = \lim_{t \rightarrow \infty} g^{t+1} x = x',$$

de modo que x' é ponto fixo de g .

Para que essa ideia funcione de fato, devemos primeiro introduzir em \mathbb{F} uma topologia, para podermos falar de limites, e essa topologia deve ser tal que a ação de g

seja contínua e limites de sequências sejam únicos. Existe em \mathbb{F} uma topologia natural, e essencialmente única, na qual a ação (1.1) de G em \mathbb{F} é contínua e \mathbb{F} é Hausdorff (ver Seção 2.3). Logo as propriedades desejadas são satisfeitas, bastando-nos garantir condições suficientes para que a convergência (1.2) aconteça. Essas condições serão dadas pelo seguinte resultado, que será demonstrado na Seção 2.4.

Teorema 1.2. *Suponha que os módulos dos autovalores de g são todos distintos. Então, para todo $x \in \mathbb{F}$, temos que $g^t x$ converge para um flag maximal g -invariante quando $t \rightarrow \infty$, com $t \in \mathbb{Z}$.*

Com o método iterativo acima não obtemos de imediato uma base de Schur para g , mas sim um flag g -invariante x' . Dada uma base ordenada $b \in \text{Gl}(n, \mathbb{C})$, podemos formar um flag maximal $\pi(b) \in \mathbb{F}$ com suas colunas, considerando os subespaços encaixantes gerados por sua primeira coluna, por sua primeira e segunda colunas, etc. Dizemos que uma base ordenada b é *adaptada ao flag x* se $\pi(b) = x$. Obtemos assim uma projeção sobrejetora

$$\pi : \text{Gl}(n, \mathbb{C}) \rightarrow \mathbb{F}$$

cuja fibra $\pi^{-1}(x)$ sobre $x \in \mathbb{F}$ é o conjunto das bases adaptadas ao flag x . Se $x' \in \mathbb{F}$ é um flag g -invariante, segue que qualquer matriz na fibra $\pi^{-1}(x')$ é uma base de Schur para g .

Observação 1.3. *A dinâmica da iteração de g aparece em diversos contextos de matemática pura e aplicada (ver a monografia [5]) e uma ferramenta essencial para estudá-la é a chamada decomposição de Bruhat (ver Seção 2.2). Essa decomposição apareceu pela primeira vez no estudo da topologia de variedades flag clássicas por Ehresmann nos anos 1930 e foi generalizada por Bruhat nos anos 50 para grupos semisimples e suas variedades flag (para um histórico mais completo, ver a nota [10]). A relação dessa decomposição com a dinâmica da iteração na variedade flag surgiu nos anos 1980 e foi generalizada desde então (para um breve histórico, ver a revisão bibliográfica feita na introdução do artigo [2]).*

1.3 Base de Schur ortogonal via iteração

Uma maneira de procurar uma base ortogonal de Schur para g é a seguinte (ver Seção 2.5). Restringindo a projeção $\pi : G \rightarrow \mathbb{F}$ ao grupo Q das matriz unitárias, obtemos a projeção sobrejetora

$$\pi_Q : Q \rightarrow \mathbb{F},$$

cuja fibra $\pi_Q^{-1}(x)$ sobre $x \in \mathbb{F}$ é o conjunto das bases ortogonais adaptadas ao flag x . Se $x' \in \mathbb{F}$ é um flag g -invariante, então qualquer matriz na fibra $\pi_Q^{-1}(x')$ é uma base ortogonal de Schur para g .

Fazendo a ortogonalização de Gram-Schmidt nas colunas de uma matriz invertível $g \in \text{Gl}(n, \mathbb{C})$, podemos escrevê-la de forma única como o produto

$$g = qr,$$

onde $q \in Q$, o grupo das matrizes unitárias e $r \in R$, o grupo das matrizes triangulares superiores com diagonal positiva (ver Seção A.4). Essa é a chamada *decomposição QR*, ou *decomposição de Iwasawa* de g .

Temos que G age naturalmente em Q da seguinte maneira. Dada uma matriz unitária $q \in Q$, fazemos $g \in G$ agir em q formando o produto gq e ortogonalizando as colunas de gq obtendo assim a matriz unitária $Q(gq)$. Isso fornece uma ação de g em Q que cobre a ação de g em \mathbb{F} de modo que o seguinte diagrama comuta

$$\begin{array}{ccc} Q & \xrightarrow{g} & Q \\ \pi_Q \downarrow & & \downarrow \pi_Q \\ \mathbb{F} & \xrightarrow{g} & \mathbb{F} \end{array}$$

Uma vez que a iteração de g em \mathbb{F} converge para um flag g -invariante x' , segue que a iteração de g em Q converge para a fibra g -invariante $\pi_Q^{-1}(x')$, que consiste de bases ortogonais de Schur para g (ver Figura 1.1). Mais precisamente, temos o seguinte resultado, que será demonstrado na Seção 2.5 desta dissertação.

Teorema 1.4. Sejam $g \in G$ com autovalores de módulos distintos e $q \in Q$. Então, $g^t q$ converge para uma fibra g -invariante de π_Q .

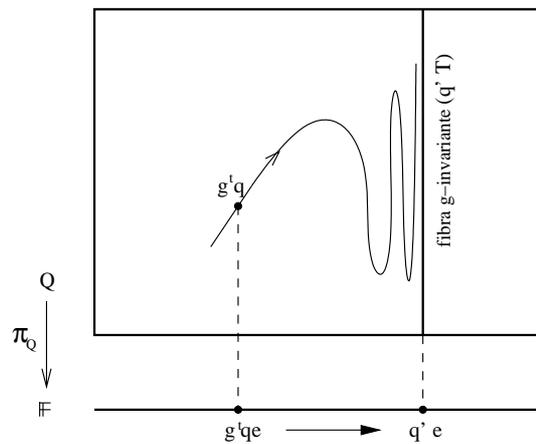


Figura 1.1: Iterações de g em Q

1.4 Forma de Schur via método QR

O método QR é um método iterativo da análise numérica que é usado para encontrar uma base ortogonal de Schur para uma matriz invertível. O método funciona da seguinte maneira.

Partindo de $g \in \text{Gl}(n, \mathbb{C})$, considere a seguinte sequência de matrizes invertíveis, obtida trocando-se a ordem dos fatores das sucessivas decomposições de Iwasawa, $q_t \in$

Q e $r_t \in R$, para todo $t \in \mathbb{N}$,

$$\begin{aligned}
 g_0 &= g &= q_0 r_0 \\
 g_1 &= r_0 q_0 &= q_1 r_1 \\
 g_2 &= r_1 q_1 &= q_2 r_2 \\
 &\vdots & \\
 g_{t+1} &= r_t q_t &= q_{t+1} r_{t+1} \\
 &\vdots &
 \end{aligned} \tag{1.3}$$

O método QR se baseia no seguinte resultado surpreendente, que será demonstrado na Seção 3.1 desta dissertação.

Teorema 1.5. Se os módulos dos autovalores de g são todos distintos, então g_t converge para uma forma de Schur para g , quando $t \rightarrow \infty$. Mais precisamente,

- (i) as entradas da diagonal de g_t convergem para os autovalores de g ,
- (ii) as entradas abaixo da diagonal de g_t convergem para zero,

Se, além disso, g é hermitiana, então g_t converge para uma matriz diagonal com os autovalores de g em suas entradas.

Exemplo 1.6. Vamos apresentar um exemplo de [11] que ilustra como as iterações do método QR aplicada a uma matriz hermitiana convergem para uma matriz diagonal. Aplicando o método QR à matriz hermitiana tridiagonal

$$g = g_0 = \begin{pmatrix} 4 & -3 & 0 & 0 \\ -3 & 2 & 3.16228 & 0 \\ 0 & 3.16228 & -1.4 & -0.2 \\ 0 & 0 & -0.2 & 1.4 \end{pmatrix},$$

obtemos

$$g_1 = \begin{pmatrix} 6.16 & 1.90116 & 0 & 0 \\ 1.90116 & -1.71378 & -2.43973 & 0 \\ 0 & -2.43973 & 0.164185 & -0.114069 \\ 0 & 0 & -0.114069 & 1.3896 \end{pmatrix}$$

⋮

$$g_{17} = \begin{pmatrix} 6.64575 & 0.000115629 & 0 & 0 \\ 0.000115629 & -3.64575 & -9.2563 \times 10^{-6} & 0 \\ 0 & -9.2563 \times 10^{-6} & 1.64567 & -0.00483899 \\ 0 & 0 & -0.00483899 & 1.35433 \end{pmatrix}$$

⋮

$$\begin{aligned}
g_{23} &= \begin{pmatrix} 6.64575 & 0 & 0 & 0 \\ 0 & -3.64575 & 0 & 0 \\ 0 & 0 & 1.64574 & -0.00150269 \\ 0 & 0 & -0.00150269 & 1.35426 \end{pmatrix} \\
&\vdots \\
g_{52} &= \begin{pmatrix} 6.64575 & 0 & 0 & 0 \\ 0 & -3.64575 & 0 & 0 \\ 0 & 0 & 1.64575 & -5.2671 \times 10^{-6} \\ 0 & 0 & -5.2671 \times 10^{-6} & 1.35425 \end{pmatrix} \\
g_{53} &= \begin{pmatrix} 6.64575 & 0 & 0 & 0 \\ 0 & -3.64575 & 0 & 0 \\ 0 & 0 & 1.64575 & 0 \\ 0 & 0 & 0 & 1.35425 \end{pmatrix} \\
&\vdots
\end{aligned}$$

É importante frisar que esse exemplo usa a representação dos números reais em um computador de modo que as entradas nulas nas matrizes significam, na verdade, entradas muito pequenas.

Mostraremos, usando as idéias do artigo [16], que o método QR pode ser compreendido com dinâmica. Denote por $Q(g)$ e $R(g)$ as componentes em Q e R de g na decomposição QR, respectivamente, de modo que $g = Q(g)R(g)$. Ponha $G = \text{Gl}(n, \mathbb{C})$ e considere o homeomorfismo

$$\Phi : G \rightarrow G, \quad g \mapsto R(g)Q(g). \quad (1.4)$$

Então, a sequência g_t do método QR é dada pela iteração

$$g_t = \Phi^t(g), \quad t \in \mathbb{N},$$

que é um fluxo de tempo discreto em G . Veremos que as iteradas de Φ começando em g convergem, a menos de conjugação por diagonais unitárias, para uma forma de Schur de g precisamente porque temos o seguinte resultado, que será demonstrado na Seção 3.1.

Teorema 1.7. Temos que

$$\Phi^t(g) = Q(g^t)^{-1}gQ(g^t), \quad t \in \mathbb{Z}.$$

Como vimos que $Q(g^t)$ converge, a menos de multiplicação à direita por matrizes diagonais unitárias, para uma base ortogonal de Schur para g , o resultado acima explica porque $\Phi^t(g)$ converge grosso modo a uma forma de Schur para g .

Observação 1.8. *O método QR foi criado nos anos 60 e é a base para métodos modernos de cálculo de autovalores e autovetores até hoje, devido a sua simplicidade, rapidez de convergência e robustez na propagação de erros numéricos (para isso ver, por exemplo, [4] e referências lá contidas). Nesta dissertação não vamos analisar essas características do método, mas queremos compreender o que está por trás do método QR.*

Além de satisfazer a nossa curiosidade, essa compreensão pode ser a base para a formulação de novos métodos práticos para o cálculo numérico de autovalores e autovetores (ver, por exemplo [7]). Além disso, o método QR possui conexões inusitadas com temas de matemática pura como, por exemplo, os chamados sistemas integráveis. Essa conexão foi descoberta nos anos 80, é explorada até hoje e será brevemente considerada aqui (ver Seção 1.6).

1.5 Forma de Schur via equações diferenciais

Outro fato surpreendente sobre o método QR é que seus passos são a discretização da solução de uma certa EDO (ver Seção 3.2.3). De fato, vimos que o método QR fornece a sequência

$$\Phi^k(g) = Q(g^k)^{-1}gQ(g^k), \quad k \in \mathbb{Z},$$

que converge grosso modo para uma forma de Schur para g quando $k \rightarrow \infty$. Escrevendo $g = \exp(X)$ com $X \in \mathfrak{gl}(n, \mathbb{C})$, para $k \in \mathbb{Z}$ temos que

$$g^k = \exp(kX)$$

e que

$$\Phi^k(g) = \exp(Q(g^k)^{-1}XQ(g^k)).$$

Consideramos então as curvas, para $t \in \mathbb{R}$, dadas por

$$g^t = \exp(tX),$$

$$\phi^t(X) = Q(g^t)^{-1}XQ(g^t) = \text{Ad}(Q(g^t)^{-1})X, \quad (1.5)$$

de modo que

$$\Phi^k(\exp(X)) = \exp(\phi^k(X)), \quad k \in \mathbb{Z}.$$

Veremos que, como era de se esperar, a curva $\phi^t(X)$ converge, a menos de conjugação por diagonais unitárias, para uma forma de Schur para X , quando $t \rightarrow \infty$ em \mathbb{R} , desde que X tenha autovalores com partes reais distintas. Mais interessante é o seguinte resultado, que será demonstrado na Seção 3.2.3 desta dissertação.

Teorema 1.9. A aplicação $\phi^t(X)$ é um fluxo diferenciável em $\mathfrak{gl}(n, \mathbb{C})$ com campo

$$F(X) = -\text{ad}(\mathfrak{q}(X))X = [X, \mathfrak{q}(X)], \quad (1.6)$$

onde $\mathfrak{q}(x)$ é a projeção de X no subespaço das matrizes anti-hermitianas de \mathfrak{g} .

Isso mostra que os passos do método QR começando em $g = \exp(X)$ vêm da discretização da solução da EDO

$$X'(t) = [X(t), \mathfrak{q}(X(t))], \quad X(0) = X, \quad (1.7)$$

cujas soluções $X(t)$ são unitariamente conjugadas à condição inicial X e convergem grosso modo para uma forma de Schur para X quando $t \rightarrow \infty$. Em outras palavras, essa é uma equação diferencial em $\mathfrak{gl}(n, \mathbb{C})$ que triangulariza a condição inicial.

Além disso, vamos mostrar (ver Seção 3.2.4) que, se a condição inicial é hermitiana, a altura do fluxo $X(t)$ com respeito a uma matriz diagonal positiva decresce a medida que t cresce (ver Figura 1.2).

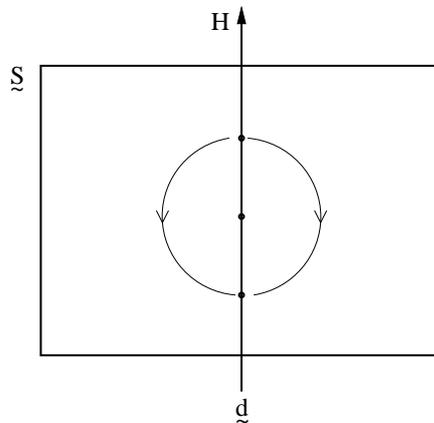


Figura 1.2: Altura de $X(t)$ com respeito à H

1.6 Perspectivas

A EDO (1.7) é conhecida como *equação de Lax*, ou *fluxo isoespectral*. Esse tipo de EDO e suas propriedades têm sido generalizados em várias direções (ver, por exemplo o artigo [6] e [14]).

Em particular, essa EDO está relacionada com uma ampla classe dos chamados sistemas integráveis descoberta nos anos 1980. Por exemplo, os autovalores da condição inicial X são quantidades que se conservam ao longo da solução $X(t)$. Algumas equações diferenciais da física se transformam numa equação de duplo-colchete por uma mudança adequada de variáveis. Assim, o método QR fornece um método numérico para integrar essas EDOs de uma maneira que acumula poucos erros a medida que t cresce. Para mais detalhes sobre a relação do método QR com sistemas integráveis, consulte a revisão bibliográfica feita na introdução do artigo [14].

1.7 Contribuições da dissertação

Esta dissertação é baseada no artigo de Shub e Vasquez [16] e suas contribuições tem um caráter expositório. Um dos objetivos foi desenvolver esse assunto de maneira

concreta usando apenas ferramentas de álgebra linear, cálculo em várias variáveis e ações de grupos topológicos, sem apelar para ferramentas mais avançadas de Teoria de Lie. Assim como Shub e Vasquez, trabalhamos com espaços vetoriais complexos para simplificar a exposição, porém, diferentemente deles, nos concentramos nos conceitos de forma e base de Schur e na dinâmica na variedade flag maximal. Aqui construímos a decomposição de Bruhat de $Gl(n, \mathbb{C})$ e de sua variedade flag maximal de um modo elementar (Seção 2.2), algo apenas esboçado nos artigos [16, 17] e no apêndice do livro [3], no qual nos baseamos. Também expusemos de modo elementar a topologia da variedade flag maximal \mathbb{F} de \mathbb{C}^n , em especial a construção da seção contínua de um aberto denso de \mathbb{F} em $Gl(n, \mathbb{C})$ (ver Seção 2.2). Mostramos diretamente que a versão contínua do método QR possui a propriedade do fluxo, o que tornou a obtenção da EDO correspondente mais transparente que no artigo de Shub e Vasquez. Por último, mostramos de modo elementar que a altura da solução com respeito a certas matrizes diagonais é uma função de Lyapunov desse fluxo (ver Seção 3.2).

1.8 Estrutura da dissertação

No Capítulo 2, veremos como obter uma base de Schur para g por meio de uma iteração de g na variedade flag maximal \mathbb{F} , demonstrando o Teorema 1.2. Nossa ferramenta principal será a chamada decomposição de Bruhat de G que mostraremos ser equivalente a uma decomposição de \mathbb{F} como união disjunta de uma quantidade finita das chamadas células de Bruhat. Na última seção veremos a relação entre a ação de g no conjunto das matrizes unitárias Q e a ação de g em \mathbb{F} , o que nos fornecerá o Teorema 1.4.

No Capítulo 3 vamos mostrar a relação entre as iteradas do método QR iniciada em g e a iteração de g em Q , obtendo o Teorema 1.7. A partir disso e dos resultados do capítulo anterior, vamos obter o Teorema 1.5 que fornece a convergência do método QR. Em seguida, vamos considerar uma versão do método QR em tempo contínuo mostraremos que ela é solução da EDO do Teorema 1.9 e que essa EDO de fato triangulariza a condição inicial. Quando a condição inicial é hermitiana, veremos que a altura da solução com respeito a certas matrizes diagonais é uma função de Lyapunov do fluxo.

Para uma melhor leitura, escrevemos dois apêndices que se encontram no final desta dissertação. O Apêndice A, de caráter mais algébrico organiza as principais ferramentas matriciais aqui utilizadas. Já o Apêndice B, de caráter topológico, tem o objetivo de introduzir a topologia dos espaços homogêneos, crucial para uma boa compreensão da dinâmica das iteradas do método QR.

Capítulo 2

Base de Schur via iterações

Neste capítulo, relembremos a noção de forma e base de Schur e apresentamos métodos iterativos para encontrar uma base de Schur. Para economizar notação, usaremos $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$ e $G = \mathrm{Gl}(n, \mathbb{C})$.

2.1 Base de Schur e ponto fixo

Uma matriz $X \in \mathfrak{g}$ sempre possui uma base de Schur $b \in G$ (ver Seção A.5). Geometricamente, uma tal base equivale a uma sequência (V_1, \dots, V_n) de n subespaços de \mathbb{C}^n estritamente encaixados

$$V_1 \subset V_2 \subset \dots \subset V_n = \mathbb{C}^n, \quad \dim V_k = k, \quad k = 1, \dots, n, \quad (2.1)$$

tal que cada subespaço é X -invariante, isto é

$$XV_k \subset V_k, \quad k = 1, \dots, n. \quad (2.2)$$

De fato, seja $b = (v_1 | \dots | v_n) \in \mathrm{Gl}(n, \mathbb{C})$ uma base de Schur para X e considere os subespaços $V_k = [v_1, \dots, v_k]$. Eles são tais que $XV_k \subset V_k$, pois $b^{-1}Xb$ é triangular superior. Reciprocamente, sejam $V_1 \subset \dots \subset V_n = \mathbb{C}^n$ subespaços tais que $\dim V_k = k$ e $XV_k \subset V_k$. Tomando uma base $b \in \mathrm{Gl}(n, \mathbb{C})$ cujos vetores coluna v_1, \dots, v_n são tais que $V_k = [v_1, \dots, v_k]$, temos que bXb^{-1} é triangular superior, uma vez que $XV_k \subset V_k$.

Uma sequência (V_1, V_2, \dots, V_n) de subespaços de \mathbb{C}^n , satisfazendo (2.1) é chamado *flag maximal de \mathbb{C}^n* . Um flag maximal que satisfaz (2.2) é chamado de *flag X -invariante*. Segue que uma base de Schur para $X \in \mathfrak{g}$ corresponde a um flag maximal de \mathbb{C}^n que é X -invariante.

Uma matriz invertível age nos flags maximais da seguinte maneira. Considere o conjunto de todos os flags maximais de \mathbb{C}^n dado por

$$\mathbb{F} = \{x = (V_1, \dots, V_n) : x \text{ é um flag maximal de } \mathbb{C}^n\},$$

também chamado de *variedade flag maximal de \mathbb{C}^n* . Temos que o grupo $G = \mathrm{Gl}(n, \mathbb{C})$ age à esquerda de \mathbb{F} por

$$g(V_1, \dots, V_n) = (gV_1, \dots, gV_n), \quad g \in G. \quad (2.3)$$

De fato, como $g \in G$ preserva a dimensão e a inclusão de subespaços de \mathbb{C}^n , decorre das propriedades do produto em G que a ação (2.3) de G em \mathbb{F} está bem definida. Um flag $x = (V_1, \dots, V_n) \in \mathbb{F}$ é invariante por $g \in G$ se, e somente se,

$$gV_k = V_k, \quad k = 1, \dots, n, \quad (2.4)$$

ou seja, $gx = x$.

Lema 2.1. Se g é invertível, então um flag maximal g -invariante é um ponto fixo da ação de g em \mathbb{F} .

As observações acima transformam o problema de encontrar uma base de Schur para $g \in G$ no problema de encontrar um ponto fixo de g em \mathbb{F} . Isso pode ser feito de maneira iterativa com as seguintes ideias. Suponha que $x \in \mathbb{F}$ é tal que suas iteradas convergem para $x' \in \mathbb{F}$, isto é,

$$\lim_{t \rightarrow \infty} g^t x = x' \quad (2.5)$$

com $t \in \mathbb{Z}$. Então x' é um ponto fixo de g em \mathbb{F} , uma vez que

$$gx' = g \left(\lim_{t \rightarrow \infty} g^t x \right) = \lim_{t \rightarrow \infty} g^{t+1} x = x'. \quad (2.6)$$

Para que (2.5) faça sentido, \mathbb{F} deve ser munido de uma topologia. Para que (2.6) funcione, devemos ter (2.5) e também que a ação de g seja contínua e o limite de seqüências em \mathbb{F} seja único.

