capítulo 2

Preliminares

2.1 Sobre este capítulo

Este capítulo apresenta a notação a ser utilizada subsequentemente neste trabalho e uma série de resultados preliminares que serão utilizados nos capítulos seguintes. Resultados e definições apresentados neste capítulo servem de forma geral aos capítulos seguintes. Resultados com pertinência apenas dentro de um problema ou técnica específicos estão agrupados dentro do capítulo que trata do referido problema ou técnica.

A seção 2.2, a seguir, apresenta a notação básica a ser utilizada em todo este trabalho. Na seção 2.3 são apresentados alguns resultados básicos, seguidos por resultados ligados ao conceito de estabilidade na média quadrática (na subseção 2.3.2), testes para se determinar estabilizabilidade e detectabilidade de um sistema markoviano (na subseção 2.3.3) e por um resultado auxiliar extremamente importante (na subseção 2.3.4). Finalmente, a seção 2.4 estende os conceitos usuais de normas H_2 e H_∞ para sistemas lineares discretos sujeitos a saltos markovianos.

2.2 Notação

A seguinte notação será usada ao longo de todo este trabalho.

$\mathbb{R}^n$ e denotam os espaços -dimensionais real e complexo respectivamente. denota o espaço linear limitado normado de todas as matrizes complexas . Da mesma forma,

$\mathbb{B}(\mathbb{R}^n, \mathbb{R}^m)$ denota o espaço linear limitado normado de todas as matrizes reais $m \times n$. Por simplicidade, $\mathbb{B}(\mathbb{C}^n, \mathbb{C}^n) := \mathbb{B}(\mathbb{C}^n)$ e $\mathbb{B}(\mathbb{R}^n, \mathbb{R}^n) := \mathbb{B}(\mathbb{R}^n)$.

A^* indica o operador adjunto aplicado à matriz complexa A. Caso a matriz A seja real, a mesma notação será utilizada para indicar a transposta de A. $L > 0$ e $L \geq 0$ significam que a matriz hermitiana L é positiva definida e positiva semi-definida respectivamente.

Para uma matriz $A \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$, a inversa generalizada de A (ou inversa de Moore-Penrose[1] de A) será denotada por $A^\dagger$. A inversa generalizada de A é a matriz única $A^\dagger \in \mathbb{B}(\mathbb{C}^m, \mathbb{C}^n)$ tal que (veja [50], pág. 13)

$$\begin{aligned} AA^\dagger A &= A \\ A^\dagger AA^\dagger &= A^\dagger \\ (AA^\dagger)^* &= AA^\dagger \\ (A^\dagger A)^* &= A^\dagger A. \end{aligned}$$

Para espaços complexos de Banach[2] $\mathbb{X}$ e $\mathbb{Y}$, $\mathbb{B}(\mathbb{X}, \mathbb{Y})$ denota o espaço de Banach de todos os operadores lineares limitados de $\mathbb{X}$ em $\mathbb{Y}$, com a norma induzida uniforme representada por $\|.\|$. Por simplicidade, $\mathbb{B}(\mathbb{X}, \mathbb{X}) := \mathbb{B}(\mathbb{X})$.

O raio espectral de um operador $\mathcal{T} \in \mathbb{B}(\mathbb{X})$ é denotado por $r_\sigma(\mathcal{T})$. $\mathcal{T}^*$ representa o operador adjunto de $\mathcal{T}$. $\mathcal{T} > 0$ e $\mathcal{T} \geq 0$ significa que o operador $\mathcal{T}$ é positivo definido e positivo semi-definido respectivamente. Se $\mathbb{X}$ for um espaço de Hilbert[3], então o produto interno será representado por $\langle .;. \rangle$.

$\mathbb{H}^{n,m}$ denota o espaço linear formado por todas as N-uplas de matrizes $V = (V_1, \ldots, V_N)$ com $V_i \in \mathbb{B}(\mathbb{C}^n, \mathbb{C}^m)$ para $i = 1, \ldots, N$. Por simplicidade, $\mathbb{H}^{n,n} := \mathbb{H}^n$. Da mesma forma, $\mathbb{G}^{n,m}$ denota o espaço linear formado por todas as N-uplas de matrizes reais.

Para $V \in \mathbb{H}^{n,m}$, as seguintes normas são consideradas:

---

1 **Eliakim Hastings Moore** (1862-1932). Matemático americano.

2 **Stefan Banach** (1892-1945). Matemático polônes (nascido no então Império Austro-húngaro). Banach foi o fundador da moderna análise funcional. Os espaços de Banach, introduzidos por ele em 1920 receberam esse nome como uma homenagem de Fréchet (matemático francês falecido em 1973).

3 **David Hilbert** (1862-1943). Matemático russo (nascido no que era à época território prussiano). Conhecido especialmente por seus trabalhos em Geometria.

$$\|V\|_1 := \sum_{i=1}^{N} \|V_i\| \tag{2.1.a}$$

$$\|V\|_2 := \left(\sum_{i=1}^{N} \text{tr}\,(V_i^* V_i) \right)^{1/2} . \tag{2.1.b}$$

Pode-se verificar facilmente que $\mathbb{H}^{n,m}$ equipado com qualquer das normas acima é um espaço de Banach, e que, de fato, $(\|.\|_2, \mathbb{H}^{n,m})$ é um espaço de Hilbert, com produto interno dado por

$$\langle V;S \rangle := \sum_{i=1}^{N} \text{tr}\,(V_i^* S_i) \tag{2.1.c}$$

para $V = (V_1, \ldots, V_N) \in \mathbb{H}^{n,m}$ e $S = (S_1, \ldots, S_N) \in \mathbb{H}^{n,m}$.