Na próxima seção introduzimos uma decomposição de G que será crucial para o estudo da topologia de \mathbb{F} introduzida na seção seguinte.

2.2 Decomposição de Bruhat (LPU)

Nesta seção veremos que, modificando a decomposição PLU (ver Apêndice A.3), obtemos uma decomposição $g = lpu$ com $l \in L$, $p \in P$ e $u \in U$ na qual a matriz permutação p é única. Essa decomposição é conhecida como *decomposição de Bruhat* ou *decomposição LPU* em G e tem diversas aplicações, das quais algumas exploraremos ainda neste capítulo.

A demonstração que daremos (baseada em [3], p. 396) consiste em aplicarmos sucessivas operações elementares sobre as colunas de g , obtendo u' triangular superior de modo que gu' é triangular inferior a menos de uma (única) permutação de colunas, ou seja, a menos de multiplicação à direita por uma permutação p . Pré-requisitos se encontram no Apêndice A.1.

Teorema 2.2. Dado $g \in G$, existem $l \in L$, $u \in U$ e única $p \in P$ tais que

$$g = lpu.$$

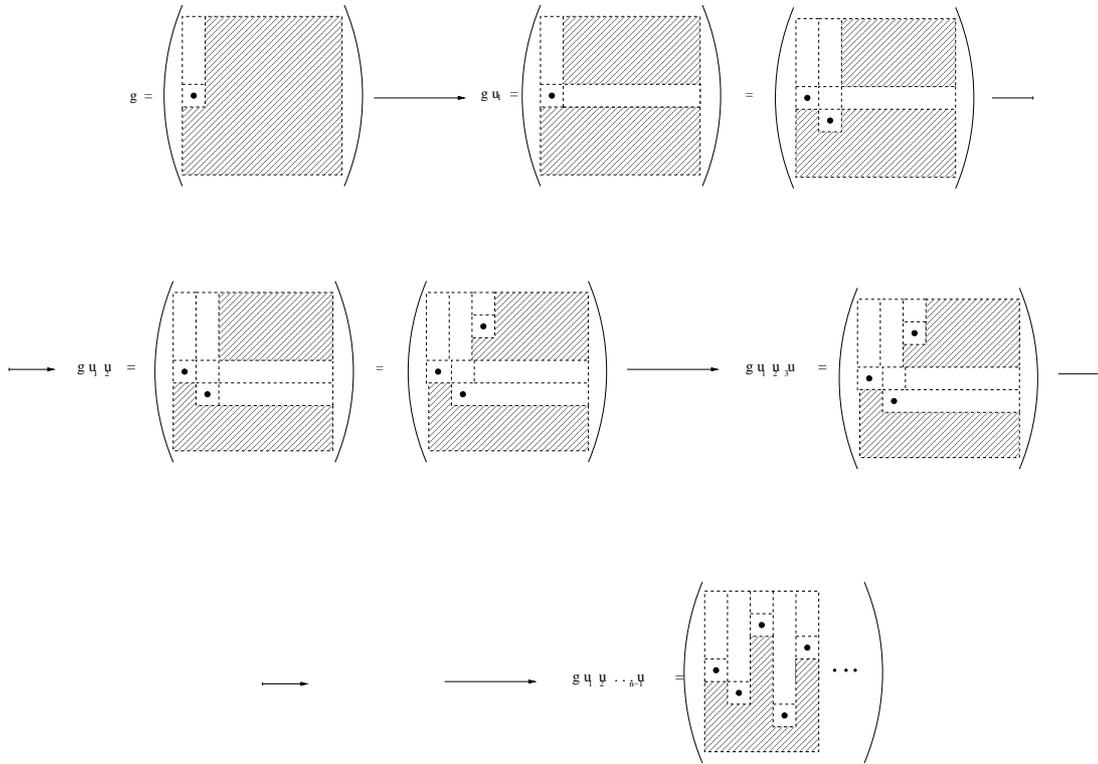


Figura 2.1: Obtendo decomposição LPU de G

Demonstração. (Existência) Acompanhe a demonstração na Figura 2.1. Como $g \in G$ tem posto n , suas colunas são linearmente independentes. Seja i_1 a linha da primeira entrada não-nula da primeira coluna de $g = (g_{ij})$. Considerando as operações elementares

$$C_k \rightarrow C_k - \frac{g_{i_1 k}}{g_{i_1 1}} C_1$$

para $1 < k \leq n$, ou seja multiplicando g à direita pela matriz triangular superior com 1 na diagonal

$$u_1 = \left(e_1 \mid e_2 - \frac{g_{i_1 2}}{g_{i_1 1}} e_1 \mid \cdots \mid e_n - \frac{g_{i_1 n}}{g_{i_1 1}} e_1 \right) \in U,$$

temos que gu_1 possui entradas nulas na linha i_1 da segunda à n -ésima entrada. Como gu_1 é invertível, existe i_2 linha da primeira entrada não-nula da segunda coluna de gu_1 . Assim,

$$i_1 \neq i_2,$$

uma vez que as entradas da linha i_1 de gu_1 são nulas, salvo a primeira. Com as operações elementares

$$C_k \rightarrow C_k - \frac{g_{i_2 k}}{g_{i_2 2}} C_2$$

para $2 < k \leq n$, obtemos $u_2 \in U$ que multiplicada à direita de gu_1 não altera sua i_1 -ésima linha, pois fixa a primeira coluna e as demais entradas são combinações de zeros que, ao substituir uma entrada nula não altera seu valor. Além disso, $gu_1 u_2$ possui entradas nulas na linha i_2 da terceira até a n -ésima entrada. Repetindo este

processo até a n -ésima coluna, obtemos valores distintos $i_1, \dots, i_n \in \{1, \dots, n\}$ e $u_1, u_2, \dots, u_{n-1} \in U$ satisfazendo as seguintes condições (ver Figura 2.2 abaixo).

- (i) Pondo $u' = u_1 u_2 \cdots u_{n-1} \in U$, gu' é tal que, na k -ésima coluna, o seu primeiro elemento não nulo está na linha i_k .
- (ii) Na linha i_k de gu' os elementos das colunas posteriores à k são nulos.
- (iii) $\sigma(k) = i_k$ define uma permutação $\sigma \in S_n$.

Afirmamos que permutando as colunas de gu' segundo σ obtemos uma matriz triangular inferior, isto é,

$$gu'I_\sigma \in L.$$

De fato, I_σ permuta as colunas de gu' segundo σ , quando multiplicada à direita,

$$gu' = (C_i) \rightarrow (C_{\sigma(i)}) = gu'I_\sigma.$$

A coluna k de gu' é tal que, sua primeira linha não-nula é $\sigma(k)$. Assim, movendo a coluna k de gu' , sua primeira linha não-nula estará na diagonal precisamente quando estivermos na coluna $\sigma(k)$ (ver figura 2.2).

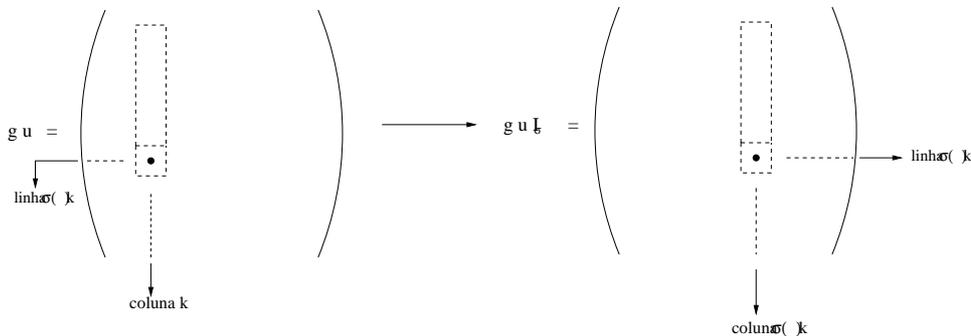


Figura 2.2: Ação de I_σ à direita de gu'

Logo, todos os elementos acima da diagonal de $gu'I_\sigma$ são nulos e seus elementos da diagonal não são nulos, de modo que

$$gu'I_\sigma = l \in L,$$

como afirmamos.

Pondo $u = (u')^{-1}$ e $p = I_{\sigma^{-1}} \in P$, temos que

$$g = lpu$$

é a decomposição desejada.

(Unicidade) Suponha que

$$g = lI_\sigma u = l'I_\sigma' u'$$

são duas decomposições de Bruhat para g e mostremos que $\sigma = \sigma'$. Como $I_\sigma = l^{-1}l'I_{\sigma'}u'u^{-1}$, é suficiente mostrarmos

$$I_\sigma = lI_\tau u \Rightarrow \sigma = \tau.$$

Se $l \in L$ e $u \in U$, pondo $l = (l_{ij})$, da Proposição A.2 temos que

$$u^{-1} = I_\sigma^{-1}lI_\tau = (l_{\sigma(i)\tau(j)}).$$

Como as entradas da diagonal de u^{-1} são não-nulas, segue que $l_{\sigma(k)\tau(k)} \neq 0$ para todo k . Como as entradas acima da diagonal de l são nulas, segue que

$$\sigma(k) \geq \tau(k), \quad \forall k.$$

Daí segue que, $\sigma(k) = \tau(k)$, $\forall k$, ou seja,

$$\sigma = \tau,$$

como queríamos. □

Observação 2.3. (a) A não unicidade dos fatores triangulares na decomposição de Bruhat para uma dada $g \in G$ pode ser vista da seguinte maneira. Para $(a_{ij}) \in \mathfrak{g}$ temos que a conjugação

$$I_{\sigma^{-1}}(a_{ij})I_\sigma = (a_{\sigma(i)\sigma(j)})$$

permuta a diagonal de (a_{ij}) segundo σ . Assim, podemos “mover” a diagonal de l para u na decomposição de Bruhat $g = lpu$ sem alterar, claro, a matriz permutação p . Mais precisamente, considere $L' = \{g \in L : g_{ii} = 1, \forall i\}$ o subgrupo das matrizes triangulares superiores com entradas na diagonal todas iguais a 1. Escrevendo $l = l'd$ com $l' \in L'$ e d diagonal (ver Lema A.4), temos que

$$lpu = l'dpu = l'p(p^{-1}dp)u = l'p(d'u),$$

onde $d' = p^{-1}dp$ é diagonal e $d'u \in U$. Uma consequência importante disso é que podemos considerar l como tendo entradas 1 em sua diagonal. Mesmo assim, ainda não temos a unicidade dos fatores triangulares na decomposição de Bruhat como mostra a relação (2.7) do Exemplo 2.5 abaixo.

(b) Dada $g \in G$, a decomposição de Bruhat (LPU) para g foi obtida de maneira natural a partir de uma modificação da Eliminação Gaussiana com pivoteamento (decomposição PLU), como ilustrado abaixo.

Decomposição PLU:

$$\begin{aligned} g &\rightarrow p_1g \rightarrow p_1gu_1 \rightarrow p_2p_1gu_1 \rightarrow p_2p_1gu_1u_2 \rightarrow \cdots \\ &\cdots \rightarrow \underbrace{p_{n-1} \cdots p_2p_1}_{\in P} g \underbrace{u_1u_2 \cdots u_{n-1}}_{\in U} \in L \end{aligned}$$

Decomposição LPU:

$$g \rightarrow gu_1 \rightarrow gu_1u_2 \rightarrow \cdots \rightarrow g \underbrace{u_1u_2 \cdots u_{n-1}}_{\in U} \in LP.$$

(c) *Obtivemos a existência da decomposição de Bruhat para $g \in G$ da seguinte maneira. Por operações elementares sobre as colunas de g , anulando suas linhas da esquerda para a direita, obtivemos $u' \in U$ tal que $gu' \in LP$. Uma outra maneira de conseguirmos tal decomposição é efetuar operações elementares sobre as linhas de g , anulando suas colunas de cima para baixo, afim de obtermos $l' \in L$ tal que $l'g \in PU$.*

Uma consequência imediata da unicidade da permutação dada pelo teorema acima é a seguinte decomposição de G .

Corolário 2.4. G se decompõe como união disjunta

$$G = \bigcup_{\sigma \in S_n} LI_\sigma U.$$

O próximo exemplo, além de enfatizar a não unicidade dos fatores triangulares na decomposição de Bruhat em G , mostra que as decomposições LPU e LPL podem ser obtidas uma da outra. Ambas decomposições são conhecidas como *decomposição de Bruhat*.

Exemplo 2.5. Considere

$$J = \begin{pmatrix} & & & 1 \\ & 0 & \cdot & \\ & & \cdot & \\ & \cdot & & 0 \\ 1 & & & \end{pmatrix} \in P.$$

Observe que $J = I_\sigma$, onde $\sigma \in S_n$ é a permutação que reverte a ordem: $\sigma(1) = n, \sigma(2) = n-1, \dots, \sigma(n) = 1$. Temos que $J^2 = I$ e, para qualquer matriz triangular inferior l , a matriz $u = JlJ$ é triangular superior, pois J age em \mathfrak{g} permutando as linhas segundo σ^{-1} ($= \sigma$) quando multiplicada à esquerda e permutando as colunas segundo σ quando multiplicada à direita (ver Proposição A.2). Segue que

$$g = lJu = l^2JI \tag{2.7}$$

são duas decomposições de Bruhat diferentes para g , claro, com a mesma permutação J .

A matriz J estabelece uma relação entre as decomposições LPU e LPL . De fato, se g é invertível, então gJ também é e, pelo Teorema 2.2, podemos escrever

$$gJ = lpu,$$

decomposição LPU para gJ , com $l \in L$, $u \in U$ e $p \in P$, e daí

$$g = l(pJ)(JuJ),$$

é uma decomposição LPL para g , com permutação pJ . Por outro lado, a partir de uma decomposição LPL

$$gJ = l_1\pi l_2$$

com $l_1, l_2 \in L$ e $\pi \in P$, podemos obter uma decomposição LPU com

$$g = l_1(\pi J)(Jl_2J),$$

com permutação πJ .

2.2.1 Decomposição LU

Nesta subseção, vamos mostrar que “a maioria” dos elementos de \mathfrak{g} admitem decomposição LU, que é a decomposição LpU com $p = I$, e vamos tirar algumas consequências disso.

Primeiro vamos caracterizar os elementos de \mathfrak{g} que admitem decomposição LU através de seus menores principais. Dada $X \in \mathfrak{g}$ e $k \in \mathbb{N}$, o k -menor principal de g , denotado por

$$X[k] \in \mathfrak{gl}(k, \mathbb{C}),$$

é formado pelas k primeiras linhas de X truncadas até a k^{a} entrada. Mais precisamente, considere a inclusão

$$\mathbb{C}^k = \{(a_1, \dots, a_k, 0, \dots, 0) : a_i \in \mathbb{C}\} \subset \mathbb{C}^n$$

e a projeção $\pi_k : \mathbb{C}^n \rightarrow \mathbb{C}^k$ nas primeiras k coordenadas. Então, $X[k]$ é dado por

$$X[k]v = \pi_k(Xv), \quad v \in \mathbb{C}^k.$$

Lema 2.6. Para todo $X \in \mathfrak{g}$, $u \in \mathfrak{u}$ e $l \in \mathfrak{l}$ valem

(a) $(Xu)[k] = X[k]u[k]$ e

(b) $(lX)[k] = l[k]X[k]$,

para cada $k \in \mathbb{N}$.

Demonstração. Se $u \in \mathfrak{u}$, então $u(\mathbb{C}^k) = \mathbb{C}^k$. O que nos dá $u[k]v = uv$ para todo $v \in \mathbb{C}^k$. Assim,

$$X[k]u[k]v = X[k]uv = \pi_k(Xuv) = (Xu)[k]v, \quad \forall v \in \mathbb{C}^k.$$

De onde decorre (a). Para (b), notemos inicialmente que

$$(X[k])^\dagger = X^\dagger[k].$$

e que $l^\dagger \in U$ se $l \in L$. Assim, pelo item (a),

$$((lX)[k])^\dagger = (X^\dagger l^\dagger)[k] = X^\dagger[k]l^\dagger[k] = (X[k])^\dagger(l[k])^\dagger = (l[k]X[k])^\dagger.$$

De onde decorre (b). □

A seguinte proposição é consequência da decomposição de Bruhat e do lema anterior.

Proposição 2.7. Seja $g \in G$. Então, existem $l \in L$ e $u \in U$ tais que $g = lu$ se, e somente se, os k -menores principais têm determinantes não-nulos para todo k .

Demonstração. Seja $g = lI_\sigma u$ decomposição de Bruhat para g . Como as entradas das diagonais de l e de u são todas não-nulas, temos que $l[k], u[k] \in \text{Gl}(k, \mathbb{C})$ para todo $k = 1, \dots, n$. Além disso, pelos itens (a) e (b) do lema anterior temos que

$$g[k] = (lI_\sigma)[k]u[k] = l[k]I_\sigma[k]u[k].$$

Segue que $g[k]$ tem determinante não-nulo para todo k se, e somente se, $I_\sigma[k]$ tem determinante não-nulo para todo k . Como $\det I[k] = 1$, para todo k , temos $g = lu$ possui $\det g[k] \neq 0$ para todo k . Reciprocamente, se $\det(I_\sigma[k]) \neq 0$ para todo k , como $\det(I_\sigma[1]) \neq 0$, temos

$$I_\sigma = (e_1 | * | \dots | *).$$

Como $\det(I_\sigma[2]) \neq 0$, temos

$$I_\sigma = (e_1 | e_2 | * | \dots | *).$$

Prosseguindo assim, temos

$$I_\sigma = I.$$

Ou seja, $g = lu$. □

Observe que LU não é subgrupo de G , o que pode ser visto a partir da proposição acima. De fato, basta tomar $u \in U$ e $l \in L$ tais que a primeira entrada da primeira coluna de ul seja nula. Então $(l^{-1}u^{-1})^{-1} \notin LU$. Por exemplo, tomando $u = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ e $l = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, temos que $l^{-1}u^{-1} \in LU$, no entanto

$$(l^{-1}u^{-1})^{-1} = ul = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \notin LU,$$

uma vez que seu 1-menor principal $(ul)[1]$ é nulo.

Da Proposição 2.7, segue que

$$LU = \{g \in \mathfrak{g} : \det(g[k]) \neq 0 \text{ para } k = 1, \dots, n\}.$$

Além disso, para cada $k = 1, \dots, n$, o determinante $\det(g[k])$ é um polinômio não-nulo nas entradas de g . Então, LU é uma interseção finita de abertos densos em \mathfrak{g} , dados pelos complementares de zeros de polinômios. Assim, obtemos o próximo resultado.

Proposição 2.8. LU é aberto e denso em \mathfrak{g} , logo em G .

Como observamos anteriormente (Observação 2.3), na decomposição LU podemos tomar o fator em L com 1 em todas as entradas da diagonal, movendo a diagonal do fator L para o fator U . Além disso, é fácil ver que L' é subgrupo de L e que $L' \cap U = \{I\}$.

Proposição 2.9. Seja $g \in LU$. Então g pode ser escrita de maneira única como $g = lu$ com $l \in L'$ e $u \in U$.

Demonstração. Pelas observações acima, temos que $g \in LU$ pode ser escrita como $g = lu$, com $l \in L'$ e $u \in U$. Suponha que $g = lu = l'u'$, com $l' \in L'$ e $u' \in U$. Então,

$$u(u')^{-1} = l^{-1}l' \in U \cap L' = \{I\},$$

e portanto, $u = u'$ e $l = l'$. □

O próximo resultado topológico é bastante importante para a próxima seção.

Teorema 2.10. A restrição da multiplicação

$$L' \times U \longrightarrow LU, \quad (l, u) \longmapsto lu$$

é sobrejetiva e aberta, onde $L' \times U$ tem a topologia produto.

Demonstração. Seja $\psi : L' \times U \rightarrow LU$, $(l, u) \mapsto lu$. Da Proposição 2.9, temos que $L'U = LU$, de onde segue a sobrejetividade de ψ . Para mostrar que ψ é aberta, basta mostrar que a imagem de uma vizinhança de (I, I) em $L' \times U$ é uma vizinhança de I em LU . De fato, se A é uma vizinhança de (l, u) em $L' \times U$, então existem V vizinhança de I em L' e W vizinhança de I em U tais que $lV \times Wu$ é uma vizinhança de (l, u) contida em A (basta usar a topologia produto e a Proposição B.8). Como $V \times W$ é vizinhança de (I, I) em $L' \times U$, por hipótese temos que sua imagem $\psi(V, W) = VW$ é uma vizinhança de $\psi(I, I) = I$ em LU . A imagem de $lV \times Wu$ é $\psi(lV, Wu) = lVWu$, que é uma vizinhança de lu em LU , uma vez que a aplicação $g \in LU \mapsto lgu \in LU$ é um homeomorfismo de LU . Segue que a imagem da vizinhança A de (l, u) contém uma vizinhança de $\psi(l, u) = lu$, mostrando que ψ é aberta.

Para mostrar que ψ leva vizinhança de (I, I) em vizinhança de I , vamos usar o teorema da função inversa da seguinte maneira. Observe que o conjunto

$$U = \{X \in \mathfrak{u} : \det(X) \neq 0\}$$

é um aberto do espaço vetorial \mathfrak{u} e que o conjunto

$$L' = \{X \in \mathfrak{l} : \text{diag}(X) = I\}$$

é um subespaço afim do espaço vetorial \mathfrak{l} , dado por

$$L' = \mathfrak{l}' + I,$$

onde

$$\mathfrak{l}' = \{X \in \mathfrak{l} : \text{diag}(X) = 0\}.$$

Observe também que

$$LU \subset \mathfrak{g} = \mathfrak{l}' \oplus \mathfrak{u},$$

uma vez que toda matriz de \mathfrak{g} se escreve de modo único como a soma de uma triangular inferior sem diagonal e uma triangular superior. Podemos então usar a estrutura de espaço afim L' e de espaço vetorial dos abertos $U \subset \mathfrak{u}$ e $LU \subset \mathfrak{g}$ para derivar ψ . Claramente, ψ é derivável e sua derivada em (I, I) é uma aplicação linear

$$\psi'(I, I) : \mathfrak{l}' \times \mathfrak{u} \longrightarrow \mathfrak{g}.$$

Para calculá-la, uma curva em $L' \times U$ que passa por (I, I) na direção de (X, Y) , com $X \in \mathfrak{l}'$ e $Y \in \mathfrak{u}$ é dada por

$$\alpha(t) = (I + tX, I + tY), t \in \mathbb{R}.$$

Derivando

$$\psi(\alpha(t)) = (I + tX)(I + tY) = I + tX + tY + t^2XY$$

em $t = 0$, obtemos

$$\psi'(I, I)(X, Y) = X + Y, \quad X \in \mathfrak{l}', \quad Y \in \mathfrak{u}.$$

Como $\mathfrak{g} = \mathfrak{l}' \oplus \mathfrak{u}$, segue que $\psi'(I, I)$ é um isomorfismo linear. Então, pelo teorema da função inversa temos que ψ é um difeomorfismo local numa vizinhança de (I, I) . Em particular, a imagem de uma vizinhança suficientemente pequena de (I, I) em $L' \times U$ é uma vizinhança de I em LU , como queríamos. \square

2.3 Topologia da variedade flag maximal

Nesta seção usaremos ferramentas do Apêndice B.3 a fim de introduzirmos em \mathbb{F} uma topologia adequada aos nossos interesses. Primeiro veremos como a decomposição de Bruhat de G equivale a uma decomposição correspondente de \mathbb{F} . Denote por

$$e = \left([e_1], [e_1, e_2], \dots, [e_1, \dots, e_n] \right)$$

o *flag canônico* em \mathbb{F} , onde e_1, \dots, e_n são os vetores canônicos de \mathbb{C}^n . Se $\sigma \in S_n$ é uma permutação, denote por

$$\sigma e = I_\sigma e = \left([e_{\sigma(1)}], [e_{\sigma(1)}, e_{\sigma(2)}], \dots, [e_{\sigma(1)}, \dots, e_{\sigma(n)}] \right)$$

o *flag canônico permutado* por σ . Finalmente considere a projeção

$$\pi : G \rightarrow \mathbb{F}, \quad g \mapsto ge,$$

que é G -equivariante, pois

$$g'(ge) = (g'g)e, \quad \forall g', g \in G,$$

e sobrejetora, de acordo com a proposição seguinte.

Proposição 2.11. A ação de G em \mathbb{F} é transitiva e a isotropia do flag canônico e é o subgrupo U . Em particular, $\pi^{-1}(e) = U$.

Demonstração. Devemos mostrar que \mathbb{F} é uma G -órbita. Afirmamos que

$$\mathbb{F} = Ge = \{ge : g \in G\}.$$

De fato, se $x = (V_1, \dots, V_n) \in \mathbb{F}$, escrevendo $V_k = [v_1, \dots, v_k]$, para $k = 1, \dots, n$, temos que $\{v_1, \dots, v_n\}$ é uma base de \mathbb{C}^n e g definido por

$$ge_k = v_k, \quad k = 1, \dots, n$$

é tal que $g \in G$ e $ge = x$. Então,

$$\mathbb{F} \subset Ge \subset \mathbb{F},$$

de onde segue nossa afirmação.

Além disso, a isotropia de e em G é

$$G_e = \{g \in G : ge = e\} = U.$$

De fato, $g \in G$ é triangular superior se, e somente se,

$$ge_k \in [e_1, \dots, e_k], \quad k = 1, \dots, n.$$

Como g é invertível, isso ocorre se, e somente se,

$$g[e_1, \dots, e_k] = [e_1, \dots, e_k], \quad k = 1, \dots, n,$$

ou seja, $ge = e$.

A última observação segue de

$$\pi^{-1}(e) = \{g \in G : ge = e\} = G_e.$$

□

A decomposição de Bruhat de G (ver Corolário 2.4)

$$G = \bigcup_{\sigma \in S_n} LI_\sigma U,$$

juntamente com a proposição anterior nos dá

$$\mathbb{F} = Ge = \bigcup_{\sigma \in S_n} LI_\sigma Ue = \bigcup_{\sigma \in S_n} L\sigma e, \quad (2.8)$$

uma vez que U é a isotropia de e . Essa é a *decomposição de Bruhat de \mathbb{F}* . Cada órbita $L\sigma e$ de L na decomposição de Bruhat de \mathbb{F} é chamada de *célula de Bruhat* e, para $\sigma = 1$, temos a *célula principal de Bruhat* $C = Le$.

Proposição 2.12. Cada célula de Bruhat é tal que $L\sigma e = L'\sigma e$. Além disso, elas são disjuntas, isto é,

$$\mathbb{F} = \bigcup_{\sigma \in S_n} L\sigma e$$

é uma união disjunta de órbitas de L . Temos ainda que as decomposições de Bruhat de \mathbb{F} e de G podem ser obtidas uma da outra.

Demonstração. Segue do item (iii) do lema A.4 que $LI_\sigma U = L'I_\sigma U$ para todo $\sigma \in S_n$. Isso, juntamente com a proposição anterior nos dá $L\sigma e = L'\sigma e$ para todo $\sigma \in S_n$.

Para mostrar que a união é disjunta, sejam

$$l\sigma e = l'\tau e \in \mathbb{F},$$

com $l, l' \in L$ e $\sigma, \tau \in S_n$. Então

$$I_\sigma e = l^{-1}l'I_\tau e = l''I_\tau e \implies I_\sigma^{-1}l''I_\tau e = e,$$

com $l'' = l^{-1}l' \in L$. Portanto,

$$I_\sigma^{-1}l''I_\tau \in G_e = U \implies I_\sigma^{-1}l''I_\tau = u \in U \implies l''I_\tau = I_\sigma u \implies \tau = \sigma$$

pela unicidade da permutação na decomposição de Bruhat de G (ver Teorema 2.2).

Já mostramos que a decomposição de Bruhat de \mathbb{F} é obtida da decomposição de Bruhat de G (equação (2.8) acima). Para a recíproca, observe que

$$\pi^{-1}(L\sigma e) = \pi^{-1}(LI_\sigma e) = LI_\sigma \pi^{-1}(e) = LI_\sigma U,$$

onde usamos que $\pi^{-1}(e) = U$ e que

$$\begin{aligned} g \in \pi^{-1}(LI_\sigma e) &\Leftrightarrow \pi(g) = lI_\sigma e, \quad \text{com } l \in L, \\ &\Leftrightarrow ge = lI_\sigma e \\ &\Leftrightarrow I_\sigma^{-1}l^{-1}g \in U = \pi^{-1}(e) \\ &\Leftrightarrow g \in LI_\sigma \pi^{-1}(e). \end{aligned}$$

Donde, $\pi^{-1}(LI_\sigma e) = LI_\sigma \pi^{-1}(e)$.

Sabendo que as órbitas $L\sigma e$ são disjuntas e exaurem \mathbb{F} , segue que os conjuntos $LI_\sigma U$ são disjuntos e exaurem G , para σ variando em S_n . \square

Como a ação de G em \mathbb{F} é transitiva, podemos munir \mathbb{F} com uma única topologia tal que essa ação é contínua e aberta (ver Teorema B.16). Em particular a projeção $\pi : G \rightarrow \mathbb{F}$, $g \mapsto ge$ é contínua e aberta. O próximo resultado fornece propriedades dessa topologia. Uma *seção* de π sobre $A \subset \mathbb{F}$ é uma aplicação $s : A \rightarrow G$ tal que $\pi(s(x)) = x$ para todo $x \in A$. Temos o seguinte diagrama comutativo

$$\begin{array}{ccc} & & G \\ & \nearrow s & \downarrow \pi \\ A & \xrightarrow{i} & \mathbb{F} \end{array}$$

onde $i : A \rightarrow \mathbb{F}$, $i(x) = x$ é a inclusão.

Teorema 2.13. Temos que \mathbb{F} é Hausdorff e que a célula principal de Bruhat, C , é aberta e densa em \mathbb{F} . Além disso, a projeção π possui uma seção contínua sobre C .

Demonstração. Para mostrar que \mathbb{F} é Hausdorff, basta mostrar que a isotropia U do flag canônico é um subgrupo fechado de G (ver Proposição B.12). Para isso, dado $X = (X_{ij}) \in \mathfrak{g}$, considere o subespaço das matrizes triangulares superiores

$$\mathfrak{u} = \{X \in \mathfrak{g} : X_{ij} = 0, i > j\},$$

que é claramente fechado em \mathfrak{g} . Uma vez que $U = G \cap \mathfrak{u}$, segue que U é fechado em G na topologia induzida por \mathfrak{g} .

Como π é contínua, aberta e sobrejetiva, temos que π leva aberto e denso de G em aberto e denso de \mathbb{F} (ver Proposição B.3). Logo,

$$\pi(LU) = Le = C$$

é aberto e denso em \mathbb{F} , uma vez que LU é aberto e denso em G (ver Proposição 2.8).

Para construir a seção em C , vamos usar que $C = L'e$ (ver Proposição 2.12). Denote por $\pi' : L' \rightarrow C$ a restrição de π . Claramente, π' é contínua e sobrejetora. Afirmamos que π' é aberta. De fato, a aplicação $L' \times U \rightarrow G$ dada pelo produto de matrizes é aberta (ver Teorema 2.10). Se V é um aberto de L' , então $V \times U$ é aberto de $L' \times U$, pela definição de topologia produto. Segue que VU é aberto em G , de modo que

$$\pi'(V) = Ve = VUe = \pi(VU)$$

é aberto em \mathbb{F} , uma vez que a ação de G em \mathbb{F} é aberta. Considere a aplicação

$$s : C \rightarrow G, \quad le \mapsto l, \quad l \in L'.$$

Ela está bem definida uma vez que $l_1e = l_2e$, com $l_1, l_2 \in L'$, implica que $l_2^{-1}l_1e = e$, logo que $l_2^{-1}l_1 \in U \cap L' = \{I\}$, de modo que $l_1 = l_2$. A aplicação s é claramente uma seção de π . Para sua continuidade, observe que seu levantamento

$$s \circ \pi' : L' \rightarrow G$$

coincide com a inclusão de L' em G , que é contínua uma vez que em L' se considera a topologia induzida. Como $\pi' : L' \rightarrow C$ é aberta, segue que $s : C \rightarrow G$ é contínua (ver Proposição B.5), como queríamos. \square

Note que a demonstração do teorema acima mostra que C é homeomorfo à L' .

Observação 2.14. Já vimos que cada flag $x \in \mathbb{F}$ possui uma base adaptada $b \in \text{Gl}(n, \mathbb{F})$ tal que $\pi(b) = x$. O significado geométrico da existência de seção contínua para π sobre um conjunto $A \subset \mathbb{F}$ pode ser entendido da seguinte maneira. Se $x(t) \in A$ é uma curva contínua de flags com $t \in (-\delta, \delta)$ então, para t suficientemente pequeno, existe uma curva contínua $b(t) \in \text{Gl}(n, \mathbb{C})$ tal que cada $b(t)$ é base adaptada para $x(t)$ (ver Figura 2.3).

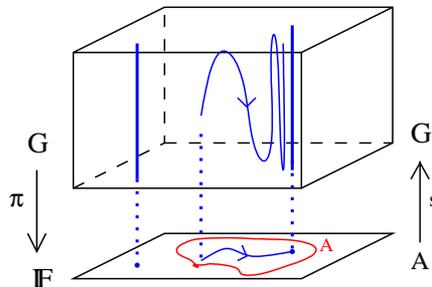


Figura 2.3: Seção contínua de π sobre A

Com a topologia de espaço homogêneo de $\text{Gl}(n, \mathbb{C})$, temos que a ação de g em \mathbb{F} é contínua e que o limite de sequências em \mathbb{F} é único (pois \mathbb{F} é Hausdorff), como queríamos.

Exemplo 2.15. Tomando $n = 2$ temos que a variedade flag maximal de \mathbb{C}^2 é

$$\mathbb{F}(\mathbb{C}^2) = \{(V, \mathbb{C}^2) : V \text{ subespaço de } \mathbb{C}^2 \text{ com } \dim V = 1\}.$$

Considerando o homeomorfismo $V \mapsto (V, \mathbb{C}^2)$ e $\mathbb{F}(\mathbb{C}^2)$, escrevemos

$$\mathbb{F}(\mathbb{C}^2) = \mathbb{P}(\mathbb{C}^2) = \left\{ \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right] : \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2 - \{0\} \right\},$$

ou seja, o conjunto $\mathbb{P}(\mathbb{C}^2)$ dos subespaços de \mathbb{C}^2 de dimensão 1, conhecido como linha projetiva complexa.

1. Temos que $G = \text{Gl}(2, \mathbb{C})$ age em $\mathbb{P}(\mathbb{C}^2)$ por

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right] = \left[\begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix} \right].$$

2. Afirmamos que $\mathbb{F}(\mathbb{C}^2)$ é homeomorfo à esfera de Riemann compacta S^2 .

Com efeito, como $\mathbb{F}(\mathbb{C}^2)$ é homeomorfo à $\mathbb{P}(\mathbb{C}^2)$ e $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ é homeomorfo à S^2 , é suficiente que $\mathbb{P}(\mathbb{C}^2)$ e $\overline{\mathbb{C}}$ sejam homeomorfos. De fato, a aplicação

$$f : \mathbb{P}(\mathbb{C}^2) \longrightarrow \overline{\mathbb{C}}, \quad \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right] \longmapsto \frac{z_1}{z_2},$$

onde $\frac{z_1}{z_2} = \infty$ se $z_2 = 0$ e $\frac{z_1}{z_2} = 0$ se $z_1 = 0$, está bem definida:

$$\frac{z_1}{z_2} = \frac{w_1}{w_2} \iff \left[\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right] = \left[\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right].$$

Deixamos como exercício para o leitor mostrar que f é um homeomorfismo. Note que basta mostrar bijeção e continuidade, visto que $\overline{\mathbb{C}}$ é compacto.

Além disso, temos que

$$f(e_1) = f\left(\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right]\right) = \infty \quad \text{e} \quad f(e_2) = f\left(\left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right]\right) = 0.$$

Temos ainda exatamente duas permutações, I e I_σ , em G que satisfazem

$$I[e_1] = [e_1] \quad \text{e} \quad I_\sigma[e_1] = [e_2]$$

3. O homeomorfismo f define a ação de G em $\overline{\mathbb{C}}$ por $gf(X) = f(gX)$, de modo que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{z_1}{z_2} = \frac{az_1 + bz_2}{cz_1 + dz_2}$$

e f é G -equivariante.

A projeção canônica $\pi : G \rightarrow \overline{\mathbb{C}}$ é dada por

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \infty = \frac{a}{c}.$$

4. Para a decomposição de Bruhat de $\mathbb{F}(\mathbb{C}^2)$, temos

$$L = \left\{ \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} : a, d \neq 0 \right\}$$

e as células de Bruhat são

$$C = L_\infty = \left\{ \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \right\} \frac{1}{0} = \left\{ \frac{a}{c} : a \neq 0 \right\} = \overline{\mathbb{C}} - \{0\} \text{ que é a célula principal de Bruhat}$$

$$\text{e } L_0 = \left\{ \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \right\} \frac{0}{1} = \left\{ \frac{0}{d} : d \neq 0 \right\} = \{0\}$$

5. De $L' = \left\{ \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \right\}$ temos que a seção contínua de π sobre C é dada por

$$C = L'_\infty = L' \frac{1}{0} = \left\{ \frac{1}{c} : c \in \mathbb{C} \right\} \mapsto \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \in L'.$$

2.4 Iterações em \mathbb{F}

O seguinte teorema estabelece uma maneira de encontrarmos pontos fixos de $g \in G$ em \mathbb{F} , sob certas condições sobre g . Os principais resultados que usamos aqui são a topologia e a decomposição de Bruhat de \mathbb{F} .

Teorema 2.16. Se os módulos dos autovalores de g são todos distintos, então para todo $x \in \mathbb{F}$ temos que $g^t x$ converge em \mathbb{F} para um ponto fixo x' de g , quando $t \in \mathbb{Z}$ tende ao infinito, isto é

$$g^t x \rightarrow x', \quad gx' = x'.$$

Demonstração. Como g tem autovalores distintos, temos que g é conjugada a uma matriz diagonal $d = \text{diag}(\lambda_1, \dots, \lambda_n)$, onde $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ são os autovalores de g e podemos ordená-los de modo que

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0.$$

Ou seja, existe $k \in G$ tal que

$$g = kdk^{-1}.$$

Basta mostrar que $d^t y$ converge em \mathbb{F} para todo $y \in \mathbb{F}$. De fato, se $d^t y$, $t \in \mathbb{Z}$, converge em \mathbb{F} quando $t \rightarrow \infty$ para todo $y \in \mathbb{F}$, temos que

$$g^t x = k(d^t k^{-1} x) = k(d^t y)$$

converge em \mathbb{F} para todo $x \in \mathbb{F}$, uma vez que a ação de k em \mathbb{F} é contínua. Podemos escrever

$$y = l\sigma e \in \mathbb{F},$$

com $l \in L'$ (ver Proposição 2.12).

Note que

$$d^t \sigma e = \sigma e,$$

para todo $\sigma \in S_n$, uma vez que uma vez que $de_{\sigma(i)} = \lambda_{\sigma(i)} e_{\sigma(i)}$ para $i = 1, \dots, n$. Segue que

$$d^t y = d^t(l\sigma e) = d^t(ld^{-t}d^t\sigma e) = (d^tld^{-t})\sigma e. \quad (2.9)$$

Como $|\lambda_i| > |\lambda_j|$ para $i < j$, temos da Observação A.5 que

$$d^tld^{-t} = \begin{pmatrix} 1 & & & & \\ \left(\frac{\lambda_2}{\lambda_1}\right)^t * & 1 & & & \\ \left(\frac{\lambda_3}{\lambda_1}\right)^t * & \left(\frac{\lambda_3}{\lambda_2}\right)^t * & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \left(\frac{\lambda_n}{\lambda_1}\right)^t * & \left(\frac{\lambda_n}{\lambda_2}\right)^t * & \dots & & 1 \end{pmatrix} \longrightarrow I$$

quando $t \rightarrow \infty$, onde os asteriscos acima são as respectivas entradas de l . Segue então de (2.9) que

$$d^t y \longrightarrow \sigma e, \quad \text{quando } t \longrightarrow \infty, \quad t \in \mathbb{Z},$$

para todo $y \in L\sigma e$.

O que fizemos acima mostra que $g^t x$ sempre tem limite x' em \mathbb{F} . Para mostrar que x' é ponto fixo de g , temos da ação de g em \mathbb{F} ser contínua e de \mathbb{F} ser Hausdorff, que

$$gx' = g\left(\lim_{t \rightarrow \infty} g^t x\right) = \lim_{t \rightarrow \infty} g^{t+1} x = x',$$

como queríamos. □

Exemplo 2.17. Voltando ao Exemplo 2.15, considere $g = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$, com $|\lambda_1| > |\lambda_2| > 0$, temos que para $\frac{a}{c} \in C$,

$$g^t \frac{a}{c} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}^t \frac{a}{c} \longrightarrow \infty,$$

quando $t \rightarrow \infty$, pois $a \neq 0$.

Na outra órbita, a ação de g é constante:

$$g^t \frac{0}{1} = \frac{\lambda_1^t 0}{\lambda_2^t} = 0 \quad \forall t.$$

Observação 2.18. Se no teorema anterior tivéssemos $X \in \mathfrak{g}$ e $g^t = \exp(tX)$ para $t \in \mathbb{R}$, a mesma demonstração se aplica para mostrar que, para qualquer $x \in \mathbb{F}$, temos

$$g^t x \rightarrow x', \quad t \rightarrow \infty \text{ em } \mathbb{R},$$

onde x' é ponto fixo de g^t para todo $t \in \mathbb{R}$.

2.5 Iterações em Q

Nesta seção veremos que as iterações de g em Q levam grosso modo a matriz identidade numa base ortogonal de Schur para g .

A unicidade da decomposição $G = QR$ nos fornece a projeção de G em Q , que denotamos por $Q : G \rightarrow Q$, que é contínua, Q -equivariante à esquerda e R -invariante à direita. Em \mathbb{F} temos que

$$ge = Q(g)e, \quad (2.10)$$

uma vez que R está contida na isotropia U de e . Considere a aplicação

$$G \times Q \longrightarrow Q, \quad (g, q) \longmapsto Q(gq) \quad (2.11)$$

e

$$\pi_Q : Q \longrightarrow \mathbb{F}, \quad q \longmapsto qe,$$

que é a restrição da projeção canônica $\pi : G \rightarrow \mathbb{F}$, $g \mapsto ge$.

Proposição 2.19. (i) A projeção π_Q é contínua, sobrejetora e G -equivariante.

(ii) A aplicação (2.11) é uma ação contínua de G em Q .

Demonstração. Para o item (i), temos que a continuidade de π_Q segue da continuidade de π e a sobrejetividade segue da equação (2.10). Por outro lado, a G -equivariância de π_Q também segue de (2.10) da seguinte maneira. Dados $g \in G$ e $q \in Q$, temos que

$$\pi_Q(gq) = \pi_Q(Q(gq)) = Q(gq)e = gqe = g\pi_Q(q),$$

onde usamos que $R \subset U$ e que $U = G_e$ é a isotropia do flag canônico e em G para a terceira igualdade. Logo,

$$g\pi_Q(q) = \pi_Q(gq), \quad \forall g \in G, q \in Q,$$

como queríamos.

Como a aplicação $Q : G \rightarrow Q$ é R -invariante à direita, temos que

$$Q(g'Q(gq)) = Q(g'gq),$$

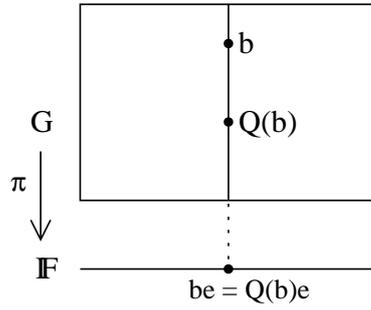
para todo $g, g' \in G$ e $q \in Q$. Com isso, mostramos o item (ii). \square

A figura abaixo ilustra a sobrejetividade de π_Q .

Corolário 2.20. A variedade flag maximal, \mathbb{F} , é compacta.

Demonstração. Como $\pi_Q : Q \rightarrow \mathbb{F}$ é sobrejetora e contínua, é suficiente mostrar que Q é compacto.

Ora, como as aplicações $q \mapsto q^*$ e $q \mapsto q$ são contínuas, segue que $f : q \mapsto q^*q$ é contínua de \mathfrak{g} em \mathfrak{g} . Assim, como $\{I\}$ é fechado em \mathfrak{g} , segue que $Q = f^{-1}(\{I\})$ é fechado em \mathfrak{g} . Por outro lado, tomando em \mathfrak{g} a norma do máximo sobre as colunas de seus elementos, vemos que Q é limitado. Portanto, Q é compacto. \square



Observação 2.21. A compacidade de \mathbb{F} foi ilustrada no exemplo 2.15, quando mostramos que $\mathbb{F}(\mathbb{C}^2) \approx S^2$ que é compacto.

Denotamos por gx e gq a ação de $g \in G$ em $x \in \mathbb{F}$ e em $q \in Q$, respectivamente. Ficará claro do contexto a qual ação estamos nos referindo. A G -equivariância de π_Q nos diz que a ação de g em Q cobre a ação de g em \mathbb{F} , de acordo com o seguinte diagrama comutativo.

$$\begin{array}{ccc} Q & \xrightarrow{g} & Q \\ \pi_Q \downarrow & & \downarrow \pi_Q \\ \mathbb{F} & \xrightarrow{g} & \mathbb{F} \end{array}$$

Para relacionar as duas ações, considere o subgrupo das matrizes triangulares superiores unitárias

$$T = Q \cap U.$$

Se $g \in T$, temos que

$$g^{-1} = g^* \in U \cap L \cap Q,$$

e portanto g é diagonal unitária, logo diagonal com entradas de módulo 1. Ou seja,

$$T = \{\text{diag}(z_1, \dots, z_n) : |z_i| = 1, \forall i\}.$$

Dizemos que a fibra $\pi_Q^{-1}(x)$ sobre $x \in \mathbb{F}$ é g -invariante se $\pi_Q(gq) = x$ para todo $q \in \pi_Q^{-1}(x)$.

Proposição 2.22. A projeção π_Q tem fibra T sobre o ponto base e . Além disso, π_Q possui seção contínua sobre $C = Le$.

Demonstração. Como a fibra de $\pi : G \rightarrow \mathbb{F}$, $g \mapsto ge$, no ponto base $e \in \mathbb{F}$ é $\pi^{-1}(e) = U$ (ver Proposição 2.11) e π_Q é a restrição de π à Q , segue que a fibra de π_Q no ponto base $e \in \mathbb{F}$ é

$$\pi_Q^{-1}(e) = Q \cap \pi^{-1}(e) = Q \cap U = T.$$

Para a segunda parte, considere s uma seção contínua para $\pi : G \rightarrow \mathbb{F}$, $g \mapsto ge$, em C , que existe (ver Teorema 2.13), e $Q : G \rightarrow G$ a projeção em Q da decomposição QR de G . Como $Q : G \rightarrow G$ é contínua (ver Corolário A.10), temos que $s' = Q \circ s : C \rightarrow Q$ é contínua e tal que

$$\pi_Q(s'(x)) = \pi(s'(x)) = Q(s(x))e = s(x)e = \pi(s(x)) = x.$$

Segue que s' é uma seção de π_Q sobre C . □

Observação 2.23. Analogamente à Observação 2.14, o significado geométrico da existência de uma seção contínua para π_Q pode ser entendido da seguinte maneira. Se $x(t)$ é uma curva contínua de flags com $t \in (-\delta, \delta)$ então, para t suficientemente pequeno, existe uma curva contínua $b(t) \in Q$ tal que cada $b(t)$ é base ortogonal adaptada para $x(t)$.

A demonstração do resultado abaixo relaciona, por meio da fibra T , a convergência em \mathbb{F} com a convergência em Q .

Corolário 2.24. Sejam q_n, q em Q tais que $q_n e \rightarrow q e$ em \mathbb{F} . Então, existe $z_n \in T$ tal que $q_n z_n \rightarrow q$ em Q .

Demonstração. Pela continuidade da ação de q^{-1} em \mathbb{F} , temos que

$$q^{-1} q_n e \rightarrow e.$$

Como C é aberto em \mathbb{F} e contém e , temos que

$$q^{-1} q_n e \in C \quad \text{para } n \text{ grande o suficiente.}$$

Pela proposição anterior, tomando a seção contínua $s : C \rightarrow Q$, temos que

$$s(q^{-1} q_n e) \rightarrow s(e).$$

Por outro lado, como

$$s(q^{-1} q_n e) \in \pi_Q^{-1}(q^{-1} q_n e) = q^{-1} q_n \pi_Q^{-1}(e) = q^{-1} q_n T \quad \text{e} \quad s(e) = I,$$

temos que existem $w_n \in T$ tais que

$$q^{-1} q_n w_n \rightarrow I.$$

Logo,

$$q_n z_n \rightarrow q \in Q,$$

pela continuidade da ação de Q em Q . □

Teorema 2.25. Sejam $g \in G$ com autovalores de módulos distintos e $q \in Q$. Então, $g^t q$ converge para uma fibra g -invariante de π_Q . Além disso, se g é hermitiana positiva, então existe $q' \in Q$ satisfazendo $g^t q \rightarrow q'$ quando $t \rightarrow \infty$, $t \in \mathbb{Z}$. (Ver Figura 2.4.)

Demonstração. Como os módulos dos autovalores de g são distintos, em \mathbb{F} temos que

$$\pi_Q(g^t q) = g^t q e \rightarrow q' e \quad \text{quando } t \rightarrow \infty,$$

onde $q' \in Q$ é tal que $q' e$ é um flag g -invariante em \mathbb{F} (ver Teorema 2.16). Daí,

$$\pi_Q^{-1}(q' e) = q' \pi_Q^{-1}(e) = q' T$$

é fibra g -invariante em Q . De fato,

$$g q' T = \pi_Q^{-1}(g q' e) = \pi_Q^{-1}(q' e) = q' T.$$

Portanto, $g^t q \rightarrow q'T$ fibra g -invariante.

Se além disso, g é hermitiana positiva, podemos adaptar a demonstração do Teorema 2.16 da seguinte maneira. Temos que

$$g = kdk^{-1},$$

com $k \in Q$ e $d = \text{diag}(\lambda_1, \dots, \lambda_n)$ diagonal real, com $\lambda_1 > \dots > \lambda_n > 0$. Como Q é Q -equivariante à esquerda, temos que

$$g^t q = Q(g^t q) = kQ(d^t k^{-1} q).$$

Considere agora a decomposição de Bruhat $k^{-1} q = lpu$, onde $l \in L$, $p = I_\sigma$ e $u \in U$. Fazendo $u = sr$, com s é diagonal com entradas de módulo 1 e r é triangular superior com diagonal positiva, obtemos

$$k^{-1} q = lpsr.$$

Como Q é R -invariante à direita temos que

$$g^t q = kQ(d^t k^{-1} q) = kQ(d^t lps).$$

Além disso,

$$(ps)^{-1} d (ps) = s^{-1} (p^{-1} dp) s = p^{-1} dp$$

é diagonal positiva, uma vez que $p^{-1} dp$ é diagonal (positiva e permutada) que comuta com s , que é diagonal. Logo,

$$(ps)^{-1} d^{-t} (ps) \in R$$

e, novamente pela R -invariancia à direita de Q , segue que

$$g^t q = kQ(d^t lps) = kQ(d^t l d^{-t} ps).$$

Como na demonstração do Teorema 2.16, temos $d^t l d^{-t} \rightarrow I$ e, pela continuidade de Q ,

$$g^t q \rightarrow kQ(ps) = q' \in Q,$$

uma vez que $k \in Q$. □

Observação 2.26. *Os resultados anteriores valem, com as mesmas demonstrações, para $X \in \mathfrak{g}$ e $g^t = \exp(tX)$ com $t \in \mathbb{R}$, fazendo $t \rightarrow \infty$ em \mathbb{R} .*

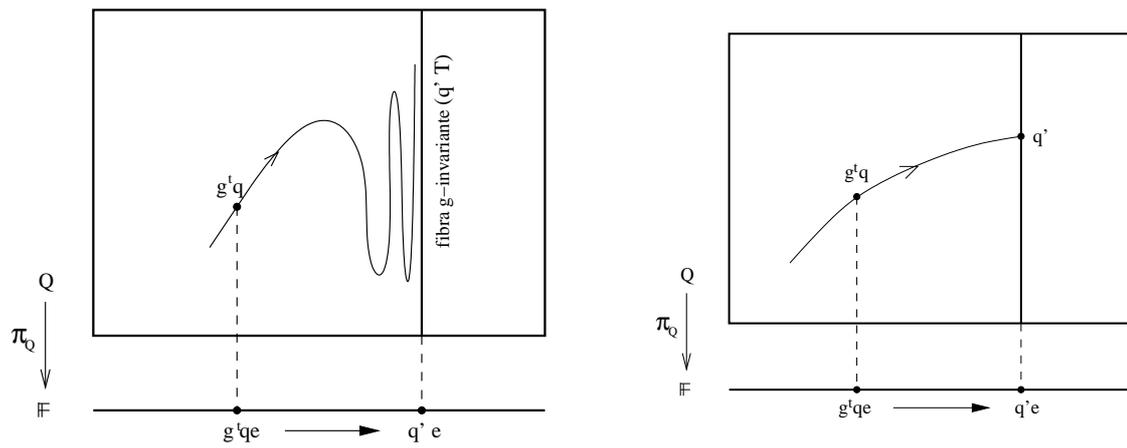


Figura 2.4: Iterações de g em Q . Na direita, g é hermitiana positiva.

Capítulo 3

Forma de Schur via método QR

Neste capítulo, trataremos do Método QR apresentado na Introdução bem como do seu análogo em tempo contínuo. Para economizar notação, usaremos $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$ e $G = \text{Gl}(n, \mathbb{C})$.

3.1 Método QR

Partindo de $g \in G$, considere a seguinte sequência de matrizes, com $q_t \in Q$ e $r_t \in R$ para todo $t \in \mathbb{N}$.

$$\begin{aligned} g_1 &= g = q_1 r_1 \\ g_2 &= r_1 q_1 = q_2 r_2 \\ &\vdots \\ g_t &= r_{t-1} q_{t-1} = q_t r_t \\ g_{t+1} &= r_t q_t = q_{t+1} r_{t+1}. \end{aligned} \tag{3.1}$$

Considere a sequência de matrizes unitárias dada por

$$k_t = q_1 \cdots q_t.$$

Proposição 3.1. Temos que a sequência do método QR é tal que

$$g_{t+1} = k_t^{-1} g k_t, \tag{3.2}$$

onde

$$k_t = Q(g^t).$$

Demonstração. Faremos a prova por indução sobre t . De (3.1) temos que

$$g_2 = r_1 q_1 = q_1^{-1} (q_1 r_1) q_1 = q_1^{-1} g q_1 = k_1^{-1} g k_1.$$

Suponha que

$$g_t = k_{t-1}^{-1} g k_{t-1} \quad \text{para algum } t \geq 1$$

e mostremos que vale (3.2). Ora, de (3.1) temos que

$$g_{t+1} = r_t q_t = q_t^{-1} (q_t r_t) q_t = q_t^{-1} g_t q_t = (k_{t-1} q_t)^{-1} g (k_{t-1} q_t) = k_t^{-1} g k_t,$$

uma vez que $k_{t-1}q_t = k_t$. De onde segue (3.2).

Por outro lado, $g = q_1r_1$ e supondo

$$g^t = k_t\rho_t, \quad \text{com} \quad \rho_t = r_t \cdots r_1$$

para algum $t \geq 1$, temos que

$$g^{t+1} = g^t g = k_t \rho_t g = k_t q_{t+1} \rho_{t+1} = k_{t+1} \rho_{t+1},$$

onde para a terceira igualdade usamos que $\rho_t g = q_{t+1} \rho_{t+1}$ para todo $t \geq 1$. De fato,

$$q_2 \rho_2 = q_2 r_2 r_1 = r_1 q_1 r_1 = \rho_1 g$$

e, por hipótese de indução,

$$g_{t+2} \rho_{t+2} = q_{t+2} r_{t+2} \rho_{t+1} = r_{t+1} q_{t+1} \rho_{t+1} = r_{t+1} \rho_t g = \rho_{t+1} g,$$

como queríamos. □

O próximo resultado mostra a convergência do método QR e dá alguns detalhes a mais sobre as sequências g_t e k_t .

Teorema 3.2 (Método QR). Se os módulos dos autovalores de g são todos distintos então g_t converge para uma forma de Schur para g e k_t converge grosso modo para uma base de Schur para g . Mais precisamente, g_t é unitariamente conjugado à g e

- (i) as entradas da diagonal de g_t convergem para os autovalores de g ,
- (ii) as entradas abaixo da diagonal de g_t convergem para zero,
- (iii) as colunas de k_t convergem para uma base ortogonal de Schur para g , a menos de multiplicação à direita por matrizes diagonais unitárias.

Demonstração. Como $k_t = Q(g^t)$ converge para uma fibra g -invariante de π_Q (ver Teorema 2.25) e uma fibra g -invariante é composta por bases ortogonais de Schur para g , segue que

$$g_{t+1} = k_t^{-1} g k_t$$

converge para uma forma de Schur para g . Além disso, decorre do Corolário 2.24 que k_t converge para uma base ortogonal de Schur para g a menos de multiplicação à direita por diagonais unitárias. □

O próximo resultado mostra que o método QR é especialmente adaptado para matrizes hermitianas positivas.

Corolário 3.3. Se g é hermitiana positiva com autovalores distintos então g_t é hermitiana para todo t e converge para uma matriz diagonal unitariamente conjugada à g . Além disso, k_t converge para uma base ortogonal de autovetores de g .

Demonstração. Se g é hermitiana positiva com autovalores distintos então o módulo de seus autovalores é distintos uma vez que eles são reais e positivos. Pela proposição anterior, g_t é hermitiana para todo t , pois é unitariamente conjugado à g hermitiano (ver Proposição A.13), as entradas abaixo da diagonal de g_t tendem a 0 e as entradas da diagonal de g_t tendem aos autovalores de g . Como $g_t = (g_t)^*$ segue que as entradas acima da diagonal de g_t também tendem a 0, logo g_t tende a uma matriz diagonal unitariamente conjugada a g .

A segunda parte decorre do Teorema 2.25. De fato, como $k_t = Q(g^t)$, existe $q \in Q$ satisfazendo $k_t \rightarrow q$. Assim,

$$g_t = k_{t-1}^{-1} g k_{t-1} \rightarrow q^{-1} g q$$

é diagonal, pela primeira parte, de modo que as colunas de q constituem uma base ortogonal de autovetores de g . \square

Considerando $Q : G \rightarrow G$ a projeção em Q da decomposição QR de G , observe que método QR fornece a sequência

$$g_t = Q(g^t)^{-1} g Q(g^t).$$

Observação 3.4. (a) *O método funciona quando os autovalores de g têm módulos distintos pois, nesse caso, temos grosso modo que $Q(g^t)$ converge a uma base de Schur ortogonal de g e que g_t tende a uma forma de Schur de g .*

(b) *Isso deve ser contrastado com o chamado método das potências que se baseia no seguinte princípio. Dado $v \in \mathbb{C}^n$ não-nulo, a sequência de vetores normalizados*

$$\frac{g^t v}{|g^t v|}$$

converge a um autoespaço de g quando $t \rightarrow \infty$. Quando consideramos a sequência de matrizes unitárias $Q(g^t)$, estamos aplicando g^t simultaneamente a toda base canônica e normalizando a base obtida via Gram-Schmidt. Assim, o método QR é uma espécie método das potências simultâneo.

(c) *O método QR também fornece uma sequência q_t de matrizes unitárias obtidas trocando-se a ordem dos fatores de sucessivas decomposições de Iwasawa de g , de modo que*

$$Q(g^t) = q_1 \cdots q_t.$$

Numericamente, a importância do método QR reside nessa maneira de normalizar os vetores coluna de g^t , uma vez que a multiplicação de uma sequência de matrizes unitárias propaga menos erro do que a iteração de uma matriz arbitrária.

Para mais detalhes sobre o método das potências ou sobre a propagação de erros no método QR consulte o artigo [4] e suas referências.

3.2 Equação diferencial

O objetivo desta seção é estudarmos uma versão de tempo contínuo do método QR. Mais precisamente, fixe $X \in \mathfrak{g}$ e ponha $g^t = \exp(tX)$, com $t \in \mathbb{R}$. Nesta seção veremos que a curva

$$X(t) = Q(g^t)^{-1}XQ(g^t) = \text{Ad}(Q(g^t)^{-1})X, \quad t \in \mathbb{R},$$

é solução de uma EDO autônoma em \mathfrak{g} e converge para uma forma de Schur de X quando $t \rightarrow \infty$, desde que certas condições sobre X sejam satisfeitas. Em seguida, investigaremos algumas propriedades das soluções dessa EDO quando a condição inicial é hermitiana. Primeiro mostramos o seguinte.

Proposição 3.5. A aplicação

$$\phi : \mathbb{R} \times \mathfrak{g} \rightarrow \mathfrak{g}, \quad (t, X) \mapsto X(t) \tag{3.3}$$

é um fluxo diferenciável em \mathfrak{g} .

Demonstração. Uma vez que

$$\phi(t, X) = \text{Ad}(Q(\exp(tX))^{-1})X$$

é imediato que ϕ é diferenciável (ver Corolário A.10) e que $\phi(0, X) = X$ para todo $X \in \mathfrak{g}$. Pondo

$$k_t = Q(\exp(tX)),$$

temos que

$$\phi(t, X) = \text{Ad}(k_t^{-1})X.$$

Usando a R -invariância à direita e Q -equivariância à esquerda de $Q(\cdot)$, temos que

$$\begin{aligned} \phi(t+s, X) &= \text{Ad}(Q(\exp(t+s)X)^{-1})X \\ &= \text{Ad}(Q(\exp(sX)\exp(tX))^{-1})X \\ &= \text{Ad}(Q(\exp(sX)k_t)^{-1})X \\ &= \text{Ad}(Q(k_t k_t^{-1} \exp(sX) k_t)^{-1})X \\ &= \text{Ad}(Q(k_t^{-1} \exp(sX) k_t)^{-1} k_t^{-1})X \\ &= \text{Ad}(Q(\exp(s \text{Ad}(k_t^{-1})X))^{-1}) \text{Ad}(k_t^{-1})X \\ &= \text{Ad}(Q(\exp(s\phi(t, X)))^{-1}) \phi(t, X) \\ &= \phi(s, \phi(t, X)), \end{aligned}$$

como queríamos. □

Para obter o campo que corresponde à esse fluxo precisamos derivar $X(t)$ em $t = 0$. Para isso, primeiro precisamos calcular algumas derivadas.

3.2.1 Conjugação e Colchete

Dado $g \in G$, temos que a *conjugação* por g , denotada por

$$\begin{aligned} \text{Ad}(g) : \mathfrak{g} &\longrightarrow \mathfrak{g} \\ Y &\longmapsto \text{Ad}(g)Y = gYg^{-1}, \end{aligned}$$

fornece a aplicação adjunta (grande)

$$\begin{aligned} \text{Ad} : G &\longrightarrow \text{Gl}(\mathfrak{g}) \\ g &\longmapsto \text{Ad}(g) \end{aligned}$$

que é um homomorfismo de grupos,

$$\text{Ad}(gh) = \text{Ad}(g)\text{Ad}(h), \quad \forall g, h \in G,$$

como é fácil verificar.

Dado $X \in \mathfrak{g}$, temos que o *colchete* por X , dado por

$$[X, Y] = XY - YX,$$

satisfaz

$$\text{Ad}(g)[X, Y] = [\text{Ad}(g)X, \text{Ad}(g)Y], \quad (3.4)$$

como é fácil verificar. Além disso, o colchete por X define o operador linear

$$\begin{aligned} \text{ad}(X) : \mathfrak{g} &\longrightarrow \mathfrak{g} \\ Y &\longmapsto \text{ad}(X)Y = [X, Y], \end{aligned}$$

que fornece a aplicação adjunta (pequena)

$$\begin{aligned} \text{ad} : \mathfrak{g} &\longrightarrow \mathfrak{gl}(\mathfrak{g}) \\ X &\longmapsto \text{ad}(X). \end{aligned}$$

Proposição 3.6. Temos que

$$\text{Ad}'(I)X = \text{ad}(X), \quad \text{para todo } X \in \mathfrak{g}.$$

Em particular, pondo $\text{Ad}_X : G \rightarrow \mathfrak{g}$, $g \mapsto \text{Ad}(g)X$, temos que

$$\text{Ad}'_X(I)Y = [X, Y].$$

Demonstração. Dado $X \in \mathfrak{g}$, temos que a curva

$$\alpha : t \in \mathbb{R} \longmapsto \exp(tX) \in G$$

é tal que

$$\alpha(0) = I \quad \text{e} \quad \alpha'(0) = X.$$

Assim,

$$\begin{aligned}
(\text{Ad}'(I)X)(Y) &= \left. \frac{d}{dt}(\text{Ad} \circ \alpha)(t)Y \right|_{t=0} \\
&= \left. \frac{d}{dt} \exp(tX)Y \exp(-tX) \right|_{t=0} \quad (\exp(tX)^{-1} = \exp(-tX)) \\
&= XY - YX \quad (\text{pela regra do produto, } t = 0) \\
&= \text{ad}(X)Y,
\end{aligned}$$

para todo $Y \in \mathfrak{g}$. Donde,

$$\text{Ad}'(I)X = \text{ad}(X).$$

□

Proposição 3.7. A inversão $i : G \rightarrow G$ definida por $i(g) = g^{-1}$ tem derivada

$$i'(I)X = -X, \quad \forall X \in \mathfrak{g}.$$

Demonstração. A curva

$$\alpha : t \in \mathbb{R} \mapsto \exp(tX) \in G$$

é tal que

$$\alpha(0) = I \quad \text{e} \quad \alpha'(0) = X.$$

Assim,

$$i'(I)X = \left. \frac{d}{dt}(i \circ \alpha)(t) \right|_{t=0} = \left. \frac{d}{dt}(\exp(tX))^{-1} \right|_{t=0} = \left. \frac{d}{dt} \exp(-tX) \right|_{t=0} = -X.$$

□

3.2.2 Decomposição de Iwasawa de \mathfrak{g}

Nesta seção vamos considerar uma decomposição de \mathfrak{g} que é análoga à decomposição QR de G e será bastante usada neste capítulo. Tal decomposição nos permitirá calcular a derivada das projeções $Q : G \rightarrow Q$ e $R : G \rightarrow R$ da decomposição QR (ver final da Seção A.4).

Para tanto, considere o subespaço \mathfrak{q} das matrizes *anti-hermitianas* e o subespaço \mathfrak{r} das matrizes triangulares superiores com entradas reais na diagonal.

Proposição 3.8. Temos que

$$\mathfrak{g} = \mathfrak{q} \oplus \mathfrak{r},$$

a qual chamamos de *decomposição de Iwasawa* de \mathfrak{g} .

Demonstração. (i) $\mathfrak{q} \cap \mathfrak{r} = \{0\}$. De fato, seja $X \in \mathfrak{q} \cap \mathfrak{r}$. Então X é diagonal tendo em vista que é triangular superior e que $\overline{X}^\dagger = -X$ também o é. Além disso, a diagonal de X deve ser nula, uma vez que suas entradas são reais e X é anti-hermitiana. Portanto, $X = 0$. Logo,

$$\mathfrak{q} \cap \mathfrak{r} = \{0\}.$$

(ii) $\mathfrak{g} = \mathfrak{q} + \mathfrak{r}$. De fato, seja $X \in \mathfrak{g}$. Escreva

$$X = X_- + X_0 + X_+$$

onde X_- é a parte triangular estritamente inferior de X , X_+ é a parte triangular estritamente superior de X e X_0 é a parte diagonal de X . Assim,

$$\begin{aligned} X &= X_- - (X_-)^* + X_0 + X_+ + (X_-)^* \\ &= (X_- - (X_-)^* + i\text{Im}X_0) + (\text{Re}X_0 + X_+ + (X_-)^*), \end{aligned}$$

onde $\text{Im}X_0$ é a matriz cujas entradas são as partes imaginárias correspondentes às entradas de X_0 e $\text{Re}X_0$ é a matriz cujas entradas são as partes reais correspondentes às entradas de X_0 .

Afirmamos que

$$X_{\mathfrak{q}} := X_- - (X_-)^* + i\text{Im}X_0 \quad \text{pertence à } \mathfrak{q}, \quad (3.5)$$

$$X_{\mathfrak{r}} := \text{Re}X_0 + X_+ + (X_-)^* \quad \text{pertence à } \mathfrak{r}. \quad (3.6)$$

Com efeito,

$$\begin{aligned} (X_{\mathfrak{q}})^* &= (X_- - (X_-)^* + i\text{Im}X_0)^* \\ &= (X_-)^* - X_- + (i\text{Im}X_0)^* \\ &= -(X_- - (X_-)^* + i\text{Im}X_0) \\ &= -X_{\mathfrak{q}}. \end{aligned}$$

Logo, $X_{\mathfrak{q}} \in \mathfrak{q}$. Por outro lado, como X_+ e $(X_-)^*$ estão em \mathfrak{r} com diagonais nulas e X_0 é diagonal, temos que $X_{\mathfrak{r}} = \text{Re}X_0 + X_+ + (X_-)^*$ é triangular superior e sua diagonal é a diagonal de $\text{Re}X_0$ cujas entradas são reais. Logo, $X_{\mathfrak{r}} \in \mathfrak{r}$ e portanto,

$$\mathfrak{g} = \mathfrak{q} + \mathfrak{r}.$$

De (i) e (ii) segue que $\mathfrak{g} = \mathfrak{q} \oplus \mathfrak{r}$. □

Assim, toda matriz $X \in \mathfrak{g}$ se decompõe unicamente como

$$X = X_{\mathfrak{q}} + X_{\mathfrak{r}}, \quad \text{com } X_{\mathfrak{q}} \in \mathfrak{q} \text{ e } X_{\mathfrak{r}} \in \mathfrak{r},$$

e portanto, estão bem definidas as projeções

$$\begin{array}{ccc} \mathfrak{q} : \mathfrak{g} & \longrightarrow & \mathfrak{q} & \text{e} & \mathfrak{r} : \mathfrak{g} & \longrightarrow & \mathfrak{r} \\ X & \longmapsto & X_{\mathfrak{q}} & & X & \longmapsto & X_{\mathfrak{r}} \end{array}$$

que, por um abuso de notação, denotamos pelas respectivas letras dos subespaços \mathfrak{q} e \mathfrak{r} .

Considere o subespaço \mathfrak{u} das matrizes triangulares superiores.

Lema 3.9. Seja $\alpha : \mathbb{R} \longrightarrow \mathfrak{g}$, uma curva diferenciável tal que $\alpha(0) = I$. Então,

(a) Se $\alpha(t) \in Q$ para todo t , então $\alpha'(0) \in \mathfrak{q}$.

(b) Se $\alpha(t) \in R$ para todo t , então $\alpha'(0) \in \mathfrak{r}$.

(c) Se $\alpha(t) \in U$ para todo t , então $\alpha'(0) \in \mathfrak{u}$.

Demonstração. Se $\alpha(t) \in Q$ para todo t , então

$$I = \alpha(t)\alpha(t)^* \quad \forall t \quad \text{e} \quad \frac{d}{dt}\alpha(t)^* = \alpha'(t)^*,$$

pela linearidade da transposta conjugada $X \mapsto X^*$. Da regra do produto obtemos

$$0 = \frac{d}{dt}(\alpha(t)\alpha(t)^*) = \alpha(t)\alpha'(t)^* + \alpha'(t)\alpha(t)^*.$$

Logo, em $t = 0$ temos que

$$0 = I\alpha'(0)^* + \alpha'(0)I.$$

Portanto, $\alpha'(0)^* = -\alpha'(0)$, ou seja, $\alpha'(0) \in \mathfrak{q}$. Com isto mostramos o item mostramos o item (a).

Para o item (b), note que \mathfrak{r} é espaço vetorial e que $R \subset \mathfrak{r}$. Logo a derivada de uma curva em R está em \mathfrak{r} . O item (c) é análogo ao (b), sendo \mathfrak{u} espaço vetorial e $U \subset \mathfrak{u}$. \square

Proposição 3.10. Temos que

$$Q'(I)X = \mathfrak{q}(X) \quad \text{e} \quad R'(I)X = \mathfrak{r}(X).$$

Demonstração. As aplicações

$$\begin{array}{ccc} m : G \times G & \longrightarrow & G & \text{e} & \psi : G & \longrightarrow & G \times G \\ (g_1, g_2) & \longmapsto & g_1 g_2 & & g & \longmapsto & (Q(g), R(g)) \end{array}$$

são ambas diferenciáveis com $m'(I, I)(X, Y) = Y + X$ (ver Proposição A.16) e

$$\psi'(g)X = (Q'(g)X, R'(g)X).$$

Além disso, segue da decomposição QR que

$$\text{Id}_G = m \circ \psi.$$

Assim, derivando $\text{Id}_G = m \circ \psi$ na direção de X , temos pela regra da cadeia que

$$\begin{aligned} X &= (m \circ \psi)'(I)X = m'(I, I) \circ \psi'(I)X \\ &= m'(I, I)(Q'(I)X, R'(I)X) \\ &= Q'(I)X + R'(I)X \end{aligned}$$

Além disso, decorre da proposição anterior que $Q'(I)X \in \mathfrak{q}$ e $R'(I)X \in \mathfrak{r}$. Assim, pela unicidade da decomposição $\mathfrak{g} = \mathfrak{q} \oplus \mathfrak{r}$, temos que

$$Q'(I)X = \mathfrak{q}(X) \quad \text{e} \quad R'(I)X = \mathfrak{r}(X).$$

□

3.2.3 Propriedades da equação diferencial

Proposição 3.11. O campo do fluxo $\phi(t, X)$, definido em (3.3), é dado por

$$\mathfrak{g} \rightarrow \mathfrak{g}, \quad X \mapsto [X, \mathfrak{q}(X)].$$

Demonstração. Devemos derivar

$$\phi(t, X) = \text{Ad}(Q(g^t)^{-1})X$$

em $t = 0$. Considere

$$\text{Ad}_X : G \rightarrow \mathfrak{g}, \quad \text{Ad}_X(g) = \text{Ad}(g)X$$

que possui derivada

$$\text{Ad}'_X(I)Y = \text{ad}(Y)X = [Y, X], \quad Y \in \mathfrak{g},$$

pela Proposição 3.6. Temos que

$$X(t) = \text{Ad}_X \circ i \circ Q \circ \exp(tX),$$

de modo que, derivando em $t = 0$, temos

$$X'(0) = \text{Ad}'_X(I) \circ i'(I) \circ Q'(I)X = \text{Ad}'_X(I) \circ (-\mathfrak{q}(X)) = [-\mathfrak{q}(X), X] = [X, \mathfrak{q}(X)],$$

onde usamos as Proposições 3.7 e 3.10. □

Teorema 3.12. Seja $X(t)$ a solução do PVI

$$\begin{cases} X'(t) &= [X(t), \mathfrak{q}(X(t))] \\ X(0) &= X. \end{cases}$$

Então,

- (i) $X(t)$ é unitariamente conjugada à X e definida para todo $t \in \mathbb{R}$.

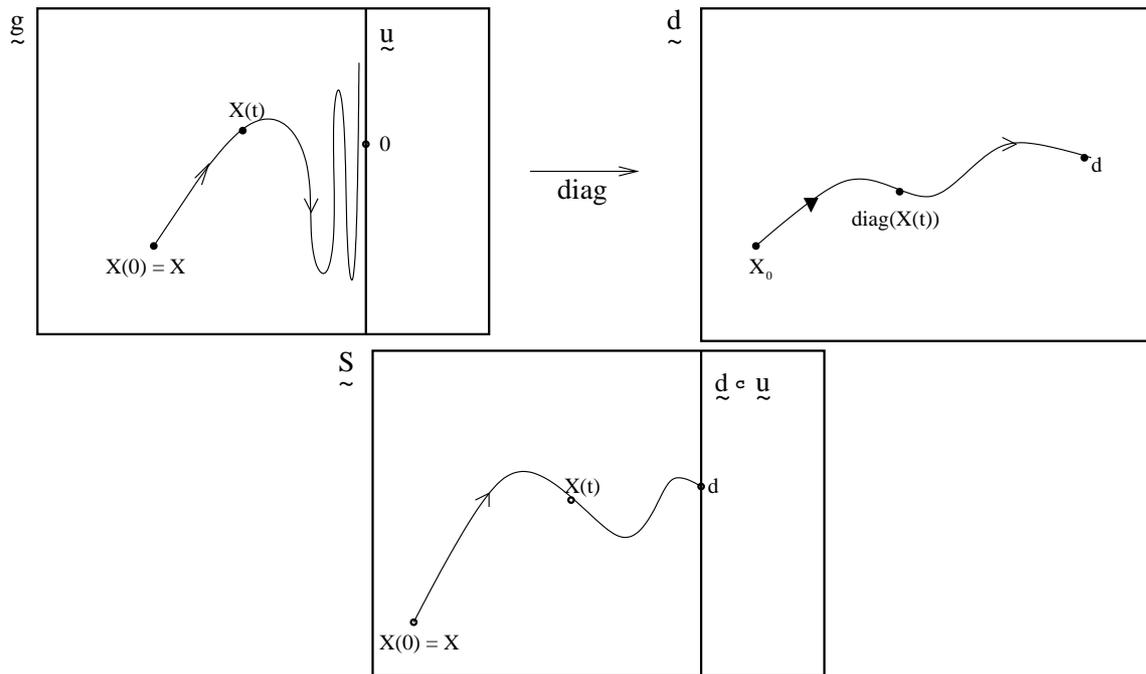


Figura 3.1: Itens (ii) e (iii) do Teorema 3.12

- (ii) Se os autovalores de X têm partes reais distintas, então $X(t)$ converge para uma forma de Schur para X . Mais precisamente,

$$X(t) \longrightarrow \mathbf{u} \quad \text{quando} \quad t \longrightarrow \infty$$

e sua diagonal converge para um vetor cujas entradas são os autovalores de X .

- (iii) Se X é hermitiana então $X(t)$ é hermitiana para todo t . Se, além disso, os autovalores de X são todos distintos, então $X(t)$ converge para uma matriz diagonal.

Demonstração. (i) Da proposição anterior, o fluxo da EDO é dado por

$$X(t) = \text{Ad}(k_t^{-1})X, \quad \text{onde} \quad k_t = Q(\exp(tX)) \quad \text{com} \quad t \in \mathbb{R}.$$

Assim, $X(t)$ é unitariamente conjugado à $X(0) = X$ para todo $t \in \mathbb{R}$.

- (ii) Se os autovalores de X têm partes reais distintas então X tem autovalores distintos, logo é diagonalizável e à cada autovalor de X corresponde um autovetor. Ponha $g = \exp(X)$. Seja λ autovalor de X com autovetor correspondente v , então

$$gv = \exp(X)v = \sum_{k=0}^{\infty} \frac{X^k v}{k!} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} v = e^{\lambda} v.$$

Então os autovalores de g são da forma e^{λ} , com λ autovalor de X . Pondo $\lambda = a + ib$, temos

$$|e^{\lambda}| = |e^a| |e^{ib}| = |e^a| = |e^{\text{Re}(\lambda)}| = e^{\text{Re}(\lambda)}.$$

Logo, os autovalores de g têm módulos distintos. Pela Observação 2.18 temos então que

$$\lim_{t \rightarrow \infty} g^t e = qe \in \mathbb{F}, \quad t \in \mathbb{R}, \quad (3.7)$$

para algum $q \in Q$. Segue que qe é ponto fixo de g^t em \mathbb{F} , para todo $t \in \mathbb{R}$, isto é

$$g^t qe = qe \implies q^{-1} g^t qe = e, \quad t \in \mathbb{R}.$$

Assim, temos que

$$q^{-1} g^t q = \exp(t \operatorname{Ad}(q^{-1})X) \in G_e = U$$

para todo $t \in \mathbb{R}$ e, derivando essa curva em $t = 0$, segue do Lema 3.9 que

$$\operatorname{Ad}(q^{-1})X \in \mathfrak{u}.$$

Da equação (3.7) temos que $Q(g^t)$ converge para q a menos de multiplicação à direita por diagonais unitárias. Então, $X(t) = \operatorname{Ad}(Q(g^t)^{-1})X$ converge para $\operatorname{Ad}(q^{-1})X \in \mathfrak{u}$ a menos de conjugação por diagonais unitárias. Assim, concluímos que

$$X(t) \longrightarrow \mathfrak{u} \quad \text{quando } t \longrightarrow \infty,$$

e que a sua diagonal converge para um vetor cujas entradas são os autovalores de X , como queríamos.

- (iii) Pelo item (i), se X é Hermitiana, então $X(t)$ também o é (ver Proposição A.13). Se X hermitiana tem autovalores distintos, como eles são todos reais, suas partes reais são distintas. Segue do item (ii) que a parte triangular estritamente inferior de $X(t)$ converge para zero e, como $X(t)^* = X(t)$, sua parte triangular estritamente superior também converge para zero. Assim $X(t)$ converge para uma matriz diagonal, como queríamos. □

3.2.4 Função altura em \mathfrak{g}

Dado um produto hermitiano (\cdot, \cdot) num espaço vetorial complexo e fixado um vetor H , a altura do vetor X com respeito a H é dada por

$$\operatorname{Re}(X, H).$$

Fixando o produto hermitiano de Frobenius em \mathfrak{g} e $H \in \mathfrak{g}$ diagonal hermitiana e positiva, veremos que, para uma condição inicial $X \in \mathfrak{g}$ hermitiana, a altura do fluxo $X(t) = Q(g^t)^{-1} X Q(g^t)$ com respeito a H é decrescente (ver Teorema 3.16).

O produto de Frobenius em \mathfrak{g} é dado pela aplicação

$$\begin{aligned} (\cdot, \cdot) : \mathfrak{g} \times \mathfrak{g} &\longrightarrow \mathbb{C} \\ (X, Y) &\longmapsto \operatorname{tr}(XY^*), \end{aligned} \quad (3.8)$$

onde $\operatorname{tr}(\cdot)$ é o funcional linear *traço* em \mathfrak{g} definido por $\operatorname{tr}(X) =$ soma das entradas da diagonal de X . Uma vez que a ij -ésima entrada de $XY^* = X\bar{Y}^\dagger$ é

$$(XY^*)_{ij} = \sum_{k=1}^n X_{ik} \bar{Y}_{jk},$$

temos que

$$(X, Y) = \operatorname{tr}(XY^*) = \sum_{i,k=1}^n X_{ik}\bar{Y}_{ik},$$

que é o produto hermitiano canônico em \mathbb{C}^{n^2} , na identificação natural de $\mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$ com \mathbb{C}^{n^2} . Dados $\lambda \in \mathbb{C}$ e $X, Y, Z \in \mathfrak{g}$, temos as seguintes propriedades

- (1) $(\lambda X + Y, Z) = \lambda(X, Z) + (Y, Z)$, em virtude da linearidade do traço e da bilinearidade na multiplicação de matrizes.
- (2) $(X, X) = \sum_{i,k} X_{ik}\bar{X}_{ik} = \sum_i |X_i|^2$, onde X_i é a i^{a} linha de X . Então, $(X, X) \geq 0$ e $(X, X) = 0$ se, e somente se, $X = 0$.
- (3) $(X, Y) = \overline{(Y, X)}$, uma vez que $\sum_{i,k} X_{ik}\bar{Y}_{ik} = \overline{\sum_{i,k} Y_{ik}\bar{X}_{ik}}$.

Portanto, (3.8) é um produto hermitiano em \mathfrak{g} . Além disso,

$$\begin{aligned} (X^*, Y^*) &= \operatorname{tr}(X^*Y^*) \\ &= \operatorname{tr}(YX^*) \quad (\operatorname{tr}(AB) = \operatorname{tr}(BA), \forall A, B \in \mathfrak{g}) \\ &= (Y, X) \\ &= \overline{(X, Y)} \quad (\text{pela propriedade (3)}). \end{aligned}$$

Logo,

$$(X^*, Y^*) = \overline{(X, Y)}, \quad \forall X, Y \in \mathfrak{g}. \quad (3.9)$$

Temos ainda que

$$\begin{aligned} ([X, Y], Z) &= -([Y, X], Z) = -\operatorname{tr}([Y, X]Z^*) \\ &= -\operatorname{tr}(Y[X, Z^*]) = \operatorname{tr}(Y[X^*, Z]^*) \\ &= (Y, [X^*, Z]), \end{aligned}$$

onde usamos que $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ e que $[A, B]^* = -[A^*, B^*]$ para todos $A, B \in \mathfrak{g}$, como é fácil verificar. Em particular, se X é anti-hermitiana então $\operatorname{ad}(X)$ também o é, isto é, se $X^* = -X$, então

$$(\operatorname{ad}(X)Y, Z) = (Y, -\operatorname{ad}(X)Z), \quad \forall Y, Z \in \mathfrak{g}. \quad (3.10)$$

Considere o conjunto das matrizes diagonais

$$\mathfrak{d} = \{X \in \mathfrak{g} : X_{ij} = 0, \text{ se } i \neq j\},$$

o conjunto das matrizes triangulares estritamente inferiores

$$\mathfrak{n}^- = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ se } i \leq j\}$$

e o conjunto das matrizes triangulares estritamente superiores

$$\mathfrak{n}^+ = \{X \in \mathfrak{g} : X_{ij} = 0 \text{ se } i \geq j\}.$$

Proposição 3.13. Com relação ao produto hermitiano (3.8), temos a decomposição ortogonal

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{d} \oplus \mathfrak{n}^+.$$

Demonstração. Que $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{d} + \mathfrak{n}^+$ é imediato. Além disso, se $d = \text{diag}(a_1, \dots, a_n) \in \mathfrak{d}$, então $(Xd^*)_{ij} = X_{ij}\bar{a}_j$ e daí,

$$(X, d) = \sum_i (Xd^*)_{ii} = \sum_i X_{ii}\bar{a}_i.$$

Como as diagonais das matrizes em \mathfrak{n}^- e \mathfrak{n}^+ são nulas, temos que

$$(\mathfrak{n}^-, \mathfrak{d}) = (\mathfrak{n}^+, \mathfrak{d}) = 0. \quad (3.11)$$

Por outro lado, se $X \in \mathfrak{n}^-$ e $Y \in \mathfrak{n}^+$, então $X_{ik} = 0$ para $i \leq k$ e $Y_{ik} = 0$ para $i \geq k$. Então,

$$(X, Y) = \sum_{i,k} X_{ik}\bar{Y}_{ik} = 0.$$

Logo,

$$(\mathfrak{n}^-, \mathfrak{n}^+) = 0. \quad (3.12)$$

As equações (3.11) e (3.12) mostram que a soma $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{d} + \mathfrak{n}^+$ é ortogonal. \square

Para $k, l = 1, \dots, n$, considere a matriz

$$E_{kl} \in \mathfrak{g},$$

cuja kl -ésima entrada é 1 e as demais são nulas. Ou seja, $E_{kl} = (x_{ij})$ com $x_{ij} = \delta_{ik}\delta_{jl}$, onde

$$\delta_{ij} = \begin{cases} 0, & \text{se } i \neq j \\ 1, & \text{se } i = j. \end{cases}$$

Proposição 3.14. Seja $H = \text{diag}(a_1, \dots, a_n) \in \mathfrak{d}$. Então, o conjunto $\{E_{kl} : 1 \leq k, l \leq n\}$ é uma base ortonormal de autovetores de $\text{ad}(H)$ em \mathfrak{g} , tal que

$$\text{ad}(H)E_{kl} = (a_k - a_l)E_{kl}.$$

Em particular,

$$[\mathfrak{d}, \mathfrak{n}^-] \subset \mathfrak{n}^- \quad \text{e} \quad [\mathfrak{d}, \mathfrak{n}^+] \subset \mathfrak{n}^+. \quad (3.13)$$

Demonstração. Seja $X = (X_{ij}) \in \mathfrak{g}$. Então,

$$X = \sum_{k,l=1}^n X_{kl}E_{kl}$$

mostra que $\{E_{kl} : 1 \leq k, l \leq n\}$ gera \mathfrak{g} . É fácil ver que este conjunto é linearmente independente e

$$(E_{ij}, E_{kl}) = \delta_{ik}\delta_{jl} = \begin{cases} 0, & \text{se } E_{ij} \neq E_{kl} \\ 1, & \text{se } E_{ij} = E_{kl}. \end{cases}$$

Portanto, $\{E_{kl} : 1 \leq k, l \leq n\}$ é uma base ortonormal de \mathfrak{g} .

Por outro lado,

$$HE_{kl} - E_{kl}H = (a_i\delta_{ik}\delta_{jl}) - (a_j\delta_{ik}\delta_{jl}) = (a_k - a_l)(\delta_{ik}\delta_{jl}) = (a_k - a_l)E_{kl}.$$

Assim, pela linearidade de $\text{ad}(H)$, vale

$$\text{ad}(H)X = \sum_{k,l=1}^n X_{kl}\text{ad}(H)E_{kl} = \sum_{k,l=1}^n (a_k - a_l)X_{kl}E_{kl}, \quad \forall X \in \mathfrak{g}, \quad (3.14)$$

de onde obtemos o resultado desejado. \square

Fixe $H = \text{diag}(a_1, \dots, a_n)$ com $a_i \in \mathbb{R}$ para todo i e considere a aplicação $f : \mathfrak{g} \rightarrow \mathbb{R}$ dada por

$$f(Y) = \text{Re}(Y, H),$$

a qual chamamos *função altura* em \mathfrak{g} com respeito a H . Dado $X \in \mathfrak{g}$, tomemos $X(t)$ solução do P.V.I.

$$\begin{cases} X'(t) &= [X(t), \mathfrak{q}(X(t))] \\ X(0) &= X. \end{cases}$$

Para calcular a derivada de $f(X(t))$ com relação à t consideramos

Proposição 3.15. Temos que

$$\frac{d}{dt}(X(t), H) = -\{(X(t)_-, [H, X(t)_-]) + (X(t)_+, [H, X(t)_-]^*)\}, \quad t \in \mathbb{R}.$$

Demonstração. Ora,

$$\begin{aligned} \frac{d}{dt}(X(t), H) &= (X'(t), H) = ([X(t), \mathfrak{q}(X(t))], H) \\ &= (-[\mathfrak{q}(X(t)), X(t)], H) \\ &= (X(t), [\mathfrak{q}(X(t)), H]) \quad (\text{por (3.10), onde } \mathfrak{q}(X(t)) \text{ é anti-hermitiana}) \\ &= -(X(t), [H, \mathfrak{q}(X(t))]) \\ &= -(Y, [H, \mathfrak{q}(Y)]), \end{aligned}$$

onde fizemos $Y = X(t)$. Da equação (3.5) temos que

$$\mathfrak{q}(Y) = Y_- - (Y_-)^* + i\text{Im}Y_0,$$

de onde segue que

$$\begin{aligned}
[H, \mathfrak{q}(Y)] &= [H, Y_-] - [H, (Y_-)^*] \quad (\text{pois } H \in \mathfrak{d} \text{ comuta com } i\text{Im}Y_0 \in \mathfrak{d}) \\
&= [H, Y_-] - [H^*, (Y_-)^*] \quad (\text{pois } H^* = H) \\
&= \underbrace{[H, Y_-]}_{\in \mathfrak{n}^-} + \underbrace{[H, Y_-]^*}_{\in \mathfrak{n}^+} \quad (\text{por (3.13)}).
\end{aligned}$$

Daí,

$$\begin{aligned}
(Y, [H, \mathfrak{q}(Y)]) &= (Y_- + Y_0 + Y_+, [H, Y_-] + [H, Y_-]^*) \\
&= (Y_-, [H, Y_-]) + (Y_+, [H, Y_-]^*) \quad (\text{ver Proposição 3.13}).
\end{aligned}$$

□

Teorema 3.16. Se $X(0) = X$ é hermitiana, então $f(X(t))$ é estritamente decrescente e é constante se, e somente se, $X(0) = X$ é diagonal. Em particular, f é função de Lyapunov para esse sistema dinâmico.

Demonstração. Sendo $X = X(0)$ hermitiana, temos que $Y = X(t)$ é hermitiana para todo $t \in \mathbb{R}$ por ser conjugação unitária de X (ver início desta seção). Logo, $Y_+ = (Y_-)^*$ e portanto,

$$\begin{aligned}
(Y, [H, \mathfrak{q}(Y)]) &= (Y_-, [H, Y_-]) + ((Y_-)^*, [H, Y_-]^*) \\
&= (Y_-, [H, Y_-]) + \overline{(Y_-, [H, Y_-])},
\end{aligned}$$

onde usamos a propriedade (3.9).

Supondo $a_1 > \dots > a_n > 0$, temos que

$$[H, Y_-] = \sum_{i < j} (a_i - a_j) Y_{ij} E_{ij},$$

o que nos dá

$$(Y_-, [H, Y_-]) = \sum_{i < j} (a_i - a_j) |Y_{ij}|^2 \in \mathbb{R}.$$

Logo,

$$(Y, [H, \mathfrak{q}(Y)]) = (Y_-, [H, Y_-]) + \overline{(Y_-, [H, Y_-])} = 2 \sum_{i < j} \underbrace{(a_i - a_j)}_{>0} |Y_{ij}|^2 \geq 0$$

e é igual a zero se, e somente se, $Y_{ij} = 0$ para $i < j$, ou seja, Y é diagonal. Como $\frac{d}{dt}(f(X(t))) = -(Y, [H, \mathfrak{q}(Y)])$, obtemos o resultado desejado. □

Apêndice A

Tópicos Matriciais

A.1 Operações elementares e matrizes de permutação

Dada uma matriz $X \in \mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$, consideramos as seguintes *operações elementares* sobre as colunas de X .

(M) Multiplicação da j^{a} coluna de X por um escalar λ e

(C) Substituição da j^{a} coluna de X pela coluna j mais λ vezes a coluna k .

Denotaremos por $X \rightarrow X'$ para indicar a matriz X' obtida a partir de X via operações elementares. Pondo C_j a j^{a} coluna de X , reescrevemos as operações acima como segue.

(M) $C_j \rightarrow \lambda C_j$ e

(C) $C_j \rightarrow C_j + \lambda C_k$.

Além disso, ponha

$$(\cdots \mid v \mid_j \cdots) \in \mathfrak{g}$$

a matriz identidade salvo a j^{a} coluna, que tem o vetor v . Podemos realizar as operações elementares (M) e (C) por multiplicação à direita de X pelas matrizes $E_M = (\cdots \mid \lambda e_j \mid_j \cdots)$ e $E_C = (\cdots \mid e_j + \lambda e_k \mid_j \cdots)$, respectivamente, como segue.

(M) $X \rightarrow X E_M$ e

(C) $X \rightarrow X E_C$.

Note que E_M é diagonal e E_C tem 1 nas entradas da diagonal, é triangular superior se $k \leq j$ e triangular inferior se $k \geq j$. Observe ainda que aplicações consecutivas de (C) nos dá

$$C_j \rightarrow \sum_k \lambda_k C_k : X \rightarrow X \left(\cdots \mid \sum_k \lambda_k e_k \mid_j \cdots \right),$$

que será realizada pela multiplicação à direita de X por uma matriz triangular superior desde que os índices k sejam tais que $k \leq j$ ou por uma matriz triangular inferior se $k \geq j$ para todo k .

Observação A.1. *Analogamente às operações elementares sobre colunas, podemos realizar operações elementares sobre linhas como segue. Se $X \in \mathfrak{g}$ tem linhas L_1, \dots, L_n , as operações*

(M') $L_j \rightarrow \lambda L_j$ e

$$(C') L_j \longrightarrow L_j + \lambda L_k$$

indicam multiplicação da j^a linha pelo escalar λ e a substituição da j^a linha pela combinação da linha j com λ vezes a linha k . Estas operações podem ser realizadas a partir de operações sobre as colunas de X^\dagger e correspondem a multiplicar X por matrizes adequadas à esquerda.

$$(M') X \rightarrow E_M^\dagger X \text{ e}$$

$$(C') X \rightarrow E_C^\dagger X,$$

onde E_M e E_C foram definidas acima.

Observações análogas as que fizemos para operações elementares sobre colunas valem para operações elementares sobre linhas. Por exemplo, E_M^\dagger é diagonal e E_C^\dagger , que tem 1 nas entradas da diagonal, é triangular inferior se $k \leq j$ e triangular superior se $k \geq j$.

Agora estudaremos as matrizes de permutação. Dado $\sigma \in S_n$, ponha

$$I_\sigma = (e_{\sigma(1)} \mid \cdots \mid e_{\sigma(n)}) \in Q,$$

onde Q é o conjunto das matrizes unitárias, ou seja, matrizes cujos vetores coluna são ortonormais entre si. Equivalentemente, I_σ é um operador linear em \mathbb{C}^n definido por

$$I_\sigma e_i = e_{\sigma(i)}, i = 1, \dots, n,$$

isto é, permuta vetores da base canônica de \mathbb{C}^n segundo σ . Chamamos os elementos de

$$P = \{I_\sigma : \sigma \in S_n\}$$

de matrizes de permutação.

Proposição A.2. Sejam $\sigma, \tau \in S_n$. Então valem as seguintes propriedades.

$$(1) I_\sigma I_\sigma^\dagger = I.$$

$$(2) I_\tau I_\sigma = I_{\tau\sigma}. \text{ Em particular, } I_\sigma^{-1} = I_{\sigma^{-1}}.$$

(3) I_σ age em $X \in \mathfrak{g}$ permutando suas colunas segundo σ quando multiplicada à direita. Ou seja,

$$(C_j) \rightarrow (C_{\sigma(j)}) : X \rightarrow XI_\sigma.$$

(4) I_σ age em $X \in \mathfrak{g}$ permutando suas linhas segundo σ^{-1} quando multiplicada à esquerda. Ou seja,

$$(L_i) \rightarrow (L_{\sigma^{-1}(i)}) : X \rightarrow I_\sigma X.$$

Em particular, temos que

$$I_\tau^{-1}(a_{ij})I_\sigma = (a_{\tau(i)\sigma(j)}).$$

Demonstração. (1) Decorre de I_σ ser unitária com entradas reais.

(2) Temos

$$(I_\tau \circ I_\sigma)e_i = I_\tau(I_\sigma e_i) = I_\tau e_{\sigma(i)} = e_{\tau\sigma(i)} = I_{\tau\sigma}e_i, \quad \text{para todo } i = 1, \dots, n.$$

Logo, $I_\tau \circ I_\sigma = I_{\tau\sigma}$ e

$$I = I_{\sigma\sigma^{-1}} = I_\sigma \circ I_{\sigma^{-1}},$$

que nos dá $I_\sigma^{-1} = I_{\sigma^{-1}}$.

(3) Para todo $Y \in \mathfrak{g}$, temos $Y = (Ye_1 | \dots | Ye_n)$, ou seja, o produto Ye_k é a k^{a} coluna de Y . Assim,

$$k^{\text{a}} \text{ coluna de } XI_\sigma = (XI_\sigma)e_k = Xe_{\sigma(k)} = \sigma(k)^{\text{a}} \text{ coluna de } X.$$

Portanto,

$$XI_\sigma = (C_1 | \dots | C_n)I_\sigma = (C_{\sigma(1)} | \dots | C_{\sigma(n)}).$$

(4) Decorre de (3), uma vez observado que (1) e (2) nos dá $I_{\sigma^{-1}} = I_\sigma^\dagger$ e então

$$(I_\sigma(a_{ij}))^\dagger = (a_{ji})I_\sigma^\dagger = (a_{ji})I_{\sigma^{-1}} = (a_{j\sigma^{-1}(i)}).$$

Logo,

$$I_\sigma(a_{ij}) = (I_\sigma(a_{ij}))^{\dagger\dagger} = (a_{j\sigma^{-1}(i)})^\dagger = (a_{\sigma^{-1}(i)j}),$$

como queríamos. □

Corolário A.3. P é subgrupo de G tal que a aplicação $\sigma \in S_n \mapsto I_\sigma \in P$ é um isomorfismo de grupos. Além disso, se $X \in \mathfrak{g}$ e $\sigma \in S_n$, então a diagonal de $I_\sigma^{-1}XI_\sigma$ é a diagonal de X permutada por σ . Em particular, se d é diagonal então $I_\sigma^{-1}dI_\sigma$ também é diagonal para todo $\sigma \in S_n$.

Demonstração. A propriedade (2) da proposição acima garante que P é um subgrupo de G , com unidade $I_1 = I$, tal que $\sigma \in S_n \mapsto I_\sigma \in P$ é um isomorfismo de grupos.

Para a segunda parte, basta notar que pondo $X = (X_{ij})$, temos que

$$I_\sigma^{-1}XI_\sigma = (X_{\sigma(i)\sigma(j)}).$$

□

A.2 Conjugações que fixam a diagonal

Considere o subgrupo de G

$$U' = \{g \in U : \text{diag}(g) = I\}$$

formado pelas matrizes triangulares superiores com entradas na diagonal todas iguais a 1 e o subgrupo de G

$$L' = \{g \in L : \text{diag}(g) = I\}$$

formado pelas matrizes triangulares inferiores com entradas na diagonal todas iguais a 1. O próximo resultado é bastante importante nesta dissertação.

Lema A.4. (i) Para d diagonal e $X \in \mathfrak{g}$, temos que a ij -ésima entrada de $d^{-1}Xd$ é

$$(d^{-1}Xd)_{ij} = d_{ii}^{-1}X_{ij}d_{jj}. \quad (\text{A.1})$$

Em particular,

$$(d^{-1}Xd)_{ii} = X_{ii}.$$

(ii) Se $u \in U'$ e $l \in L'$, então

$$\text{diag}(u^{-1}du) = \text{diag}(l^{-1}dl) = \text{diag}(d).$$

(iii) Toda matriz em L se decompõe como produto de uma matriz diagonal por uma matriz em L' . Analogamente, toda matriz em U se decompõe como produto de uma matriz diagonal por uma matriz em U' .

Em particular, matrizes diagonais são fixadas por conjugação por matrizes triangulares invertíveis.

Demonstração. (i) Imediato.

(ii) Se $u \in U'$ então $du = (d_{ii}u_{ij})$ pertence a U com diagonal igual a diagonal de d . Como U' é grupo, temos u^{-1} triangular superior com 1 nas entradas da diagonal, e portanto

$$\text{diag}(u^{-1}du) = \text{diag}(du) = \text{diag}(d).$$

Analogamente se mostra que se $l \in L'$, então

$$\text{diag}(l^{-1}dl) = \text{diag}(d).$$

(iii) Cada $l \in L$ pode ser escrita como o produto $l = l'd$ ou $l = dl'$, com $l' \in L'$ e d diagonal não necessariamente as mesmas nas duas decomposições. Com efeito, se as entradas da diagonal de l são $\lambda_1, \dots, \lambda_n$ e seus vetores colunas são c_1, \dots, c_n , então

$$l = (c_1 | \dots | c_n) = \underbrace{(\lambda_1^{-1}c_1 | \dots | \lambda_n^{-1}c_n)}_{\in L'} \underbrace{(\lambda_1 e_1 | \dots | \lambda_n e_n)}_{\text{diagonal}},$$

onde e_1, \dots, e_n são os vetores canônicos de \mathbb{C}^n . O mesmo processo aplicado em l^t nos dá $l = dl'$.

Analogamente, temos que cada $u \in U$ pode ser escrito como o produto $u = du'$ ou $u = u'd$, com $u' \in U'$ e d diagonal. □

Observação A.5. (1) Seja $d = \text{diag}(\lambda_1, \dots, \lambda_n)$ com $\lambda_i \neq 0$ para todo $i = 1, \dots, n$.

Então, para todo t temos que a ij -ésima entrada da matriz $d^t X d^{-t}$ é $\left(\frac{\lambda_i}{\lambda_j}\right)^t X_{ij}$.

Em particular, a diagonal de $d^t X d^{-t}$ é a diagonal de X .

(2) Se $u \in U$, podemos ainda escrever $u = sr$, onde s é diagonal com entradas de módulo 1 e r é triangular superior com diagonal positiva. De fato, basta escrever $u = du'$ com d diagonal e $u' \in U'$ (ver Lema A.4) e fazer a decomposição polar das entradas de d , pondo $d = sd'$ onde s é diagonal com entradas de módulo 1 e d' diagonal com entradas reais positiva, tomando $r = d'u'$ triangular superior com diagonal positiva.

A.3 Decomposição de Gauss (PLU)

Esta decomposição é o correspondente matricial do pivoteamento Gaussiano com permutações. Sejam $L = \{(g_{ij}) \in G : g_{ij} = 0, \forall i < j\}$ e $U = \{(g_{ij}) \in G : g_{ij} = 0, \forall i < j\}$ os conjuntos das matrizes invertíveis triangulares inferiores e superiores, respectivamente. Como os elementos de L e U são invertíveis e o determinante de uma matriz triangular é o produto de suas entradas na diagonal, segue que os elementos de L e U não contém zeros na diagonal. Além disso,

Proposição A.6. L e U são subgrupos de G .

Demonstração. Seja $g \in G$ e escreva $\text{diag}(g) = (\alpha_1, \dots, \alpha_n)$. Então $g \in U$ se, e somente se, sua k^{a} coluna é

$$ge_k = *e_1 + \dots + *e_{k-1} + \alpha_k e_k, \quad \text{e } \alpha_k \neq 0 \quad (\text{A.2})$$

para todo $k = 1, \dots, n$, onde os asteriscos são escalares. Assim,

$$g[e_1, \dots, e_k] = [e_1, \dots, e_k] \quad \text{para todo } k, \quad (\text{A.3})$$

em que $[e_1, \dots, e_k]$ é o subespaço gerado pelos vetores e_1, \dots, e_k . Reciprocamente, se $g \in G$ satisfaz (A.3), então de $g[e_1] = [e_1]$ temos que $ge_1 = \alpha_1 e_1$ com $\alpha_1 \neq 0$. Supondo que (A.2) vale para $k = 1, \dots, n-1$, de $g[e_1, \dots, e_{k+1}] = [e_1, \dots, e_{k+1}]$ e $g[e_1, \dots, e_k] = [e_1, \dots, e_k]$, temos que (A.2) vale para $k+1$. Assim, $g \in U$ se e somente se satisfaz (A.3). Portanto, se $u_1, u_2 \in U$, ou seja, satisfazem (A.3), é fácil ver que $u_1^{-1}u_2$ também satisfazem, ou seja, $u_1^{-1}u_2 \in U$. O que significa ser U um subgrupo de G .

Para vermos que L é subgrupo de G , basta notarmos que $L = U^\dagger = \{u^\dagger : u \in U\}$. De fato, tomando $l_1 = u_1^\dagger$ e $l_2 = u_2^\dagger$ em L , com $u_1, u_2 \in U$, temos que

$$l_1^{-1}l_2 = (u_1^\dagger)^{-1}u_2^\dagger = (u_1^{-1})^\dagger u_2^\dagger = (u_2 u_1^{-1})^\dagger$$

pertence à L , uma vez que $u_2 u_1^{-1} \in U$. □

A decomposição PLU para G que trataremos aqui, via operações elementares sobre as colunas de uma dada $g \in G$ (ver Seção A.1), não é mais que a eliminação Gaussiana com pivoteamento.

Proposição A.7. Se $g \in G$, então existem $p \in P$, $l \in L$ e $u \in U$ tais que $g = plu$.

Demonstração. Seja i_1 o número da primeira linha não nula da coluna 1 de g , que existe por g ser invertível. Fazendo $L_1 \leftrightarrow L_{i_1}$, isto é, transpondo as linhas 1 e i_1 de g , existe $p_1 \in P$ que determina esta transposição quando multiplicado à esquerda de g . Então, o elemento (11) de $p_1 g$ é não nulo. Daí existe $u_1 \in U$ (ver Seção A.1) que, multiplicada à direita de $p_1 g$, fixa a primeira coluna e anula as entradas restantes da primeira linha. Ver Figura A.1 abaixo.

Agora seja i_2 o número da primeira linha não nula da coluna 2 de $p_1 g u_1$, que existe por $p_1 g u_1$ ser invertível. Fazendo $L_2 \leftrightarrow L_{i_2}$, isto é, transpondo as linhas 2 e i_2 de g obtemos $p_2 \in P$ tal que o elemento (22) de $p_2 p_1 g u_1$ é não nulo. Daí existe $u_2 \in U$ que,

multiplicada à direita de $p_2 p_1 g u_1$, fixa as colunas 1 e 2 e anula as entradas restantes da segunda linha. Como está ilustrado na Figura A.1.

Repetindo o processo acima para as colunas subsequentes, até a n^a , obtemos $p' = p_{n-1} \cdots p_2 p_1 \in P$ e $u' = u_1 u_2 \cdots u_{n-1} \in U$ tais que $p' g u' = l \in L$. Portanto existem $p = (p')^{-1} \in P$, $l \in L$ e $u = (u')^{-1} \in U$ tais que

$$g = plu.$$

□

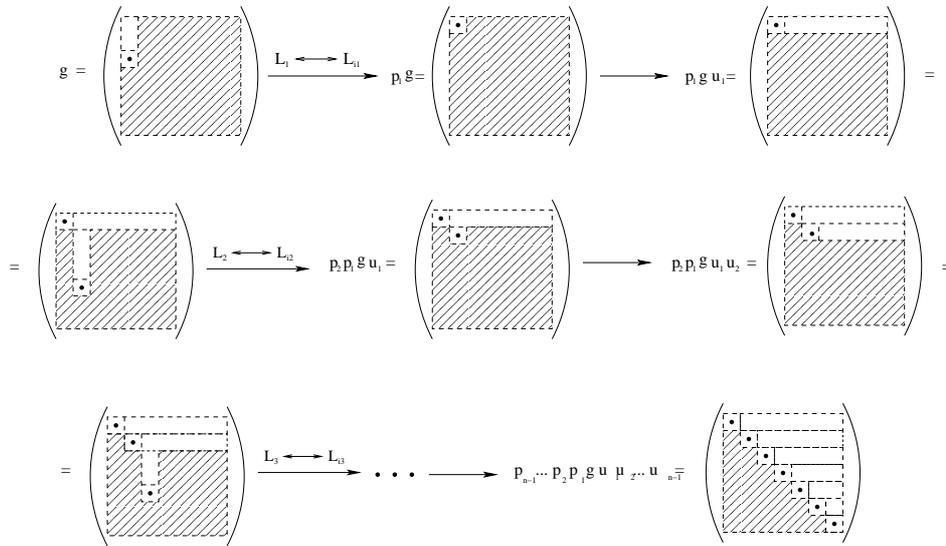


Figura A.1: Eliminação Gaussiana com pivoteamento.

Observação A.8. Na decomposição PLU de G tanto os fatores triangulares quanto a permutação não são unicamente determinados pela matriz em G , como mostra o seguinte exemplo.

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 3 & 0 \\ 2 & 4 & 0 \end{pmatrix} = \overbrace{\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}^p \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^l \overbrace{\begin{pmatrix} 1 & 3 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^u$$

$$= \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{p'} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{l'} \underbrace{\begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{u'}.$$

Isto ocorre porque as matrizes triangulares dependem das linhas transpostas, de acordo com a demonstração da proposição anterior. As transposições por sua vez determinam uma permutação, mas não necessariamente de maneira única como está claro no exemplo acima.

A.4 Decomposição de Iwasawa (QR)

Esta decomposição é o correspondente matricial do processo de Gram-Schmidt. Considere

$$Q = \{g \in G : g^*g = I\}$$

o subgrupo de $G = \text{Gl}(n, \mathbb{C})$ constituído pelas matrizes unitárias, ou seja, com colunas ortonormais, e

$$R = \{g \in U : g_{ii} > 0, \forall i\}$$

o subgrupo de G constituído pelas matrizes triangulares superiores com entradas positivas na diagonal principal.

Nesta seção mostraremos que $G = QR$, isto é, todo elemento $g \in G$ se decompõe como um produto $g = qr$, com $q \in Q$ e $r \in R$. Além disso, como $Q \cap R = \{I\}$ temos que esta decomposição é única. De fato, se $g \in Q \cap R$ então $g \in R$ é triangular superior, idem para $\bar{g}^\dagger = g^* = g^{-1} \in R$. Assim, $g = \text{diag}(\lambda_1, \dots, \lambda_n)$, com $\lambda_i > 0$ e $|\lambda_i| = 1$, para todo i , ou seja, $g = I$.

Para a decomposição $G = QR$, relembremos o processo de ortogonalização de Gram-Schmidt (ver Seção 8.2 de [?]). Seja $\{g_1, \dots, g_n\}$ um conjunto de vetores linearmente independentes em \mathbb{C}^n . Então existe uma base ortogonal $\{w_1, \dots, w_n\}$ de \mathbb{C}^n tal que

$$\text{ger}\{g_1, \dots, g_k\} = \text{ger}\{w_1, \dots, w_k\}, \quad k = 1, \dots, n.$$

Os vetores w_1, \dots, w_n são obtidos indutivamente da seguinte maneira. Ponha $w_1 = g_1$ e suponhamos que w_1, \dots, w_m ($1 \leq m < n$) tenham sido escolhidos de modo que para cada k entre 1 e m , $\{w_1, \dots, w_k\}$ seja uma base ortogonal do subespaço $\text{ger}\{g_1, \dots, g_k\}$. E então ponha

$$w_{m+1} = g_{m+1} - \sum_{k=1}^m \frac{(g_{m+1}, w_k)}{|w_k|^2} w_k,$$

onde (\cdot, \cdot) é o produto interno canônico em \mathbb{C}^n . Note que os vetores w_1, \dots, w_n obtidos acima são ortogonais, mas não necessariamente ortonormais. Este processo pode ser realizado via operações elementares sobre as colunas de $g = (g_1 | \dots | g_n) \in G$ substituindo cada coluna de g por combinações lineares adequadas com as colunas anteriores, ou seja, multiplicando g à direita por matrizes triangulares superiores com 1 na diagonal de modo a obter

$$gr' = (w_1 | \dots | w_n),$$

onde $r' \in R$ tem 1 nas entradas da diagonal e w_1, \dots, w_n são os vetores ortogonais

descritos acima. Mais precisamente, temos

$$g_1 \longrightarrow w_1 \qquad : \quad g \longrightarrow g$$

$$g_2 \longrightarrow w_2 = g_2 - \frac{(g_2, w_1)}{|w_1|^2} w_1 \qquad : \quad g \longrightarrow gr_1,$$

$$\text{onde } r_1 = \left(\cdots \left| e_2 - \frac{(g_2, w_1)}{|w_1|^2} e_1 \right|_2 \cdots \right) \in R$$

...

$$g_{m+1} \longrightarrow w_{m+1} = g_{m+1} - \sum_{k=1}^m \frac{(g_{m+1}, w_k)}{|w_k|^2} w_k \qquad : \quad g \longrightarrow gr_1 \cdots r_{m-1} r_m,$$

$$\text{onde } r_m = \left(\cdots \left| e_{m+1} - \sum_{k=1}^m \frac{(g_{m+1}, w_k)}{|w_k|^2} e_k \right|_{m+1} \cdots \right) \in R.$$

Observe que cada r_i é triangular superior com 1 na diagonal. Ao final, pondo $r' = r_1 \cdots r_{n-1}$, obtemos $gr' = (w_1 | \cdots | w_n)$. Logo,

$$g = (w_1 | \cdots | w_n) (r')^{-1} = \left(\frac{w_1}{|w_1|} \left| \cdots \right| \frac{w_n}{|w_n|} \right) (|w_1| e_1 | \cdots | |w_n| e_n) (r')^{-1}.$$

Pondo $q = \left(\frac{w_1}{|w_1|} \left| \cdots \right| \frac{w_n}{|w_n|} \right) \in Q$ e $r = (|w_1| e_1 | \cdots | |w_n| e_n) (r')^{-1} \in R$, temos a decomposição $g = qr$ desejada, conhecida como *decomposição de Iwasawa* ou *decomposição QR* de G . Mostramos o seguinte

Teorema A.9. Uma matriz $g \in G$ se decompõe unicamente como o produto $g = qr$, com $q \in Q$ e $r \in R$.

Pela unicidade da decomposição QR de G , estão bem definidas as aplicações

$$\begin{aligned} Q : G &\longrightarrow Q & \text{ e } & R : G &\longrightarrow R \\ g &\longmapsto q & & g &\longmapsto r. \end{aligned}$$

que, por um abuso de notação, denotamos pelas respectivas letras dos subgrupos Q e R de G . Pela construção da decomposição QR para o teorema acima, uma vez que são diferenciáveis o produto interno em \mathbb{C}^n e a norma em $\mathbb{C}^n \setminus \{0\}$, obtemos o seguinte

Corolário A.10. As aplicações Q e R são diferenciáveis.

Temos ainda da unicidade da decomposição acima que a aplicação Q satisfaz: $Q \circ Q = Q$, é Q -equivariante à esquerda e R -invariante à direita, ou seja,

$$Q(qgr) = qQ(g), \quad q \in Q, g \in G \text{ e } r \in R.$$

Em particular, temos que

$$Q(gQ(h)) = Q(gh). \tag{A.4}$$

A.5 Forma de Schur

Dada uma base ordenada $b \in \text{Gl}(n, \mathbb{C})$ e um operador linear $X \in \mathfrak{gl}(n, \mathbb{C})$, a matriz de X na base b é dada por

$$b^{-1}Xb,$$

onde os vetores da base são as colunas de b . Uma base b na qual X é triangular superior será chamada de *base de Schur* para X e a matriz triangular superior $b^{-1}Xb$, uma *forma de Schur* para X .

Teorema A.11. Dada $X \in \mathfrak{gl}(n, \mathbb{C})$, existe uma base de Schur $b \in \text{Gl}(n, \mathbb{C})$. Em particular,

- (i) X possui ao menos um autovetor.
- (ii) A diagonal de uma forma de Schur para X é composta pelos autovalores de X .
- (iii) Uma base de Schur b para X pode ser ortogonalizada. Mais precisamente, aplicando o processo de Gram-Schmidt às colunas de b , obtemos uma base ortogonal $q \in Q$ que ainda é base de Schur para X .
- (iv) Se X é hermitiana, então X é unitariamente conjugada a uma matriz diagonal, ou seja, $q^{-1}Xq$ é diagonal para algum $q \in Q$.

Demonstração. Como \mathbb{C} é algebricamente fechado e $p(\lambda) = \det(X - \lambda I)$ é um polinômio em λ , existe $\lambda_1 \in \mathbb{C}$ tal que $\det(X - \lambda_1 I) = 0$. Assim obtemos $v_1 \in \mathbb{C}^n$ com $v_1 \neq 0$ no núcleo de $X - \lambda_1 I$, isto é, tal que

$$Xv_1 = \lambda_1 v_1. \tag{A.5}$$

Então, $X[v_1] \subseteq [v_1]$ e X induz um operador linear no espaço vetorial $\mathbb{C}^n/[v_1] = \{v + [v_1] : v \in \mathbb{C}^n\}$ definido por

$$\begin{aligned} \mathbb{C}^n/[v_1] &\longrightarrow \mathbb{C}^n/[v_1] \\ v + [v_1] &\longmapsto Xv + [v_1]. \end{aligned}$$

Por um argumento análogo ao que fizemos acima, existem $v_2 + [v_1] \in \mathbb{C}^n/[v_1]$ com $v_2 \notin [v_1]$ e $\lambda_2 \in \mathbb{C}$ tal que

$$Xv_2 + [v_1] = \lambda_2(v_2 + [v_1]).$$

Assim obtemos $v_2 \in \mathbb{C}^n$ linearmente independente de v_1 , uma vez que $v_2 \notin [v_1]$, tal que $Xv_2 \in \lambda_2 v_2 + [v_1]$. Em particular temos que $X[v_1, v_2] \subseteq [v_1, v_2]$. Repetindo este processo sucessivamente, obtemos uma base $\beta = \{v_1, v_2, \dots, v_n\}$ de \mathbb{C}^n tal que, para todo $k = 1, \dots, n$ temos

$$X[v_1, \dots, v_k] \subseteq [v_1, \dots, v_k].$$

Pondo b a matriz invertível cujas colunas são os vetores da base ordenada β , segue que $b^{-1}Xb$ é triangular superior, como queríamos.

- (i) Segue de (A.5) que v_1 é um autovetor de X , visto que $v_1 \neq 0$.

- (ii) Se $X \in \mathfrak{g}$ é triangular, seus autovalores são as entradas da diagonal com multiplicidade igual a quantidade de vezes que aparece, tendo em vista que $X - \lambda I$ permanece triangular e, portanto, $\det(b^{-1}Xb - \lambda I) = \det(X - \lambda I)$ é o produto dos elementos de sua diagonal.
- (iii) Faça a decomposição QR de $b = qr$, com q unitária e r triangular superior (ver Seção A.4). Então a base ortogonal dada pelas colunas de q são o resultado do processo de Gram-Schmidt aplicado aos vetores de β . Temos que q é uma base de Schur para X . De fato, de $q = br^{-1}$ segue que

$$q^{-1}Xq = r(b^{-1}Xb)r^{-1},$$

é triangular superior, uma vez que r e $(b^{-1}Xb)$ são triangulares superiores e as matrizes triangulares superiores são fechadas para composição.

- (iv) Sendo X hermitiana e q uma base ortogonal, $q^{-1}Xq$ também é hermitiana (ver Proposição A.13). Temos então que a forma de Schur $q^{-1}Xq$ é hermitiana e triangular superior e, portanto, diagonal.

□

A.6 Algumas funções matriciais

Dado $X \in \mathfrak{g} = \mathfrak{gl}(n, \mathbb{C})$, considere $X^* = \overline{X}^\dagger$ o transposto conjugado de X . Dizemos que X é *hermitiana* (respectivamente *anti-hermitiana*) se $X^* = X$ (respectivamente $X^* = -X$).

Proposição A.12. Valem as seguintes propriedades.

- (1) Se X é hermitiana, então sua diagonal tem entradas reais.
- (2) Se X é anti-hermitiana, então sua diagonal tem entradas imaginárias.
- (3) $(X^*)^* = X$
- (4) $(tX)^* = tX^*$, $t \in \mathbb{R}$
- (5) $(X + Y)^* = X^* + Y^*$
- (6) $(XY)^* = Y^*X^*$

Para demonstrações das propriedades acima, ver [9] páginas 15 e 178. Considere o conjunto Q das *matrizes unitárias*.

Proposição A.13. Se $X \in \mathfrak{g}$ é hermitiana e $q \in Q$, então $q^{-1}Xq$ é hermitiana

Demonstração. De $q \in Q$, temos $q^*q = I$, de onde temos

$$q^* = q^{-1} \quad \text{e} \quad (q^{-1})^* = q^{**} = q.$$

Assim, como $X^* = X$, segue que

$$(q^{-1}Xq)^* = q^*X^*(q^{-1})^* = q^{-1}Xq.$$

Portanto, $q^{-1}Xq$ é hermitiana. □

Para todo $X \in \mathfrak{g}$, definimos a *exponencial de X* pela série convergente

$$\exp(X) = \sum_{k=0}^{\infty} \frac{X^k}{k!}.$$

Proposição A.14. A imagem de \exp é todo o grupo $\text{Gl}(n, \mathbb{C})$. Dados $X, Y \in \mathfrak{g}$ e $g \in G$, valem as seguintes propriedades.

- (i) $g^{-1}\exp(X)g = \exp(g^{-1}Xg)$
- (ii) $\exp(X)^* = \exp(X^*)$
- (iii) Se $[X, Y] = 0$, então $\exp(X + Y) = \exp(X)\exp(Y)$. Em particular, $\exp((t + s)X) = \exp(tX)\exp(sX)$
- (iv) $\exp(X)^{-1} = \exp(-X)$

Demonstração. A primeira afirmação vem do fato de $\exp : \mathfrak{gl}(n, \mathbb{C}) \rightarrow \text{Gl}(n, \mathbb{C})$ ser sobrejetora para matrizes complexas (ver [9]). Para as outras afirmação, vamos apenas dar as idéias (para mais detalhes consulte [9]).

(1) Para cada $k \in \mathbb{N}$, temos que

$$(g^{-1}Xg)^k = g^{-1}X^k g.$$

Portanto,

$$\begin{aligned} g^{-1}\exp(X)g &= g^{-1} \left(\sum_{k=0}^{\infty} \frac{X^k}{k!} \right) g \\ &= \sum_{k=0}^{\infty} g^{-1} \frac{X^k}{k!} g \\ &= \sum_{k=0}^{\infty} \frac{(g^{-1}Xg)^k}{k!} \\ &= \exp(g^{-1}Xg). \end{aligned}$$

(2) Note que

$$\sum_{k=0}^N \frac{(X^\dagger)^k}{k!} = \left(\sum_{k=0}^N \frac{X^k}{k!} \right)^\dagger, \quad \forall N \in \mathbb{N}.$$

Portanto, como a aplicação $X \mapsto X^\dagger$ é contínua, segue que

$$\exp(X^\dagger) = \sum_{k=0}^{\infty} \frac{(X^\dagger)^k}{k!} = \left(\sum_{k=0}^{\infty} \frac{X^k}{k!} \right)^\dagger = \exp(X)^\dagger.$$

Assim,

$$(\exp(X))^* = \overline{\exp(X)}^\dagger = \exp(\overline{X})^\dagger = \exp(\overline{X}^\dagger) = \exp(X^*)$$

(3) Se $[X, Y] = 0$, então pela regra do binômio de Newton, vale

$$(X + Y)^k = \sum_{m=0}^k \binom{k}{m} X^m Y^{k-m} = \sum_{m=0}^k \frac{k!}{(k-m)!m!} X^m Y^{k-m}$$

e daí,

$$\begin{aligned} \exp(X + Y) &= \sum_{k=0}^{\infty} \frac{(X + Y)^k}{k!} \\ &= \sum_{k=0}^{\infty} \sum_{m=0}^k \frac{1}{(k-m)!m!} X^m Y^{k-m} \\ &= \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} \frac{1}{(k-m)!m!} X^m Y^{k-m} \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} X^m \left(\sum_{k=m}^{\infty} \frac{1}{(k-m)!} Y^{k-m} \right) \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} X^m \sum_{l=0}^{\infty} \frac{1}{l!} Y^l \quad (l := k - m) \\ &= \exp(X)\exp(Y). \end{aligned}$$

(4) Basta tomar $t = 1$ e $s = -1$ em (3). □

Proposição A.15. Seja $X \in \mathfrak{g}$. Então a curva $\alpha : \mathbb{R} \rightarrow \mathfrak{g}$ definida por

$$\alpha(t) = \exp(tX), \quad t \in \mathbb{R},$$

satisfaz o PVI

$$\begin{cases} \alpha'(t) = X\alpha(t) \\ \alpha(0) = 0. \end{cases}$$

Demonstração. Ora,

$$\alpha(t) = \exp(tX) = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!} = \sum_{k=0}^{\infty} \frac{X^k}{k!} t^k.$$

Como podemos derivar essa série termo a termo (ver [9]), temos que

$$\begin{aligned} \alpha'(t) &= \frac{d}{dt} \left(\sum_{k=0}^{\infty} \frac{X^k}{k!} t^k \right) = \frac{d}{dt} \left(I - \sum_{k=1}^{\infty} \frac{X^k}{k!} t^k \right) \\ &= \sum_{k=1}^{\infty} \frac{d}{dt} \left(\frac{X^k}{k!} t^k \right) = \sum_{k=1}^{\infty} \frac{X^k}{k!} k t^{k-1} \\ &= \sum_{k=1}^{\infty} \frac{X^k}{(k-1)!} t^{k-1} = X \sum_{k=0}^{\infty} \frac{X^k}{k!} t^k \\ &= X\alpha(t). \end{aligned}$$

□

Proposição A.16. Se $B : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ é bilinear, então B é diferenciável com

$$B'(X, Y)(Z, W) = B(X, W) + B(Z, Y).$$

Demonstração. Temos que

$$B((X, Y) + (Z, W)) = B(X, Y) + B(X, W) + B(Z, Y) + B(Z, W)$$

e que $\frac{B(Z, W)}{|(Z, W)|} \rightarrow 0$ quando $(Z, W) \rightarrow 0$, onde tomamos em $\mathfrak{g} \times \mathfrak{g}$ qualquer norma $|\cdot|$.

Portanto, B tem derivada

$$B'(X, Y)(Z, W) = B(X, W) + B(Z, Y),$$

como queríamos.

□

Apêndice B

Topologia

Este apêndice tem por objetivo o estudo de algumas ferramentas de topologia que aparecem ao longo da dissertação, tais como espaços homogêneos e fluxos.

B.1 Espaços Topológicos

Seja X um conjunto não vazio. Uma *topologia* em X é uma família τ de subconjuntos de X que satisfaz as seguintes propriedades.

- (1) \emptyset e X pertencem à τ .
- (2) τ é fechado para união arbitrária, ou seja,

$$\{A_\lambda\}_\lambda \subset \tau \implies \bigcup_\lambda A_\lambda \in \tau.$$

- (3) τ é fechado para intersecção finita, ou seja,

$$A_1, \dots, A_n \in \tau \implies \bigcap_{i=1}^n A_i \in \tau.$$

O par (X, τ) é chamado *espaço topológico*, os elementos de τ são chamados *abertos* e seus complementares são chamados *fechados*. Quando não houver confusão sobre a topologia escreveremos simplesmente X no lugar de (X, τ) .

Sejam (X, τ) um espaço topológico e $Y \subset X$ um subconjunto não vazio. Então

$$\tau_Y = \{A \cap Y : A \subset X\}$$

é uma topologia em Y , a qual chamamos de *topologia induzida* por τ , ou por X . Neste caso, dizemos que Y é um *subespaço* de X . É fácil ver que se Y é um subespaço de X e F é fechado em X , então $F \cap Y$ é fechado em Y na topologia induzida por X .

Observe que a topologia induzida é a menor topologia de Y na qual a inclusão $i : Y \rightarrow X$, $y \mapsto y$, é contínua.

Se (X, τ_X) e (Z, τ_Z) são espaços topológicos, podemos munir o produto cartesiano $X \times Z = \{(x, z) : x \in X \text{ e } z \in Z\}$ com a topologia

$$\tau = \{A \times B : A \in \tau_X \text{ e } B \in \tau_Z\},$$

a qual chamamos de *topologia produto* em $X \times Z$.

Definição B.1. (a) Uma *vizinhança* de um ponto $x \in X$ é um subconjunto V de X tal que x está em algum aberto contido em V . V é dita *vizinhança aberta* se é uma vizinhança de cada um de seus pontos.

(b) X é *espaço de Hausdorff*, ou apenas *Hausdorff*, se para quaisquer pontos $x, y \in X$, com $x \neq y$, existem vizinhanças V e W de x e y , respectivamente, tais que $V \cap W = \emptyset$. Dizemos, neste caso, que X é *separável* por abertos.

(c) Uma sequência $\{x_k\}_{k \in \mathbb{N}}$ em um espaço topológico X *converge* para $x \in X$, e escrevemos

$$\lim_{k \rightarrow \infty} x_k = x \quad \text{ou} \quad x_k \longrightarrow x,$$

se para toda vizinhança V de x temos $x_k \in V$ para k grande o suficiente.

Temos que se Y é um subespaço de um espaço de Hausdorff, então Y é Hausdorff ([12], p. 100).

Proposição B.2. Se X é Hausdorff, então o limite de sequências convergentes em X é único.

Demonstração. Se $x_k \rightarrow x$ e $x_k \rightarrow y$ com $x \neq y$, então tomando em X vizinhanças disjuntas V e W de x e de y , respectivamente, temos que $x_k \in V \cap W$ para k grande o suficiente, o que é absurdo. \square

B.2 Aplicações contínuas

Sejam X e Y espaços topológicos. Uma aplicação $f : X \rightarrow Y$ é dita *contínua* se a imagem inversa $f^{-1}(A) = \{x \in X : f(x) \in A\}$ é um conjunto aberto em X para todo A aberto em Y . Dizemos ainda que f é contínua num dado $x \in X$ se $f^{-1}(V)$ é uma vizinhança de x em X para toda vizinhança V de $f(x)$ em Y . Assim, $f : X \rightarrow Y$ é contínua se e somente se, é contínua em todo $x \in X$ (ver [12], p. 104).

Proposição B.3. Se a aplicação $f : X \rightarrow Y$ é contínua e sobrejetora, então a imagem por f de um conjunto denso em X é um conjunto denso em Y .

Demonstração. Seja $D \subset X$ um conjunto denso. Dado um aberto não vazio $A \subset Y$, afirmamos que $A \cap f(D) \neq \emptyset$. De fato, temos que $f^{-1}(A)$ é aberto e não vazio de X , pois f é contínua e sobrejetora. Daí, $f^{-1}(A) \cap D \neq \emptyset$, já que D é denso em X . Novamente usando a hipótese de f ser sobrejetora, temos que $f(f^{-1}(A)) = A$. Assim, tomando $x \in f^{-1}(A) \cap D$, temos que $f(x) \in A \cap f(D)$. Logo, $A \cap f(D) \neq \emptyset$, como afirmamos. \square

Dizemos que uma aplicação $f : X \rightarrow Y$, entre os espaços topológicos X e Y , é um *homeomorfismo* quando f é uma bijeção contínua com inversa contínua.

Proposição B.4 ([12], p. 167). Se X é compacto e Y é Hausdorff, então uma bijeção contínua $X \rightarrow Y$ é um homeomorfismo.

Sejam $X = (X, \tau_X)$ um espaço topológico, Y um conjunto e $\pi : X \rightarrow Y$ uma aplicação sobrejetora. Definimos em Y a *topologia quociente* dada por

$$\tau_\pi = \{B \subset Y : \pi^{-1}(B) \in \tau_X\}.$$

Assim, temos que π é contínua. Mais ainda, temos que essa é a maior topologia de Y na qual π é contínua pois, se τ_Y é uma topologia em Y para a qual $\pi : (X, \tau_X) \rightarrow (Y, \tau_Y)$ é contínua, então claramente $\tau_Y \subset \tau_\pi$.

Proposição B.5. Se $\pi : (X, \tau_X) \rightarrow (Y, \tau_\pi)$ é aberta na topologia quociente de Y , então essa é a única topologia na qual π é contínua e aberta. Neste caso, π possui a seguinte propriedade de levantamento. Uma aplicação $f : Y \rightarrow Z$ num espaço topológico Z é contínua se, e somente se, seu levantamento $f \circ \pi : X \rightarrow Z$ é uma aplicação contínua.

$$\begin{array}{ccc} X & & \\ \pi \downarrow & \searrow f \circ \pi & \\ Y & \xrightarrow{f} & Z \end{array}$$

Demonstração. Seja τ_Y uma topologia em Y na qual π é contínua e aberta. Tome $B \in \tau_\pi$. Então $\pi^{-1}(B) \in \tau_X$ e, como $\pi : X \rightarrow (Y, \tau_Y)$ é aberta e sobrejetora, $B = \pi(\pi^{-1}(B)) \in \tau_Y$. Isso mostra que $\tau_\pi \subset \tau_Y$. Como $\pi : X \rightarrow (Y, \tau_Y)$ é contínua, temos que $\tau_Y \subset \tau_\pi$, o que mostra a igualdade.

Para a segunda parte, se $f : Y \rightarrow Z$ é contínua, segue que o levantamento $f \circ \pi : X \rightarrow Z$ também o é, uma vez que $\pi : X \rightarrow Y$ é contínua. Reciprocamente, se o levantamento $f \circ \pi : X \rightarrow Z$ é contínuo, então dado $A \subset Z$ aberto, temos que $f^{-1}(A) = \pi((f \circ \pi)^{-1}(A))$ é aberto em Y , uma vez que π é aberta e $(f \circ \pi)^{-1}(A)$ é aberto em X . \square

Observação B.6. Se $\pi : (X, \tau_X) \rightarrow (Y, \tau_\pi)$ é aberta então a topologia quociente é gerada por $\{\pi(A) : A \in \tau_X\}$.

B.3 Topologia dos Espaços Homogêneos

Um grupo abstrato G é *topológico* se G é espaço topológico tal que a multiplicação $(g, g') \in G \times G \mapsto gg' \in G$ e a inversão $g \in G \mapsto g^{-1} \in G$ são contínuas.

Proposição B.7. Seja G um grupo topológico. Se $V \subset G$ é aberto em G , então

$$VX = \{vx : v \in V \text{ e } x \in X\} \quad \text{e} \quad XV = \{xv : x \in X \text{ e } v \in V\}$$

são abertos em G para todo $X \subset G$

Demonstração. Dado $x \in G$, a aplicação $g \in G \mapsto gx \in G$ é um homeomorfismo, com inversa $g \in G \mapsto gx^{-1} \in G$. Então, se V é aberto em G , temos que Vx é aberto em G . Portanto,

$$VX = \bigcup_{x \in X} Vx$$

é aberto em G . De maneira análoga mostramos que XV é aberto em G . \square

Proposição B.8. Seja A uma vizinhança de $g \in G$. Então, existe uma vizinhança V de 1 em G tal que Vg (ou gV) é uma vizinhança de g contida em A .

Demonstração. Temos que a aplicação $g' \in G \mapsto g'g^{-1} \in G$ é um homeomorfismo. Logo, se A é uma vizinhança de g em G , então $V = Ag^{-1}$ é uma vizinhança de 1 em G com Vg contida em A . \square

Sejam G um grupo e X um conjunto. Uma *ação à esquerda* de G em X é uma aplicação

$$m : G \times X \longrightarrow X, \quad (g, x) \longmapsto g \cdot x,$$

tal que

$$\begin{cases} 1 \cdot x & = x \\ g \cdot (g' \cdot x) & = (gg') \cdot x, \end{cases}$$

para todos $g, g' \in G$ e todo $x \in X$. Escreveremos a justaposição gx no lugar de $g \cdot x$ quando a ação em questão estiver clara no contexto.

Se G é grupo topológico e X é um espaço topológico, a ação de G em X é *contínua* se a aplicação m é contínua, onde em $G \times X$ consideramos a topologia produto. De agora em diante vamos considerar apenas ações contínuas. A partir da ação de G em X definimos

- ação de G em um *ponto base* $x_0 \in X$, $m_{x_0} : G \rightarrow X$, por $g \mapsto gx_0$,
- a ação de um dado $g \in G$ em X , $g : X \rightarrow X$, por $x \mapsto gx$.

Segue das propriedades da ação que $g : X \rightarrow X$ é um homeomorfismo de X com inversa $g^{-1} : X \rightarrow X$.

Exemplo B.9. G age em si via multiplicação, $(g, g') \in G \times G \mapsto gg' \in G$. Dado $g \in G$, as aplicações $E_g : g' \in G \mapsto gg' \in G$ e $D_g : g' \in G \mapsto g'g \in G$ são homeomorfismos com inversas $E_g^{-1} = E_{g^{-1}}$, $D_g^{-1} = D_{g^{-1}}$. E_g dá a ação à esquerda de $g \in G$ no espaço G e D_g dá a ação do grupo G no ponto base $g \in G$.

Considere que G age nos espaços topológicos X e Y . Uma aplicação $f : X \rightarrow Y$ é *G -equivariante* se $f \circ g = g \circ f$, $\forall g \in G$, ou seja, $gf(x) = f(gx)$, $\forall g \in G$, $\forall x \in X$. Dizemos que f é *G -invariante* se $f \circ g = f$, $\forall g \in G$, isto é, $f(gx) = f(x)$, $\forall g \in G$, $\forall x \in X$.

Observação B.10. $f : X \rightarrow Y$ é G -equivariante quando o diagrama

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ g \downarrow & & \downarrow g \\ X & \xrightarrow{f} & Y \end{array}$$

comuta, onde $g : X \rightarrow X$ e $g : Y \rightarrow Y$ são ações de $g \in G$ em X e Y , respectivamente.

Se H é um subgrupo de G , temos o quociente

$$G/H = \{g'H : g' \in G\}$$

das classes laterais à esquerda de H em G e a projeção natural $\pi : G \rightarrow G/H$, $g' \mapsto g'H$. Em virtude da sobrejetividade de π , munimos G/H com a topologia quociente e $G \times G/H$ com a topologia produto.

Proposição B.11. A projeção $\pi : G \rightarrow G/H$ e a ação sobrejetora

$$\begin{aligned} \bar{m} : G \times G/H &\longrightarrow G/H \\ (g, g'H) &\longmapsto g(g'H) = gg'H. \end{aligned}$$

são contínuas e abertas.

Demonstração. Para provar que π é aberta, tome um aberto $A \subset G$. Então $\pi(A)$ é aberto na topologia quociente se, e somente se, $\pi^{-1}(\pi(A))$ é aberto em G . Temos que

$$\pi^{-1}(\pi(A)) = AH = \bigcup_{h \in H} Ah$$

é aberto, pois é a união dos abertos Ah .

Seja $\varphi = id \times \pi : G \times G \rightarrow G \times G/H$, $(g, g') \mapsto (g, g'H)$. Note que φ é aplicação sobrejetora, contínua e aberta. Portanto, $G \times G/H$ tem a topologia quociente induzida por φ . Assim, o seguinte diagrama é comutativo

$$\begin{array}{ccc} G & \xrightarrow{\pi} & G/H \\ m \uparrow & & \uparrow \bar{m} \\ G \times G & \xrightarrow{\varphi} & G \times G/H \end{array}$$

ou seja, $\bar{m} \circ \varphi = \pi \circ m$, e \bar{m} é contínua se, e somente se, seu levantamento $\pi \circ m$ para $G \times G$ é contínua. Logo, se $U \subset G/H$ é aberto, como

$$\varphi^{-1} \circ \bar{m}^{-1} = (\bar{m} \circ \varphi)^{-1} = (\pi \circ m)^{-1},$$

temos que $\bar{m}^{-1}(U) = \varphi \circ (\pi \circ m)^{-1}(U)$ é aberto. Como m e π são contínuas, segue que \bar{m} é contínua.

Para mostrar que \bar{m} é aberta, considere $x_0 = H \in G/H$. Então $m_{x_0} : G \rightarrow G/H$, $g \mapsto gH$, coincide com a projeção natural π que é aberta. Como G/H é homogêneo e m_{x_0} é aberta para um $x_0 \in G/H$, segue que a ação de G em G/H é aberta. \square

Proposição B.12. Se $H \subset G$ é um subgrupo fechado, então $G/H = \{gH : g \in G\}$ munido com a topologia quociente é Hausdorff.

Demonstração. Dados $gH \neq g'H$ em G/H , ou equivalentemente, $g^{-1}g' \notin H$, queremos $U \ni gH$ e $U' \ni g'H$ abertos e disjuntos em G/H . Ora, considerando a projeção natural $\pi : G \rightarrow G/H$, $g \mapsto gH$, temos que os abertos em G/H são imagens de abertos em G . Assim, desejamos $W \ni g$ e $W' \ni g'$ abertos em G tais que $\pi(W) \cap \pi(W') = \emptyset$, o que ocorre se, e somente se, $x^{-1}x' \notin H$, $\forall x \in W$, $\forall x' \in W'$, uma vez que $xH \neq x'H \iff x^{-1}x' \notin H$. Então é suficiente que existam abertos em G , $W \ni g$ e $W' \ni g'$ satisfazendo $W^{-1}W' \subset G - H$. Encontremos tais W e W' . Para isto, considere a aplicação

$$\begin{aligned} \varphi : G \times G &\longrightarrow G \\ (x, x') &\longmapsto x^{-1}x' \end{aligned}$$

que é contínua, já que G é grupo topológico. Como H é fechado em G , temos que $G - H \ni g^{-1}g'$ é aberto em G . Como φ é contínua, $\varphi^{-1}(G - H) \subset G \times G$ é aberto na topologia produto de $G \times G$ que contém (g, g') . Logo existem $W \ni g$ e $W' \ni g'$ abertos em G tais que $W \times W' \subset \varphi^{-1}(G - H)$, isto é, $W^{-1}W' = \varphi(W \times W') \subset G - H$, conforme queríamos. \square

Fixado $x_0 \in X$, a *isotropia* de x_0 em G , dada por

$$G_{x_0} = \{g \in G : gx_0 = x_0\},$$

é claramente um subgrupo de G .

Proposição B.13. Valem as seguintes afirmações.

- (1) Se X é Hausdorff, então G_{x_0} é fechado em G , $\forall x_0 \in X$.
- (2) $G_{gx_0} = gG_{x_0}g^{-1}$, $\forall g \in G$, $\forall x_0 \in X$.

Demonstração. (1) Como a ação $m : G \times X \rightarrow X$ é contínua, $m_{x_0} : g \in G \mapsto gx_0 \in X$ é contínua e daí $G_{x_0} = m_{x_0}^{-1}(x_0)$ é fechado em G , já que $\{x_0\}$ é fechado em X , pois X é Hausdorff.

(2) Temos $g' \in G_{gx_0} \iff g'(gx_0) = gx_0 \iff (g^{-1}g'g)x_0 = x_0 \iff g^{-1}g'g \in G_{x_0} \iff g' \in gG_{x_0}g^{-1}$. Portanto, $G_{gx_0} = gG_{x_0}g^{-1}$. \square

Dado $x \in X$, chamamos o subconjunto

$$Gx = \{gx : g \in G\} = m_x(G) \subset X$$

de G -órbita de x . A ação de G em X é *transitiva* se $Gx = X$ para um, e portanto para todo, $x \in X$, ou seja, X é uma G -órbita. Equivalentemente, dados $x, y \in X$ existe $g \in G$ tal que $gx = y$.

Definição B.14. Um espaço topológico X é dito *homogêneo* quando a ação de G em X é transitiva. A ação de G em X é dita *aberta* se a ação m_x , de G em um ponto base x , é aplicação aberta para todo $x \in X$.

Para que a ação de G em um espaço homogêneo X seja aberta é suficiente que m_{x_0} seja aplicação aberta para algum $x_0 \in X$. Com efeito, se $x \in X = Gx_0$ então $x = gx_0$ para algum $g \in G$. Logo $m_x = m_{gx_0} = m_{x_0} \circ D_g$ é aberta, pois são abertas m_{x_0} e D_g que é homeomorfismo.

Seja $x_0 \in X$ ponto base fixado no espaço homogêneo X e considere a isotropia G_{x_0} . A aplicação

$$\begin{aligned} \phi_{x_0} : G/G_{x_0} &\longrightarrow X \\ gG_{x_0} &\longmapsto gx_0 \end{aligned}$$

é uma bijeção contínua G -equivariante. De fato,

$$gG_{x_0} = g'G_{x_0} \iff g^{-1}g' \in G_{x_0} \iff g^{-1}g'x_0 = x_0 \iff gx_0 = g'x_0$$

mostra que ϕ_{x_0} é bem definida e injetiva. A sobrejetividade de ϕ_{x_0} decorre de X ser espaço homogêneo de G . Além disso, como $\phi_{x_0} : G/G_{x_0} \longrightarrow X$ é contínua se, e somente se, seu levantamento $m_{x_0} = \phi_{x_0} \circ \pi : G \longrightarrow X$ é contínuo, onde $\pi : G \rightarrow G/G_{x_0}$ é a projeção natural (ver diagrama abaixo), segue que ϕ_{x_0} é contínua. Finalmente, que ϕ_{x_0} é G -equivariante, decorre de a ação de G em X ser conjugada por ϕ_{x_0} à ação de G em G/G_{x_0} : dados $g \in G$ e $g'G_{x_0} \in G/G_{x_0}$,

$$g\phi_{x_0}(g'G_{x_0}) = g(g'x_0) = (gg')x_0 = \phi_{x_0}(gg'G_{x_0}) = \phi_{x_0}(gg'G_{x_0}).$$

Logo,

$$g \circ \phi_{x_0} = \phi_{x_0} \circ g, \quad \forall g \in G.$$

Temos então o seguinte diagrama comutativo de aplicações G -equivariantes:

$$\begin{array}{ccc} G & & \\ \pi \downarrow & \searrow m_{x_0} & \\ G/G_{x_0} & \xrightarrow{\phi_{x_0}} & X \end{array}$$

Proposição B.15. Seja X um espaço homogêneo de G e fixe $x_0 \in X$. A bijeção contínua ϕ_{x_0} é homeomorfismo se, e somente se, a ação de G em X é aberta.

Demonstração. Suponha que ϕ_{x_0} seja homeomorfismo. Como o diagrama acima comuta, isto é, $m_{x_0} = \phi_{x_0} \circ \pi$ e $\pi : G \rightarrow G/G_{x_0}$ é aplicação aberta, segue que m_{x_0} é aberta e portanto a ação de G em X é aberta. Reciprocamente, suponha que a aplicação m_{x_0} é aberta. Se $U \subset G/G_{x_0}$ é aberto em G/G_{x_0} , então $U = \pi(W)$ com $W \subset G$ aberto em G . Assim, $(\phi_{x_0}^{-1})^{-1}(U) = \phi_{x_0} \circ \pi(W) = m_{x_0}(W) \subset X$ é aberto em X . Portanto, imagem inversa de aberto em G/G_{x_0} por $\phi_{x_0}^{-1}$ é aberto em X , ou seja, $\phi_{x_0}^{-1}$ é contínua e portanto ϕ_{x_0} é homeomorfismo. \square

Teorema B.16. Sejam X um conjunto e G um grupo topológico agindo transitivamente em X . Então existe única topologia em X tal que a ação de G em X é contínua e aberta. Dado $x_0 \in X$, esta é a topologia tal que a aplicação $\phi_{x_0} : G/G_{x_0} \rightarrow X$, $gG_{x_0} \mapsto gx_0$, é homeomorfismo. Além disso, X é Hausdorff nesta topologia se, e somente se, a isotropia G_{x_0} é fechada em G .

Demonstração. (Unicidade) Considere uma topologia em X para a qual a ação de G em X é contínua e aberta. Pela proposição anterior ϕ_{x_0} é homeomorfismo. Assim, a topologia de X é a mesma do quociente G/G_{x_0} , que não depende de X .

(Existência) Considere em X a topologia tal que a bijeção ϕ_{x_0} é um homeomorfismo. A ação de G em X se conjuga pelo homeomorfismo ϕ_{x_0} à ação de G em G/G_{x_0} , ou seja, $g \circ \phi_{x_0} = \phi_{x_0} \circ g$, $\forall g \in G$. Assim é suficiente mostrar que a ação de G em G/G_{x_0} é contínua e aberta. Mas isto decorre da Proposição B.11, considerando $H = G_{x_0}$.

Finalmente, se X com a topologia construída acima é Hausdorff, então a isotropia G_{x_0} é fechada em G , $\forall x_0 \in X$, conforme a Proposição B.13 (1). Reciprocamente, se a isotropia G_{x_0} é fechada em G , segue da Proposição B.13 (3) que G/G_{x_0} é Hausdorff e portanto X também o é, já que é homeomorfo a G/G_{x_0} . \square

Observação B.17. *Vimos que se X é um espaço homogêneo do grupo topológico G , a aplicação $\phi_{x_0} : G/G_{x_0} \rightarrow X$ é homeomorfismo se, e somente se, a ação de G em X é aberta. Uma condição suficiente sobre G e X para que uma ação transitiva e contínua de G em X seja aberta é a seguinte: G é separável, localmente compacto, separa fechados e pontos por abertos e X é espaço de Baire Hausdorff. Para uma demonstração, ver Proposição 2.28 de [15].*

B.4 Fluxos e sistemas dinâmicos

Sejam X um espaço topológico e $\mathbb{T} = \mathbb{R}$ ou \mathbb{Z} . Dizemos que uma aplicação $\Phi : \mathbb{T} \times X \rightarrow X$ é um *fluxo* em X , se ela é contínua e satisfaz as seguintes condições, denominadas *propriedades do fluxo*.

$$\begin{aligned}\Phi(t + s, x) &= \Phi(t, \Phi(s, x)), \\ \Phi(0, x) &= x,\end{aligned}$$

para quaisquer $t, s \in \mathbb{T}$ e qualquer $x \in X$. De outro modo, pondo $\Phi^t : X \rightarrow X$, $\Phi^t(x) = \Phi(t, x)$, para cada $t \in \mathbb{T}$, as condições acima ficam

$$\begin{aligned}\Phi^{t+s} &= \Phi^t \circ \Phi^s, \quad \forall t, s \in \mathbb{T} \\ \Phi^0 &= \text{id}_X.\end{aligned}$$

No caso em que $\mathbb{T} = \mathbb{Z}$, temos que $\phi : X \rightarrow X$ definida por $\phi = \Phi^1$ é homeomorfismo com inversa Φ^{-1} . Além disso, é fácil ver das propriedades do fluxo que $\Phi(k, x) = \phi^k(x)$ para todo $k \in \mathbb{Z}$ e todo $x \in X$. Reciprocamente, dado um homeomorfismo $\phi : X \rightarrow X$, temos que $\Phi(k, x) = \phi^k(x)$, com $k \in \mathbb{Z}$, define um fluxo em X . Portanto, um fluxo de tempo discreto é dado pela iteração de um homeomorfismo.

No caso em que $\mathbb{T} = \mathbb{R}$, X é um aberto U de um espaço vetorial V e Φ é diferenciável, podemos considerar o campo tangente $F : U \rightarrow V$ às curvas do fluxo dado por

$$F(x) = \left. \frac{d}{dt} \Phi^t(x) \right|_{t=0}.$$

Nesse caso, a curva é solução da EDO com condição inicial

$$x'(t) = F(x(t)), \quad x(0) = x.$$

De fato,

$$\frac{d}{dt}\Phi^t(x) = \frac{d}{ds}\Phi^{t+s}(x)\Big|_{s=0} = \frac{d}{ds}\Phi^s(\Phi^t(x))\Big|_{s=0} = F(\Phi^t(x)), \quad \Phi^0(x) = x.$$

Reciprocamente, as soluções uma EDO autônoma dão origem a um fluxo local $\Phi^t(x)$ que, para cada x , está definida para t suficientemente próximo de zero. Para mais detalhes, ver [8].

Um *sistema dinâmico* em X é um fluxo em X , de tempo discreto ou contínuo. O estudo de um sistema dinâmico consiste, a grosso modo, em analisar os possíveis comportamentos de longo prazo de suas trajetórias, isto é, os possíveis limites de $\Phi^t(x)$ quando $t \rightarrow \infty$. Uma ferramenta importante para esse estudo é uma *função de Lyapunov* para Φ^t , que é uma função real $f : X \rightarrow \mathbb{R}$ que decresce ao longo das trajetórias do fluxo, isto é, tal que para cada $x \in X$, a aplicação $t \mapsto f(\Phi^t(x))$ é decrescente.

Referências Bibliográficas

- [1] Duistermaat, J.J., Kolk, J.A.C. e Varadarajan, V.S.: Functions, flows and oscillatory integral on flag manifolds. *Compositio Math.* 49, p. 309 à 398. 1983.
- [2] Ferraiol, T., Patrão, M. e Seco, L.: Jordan decomposition and dynamics on flag manifolds, *Discrete Contin. Dynam. Systems A*, 26(3). 2010.
- [3] Fulton, W. and Harris, J.: Representation Theory, A first course. Graduate Texts in Mathematics. Springer-Verlag. 1991.
- [4] Golub, G. e Frank, U.: The QR algorithm: 50 years later its genesis by John Francis and Vera Kublanovskaya and subsequent developments. *IMA Journal of Numerical Analysis* 29, p. 467 à 485. 2009.
- [5] Guest, M.: Morse Theory in the 1990's, In. *Invitations to Geometry and Topology*, M. Bridson S. Salamon (Eds). Oxford Graduate Texts in Mathematics 7. 2002.
- [6] Helmke, U. e Kleinsteuber, M.: A differential equation for diagonalizing complex semisimple Lie algebra elements. *Systems & Control Letters* 59. Elsevier p. 72 à 78. 2010.
- [7] Helmke, U. e Moore, J.B.: Optimization and Dynamical Systems. Springer-Verlag, Londres. 1994.
- [8] Hirsch, M. W. e Smale, S.: Differential Equations, Dynamical Systems, and Linear Algebra. Academic Press. 1974.
- [9] Lancaster, P. and Tismenetsky, M.: The Theory of Matrices, 2nd ed., with Applications. Academic Press. 1985.
- [10] Lusztig, G.: Bruhat decomposition and applications. Lecture given at a one day conference in memory of F. Bruhat held at the Institut Poincaré, Paris. 2010.
- [11] Mathews, J. H.: The QR Method for Eigenvalues. Módulo disponível em <http://math.fullerton.edu/mathews/n2003/QRMethodMod.html>
- [12] Munkres, J. R.: Topology, 2nd ed. Prentice Hall. 2000.
- [13] Peter, J. O.: Orthogonal Bases and the QR Algorithm. Notas disponíveis em http://www.math.umn.edu/~olver/aims_/qr.pdf

- [14] Przybylska, M.: On matrix differential equations and abstract FG algorithm. *Lin. Algebra Appl.* 346 (1 - 3) p. 155 à 175. 2002.
- [15] San Martin, L. A. B.: Grupos de Lie. Notas de aula disponíveis em <http://www.ime.unicamp.br/~smartin/cursos/grupolie-2011/gruplie0.pdf>
- [16] Shub, M. and Vasquez, A. T.: Some linearly induced Morse-Smale systems, the QR algorithm and the Toda lattice. Cambridge, Mass. and Amherst, Mass. p. 181 à 194. 1985.
- [17] Tyrtysnikov, E.: Matrix Bruhat decompositions with a remark on the QR (GR) algorithm. *Linear Algebra and Its Applications*, 250. p. 61 à 68. 1997.