$V = (V_1, \ldots, V_N) \in \mathbb{H}^n$ é dita hermitiana se $V_i = V_i^*$ para $i = 1, \ldots, N$. Define-se também

$$\mathbb{H}^{n^*} := \{ V = (V_1, \ldots, V_N) \in \mathbb{H}^n ; V_i = V_i^* \text{ para } i = 1, \ldots, N \},$$

$$\mathbb{H}^{n+} := \{ V = (V_1, \ldots, V_N) \in \mathbb{H}^n ; V_i \geq 0 \text{ para } i = 1, \ldots, N \}.$$

Para $V = (V_1, \ldots, V_N) \in \mathbb{H}^n$ e $S = (S_1, \ldots, S_N) \in \mathbb{H}^n$, diz-se que $V \geq S$ se $(V - S) \in \mathbb{H}^{n+}$ onde $(V - S) = (V_1 - S_1, \ldots, V_N - S_N)$. Da mesma forma, $V > S$ se $V_i - S_i > 0$ para $i = 1, \ldots, N$.

Para uma filtragem crescente $\{\Im_k\}$ definida num espaço de probabilidades $(\Omega, \Im, \mathcal{P})$ denota-se por l_2^r o espaço de Hilbert formado pela sequência de variáveis aleatórias de segunda ordem $z = (z(0), z(1), \ldots)$ com $z(k) \in \mathbb{R}^r$ e $\Im_k$-adaptadas para $k = 0, 1, \ldots$, tais que

$$\|z\|_2^2 := \sum_{k=0}^{\infty} \|z(k)\|_2^2 < \infty ,$$

onde

$$\|z(k)\|_2^2 := E(\|z(k)\|^2) .$$

Esta norma se origina do produto interno

$$\langle y;z \rangle := \sum_{k=0}^{\infty} E(y(k)^* z(k))$$

onde $y(k), z(k) \in \mathbb{C}^r$ para $k = 0, 1, \ldots$. Veja [17] para maiores detalhes sobre a construção do espaço de probabilidades $(\Omega, \Im, \mathcal{P})$ e da filtragem crescente $\Im_k$.

2.3 Definições básicas e resultados auxiliares

2.3.1 Preliminares

Alguns resultados conhecidos são apresentados nos comentários abaixo, os quais serão úteis posteriormente. O primeiro deles é um resultado bastante usado em DML´s (Desigualdades Matriciais Lineares), enquanto que o segundo será utilizado na prova do lema 2.13, a seguir.

Comentário 2.1: Complemento de Schur[1].

$W = \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \geq 0$ se e somente se $Q \geq 0$, $S = QQ^{\dagger}S$, e $R - S^*Q^{\dagger}S \geq 0$. Em particular, se $R > 0$, então $W = \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \geq 0$ se e somente se $Q \geq SR^{-1}S^*$ (veja [50], p.13).

Comentário 2.2: Para qualquer $W \in \mathbb{B}(\mathbb{C}^n)$, existem W^j, $j = 1, 2, 3, 4$, tais que $W^j \geq 0$ para $j = 1, 2, 3, 4$ e $W = (W^1 - W^2) + \sqrt{-1}(W^3 - W^4)$ (cf. [43]). Naturalmente, para qualquer $S = (S_1, \ldots, S_N) \in \mathbb{H}^n$, é possível encontrar $S^j \in \mathbb{H}^{n+}$, $j = 1, 2, 3, 4$, tais que $S = (S^1 - S^2) + \sqrt{-1}(S^3 - S^4)$.

2.3.2 Estabilidade para sistemas markovianos

Esta subseção apresenta os principais fatos e resultados relacionados à estabilidade de sistemas lineares discretos sujeitos a saltos markovianos. O conceito de estabilidade apresentado aqui, o de estabilidade no sentido da média quadrada (veja [14]), será utilizado em todo o trabalho.

Seja o seguinte sistema linear estocástico a tempo discreto definido em um espaço de probabilidades apropriado dado por $(\Omega, \{\Im_k\}, \Im, \mathcal{P})$,

$$\begin{aligned} x(k+1) &= \Gamma_{\theta(k)}x(k) \\ x(0) = x_0 \qquad & \theta(0) = \theta_0 \end{aligned} \tag{2.2}$$

e seja também uma versão controlada do mesmo sistema,

$$\begin{aligned} x(k+1) &= A_{\theta(k)}x(k) + B_{\theta(k)}u(k) \\ x(0) = x_0 \qquad & \theta(0) = \theta_0 \end{aligned} \tag{2.3}$$

onde $\theta(k)$, $k = 0, 1, \ldots$, é uma cadeia de Markov discreta com um espaço de estados finito, dado

[1] **Issai Schur** (1875-1941). Nasceu em Belarus (então Bielo-Rússia), mas fez sua carreira na Universidade de Berlim, onde foi aluno e colaborador de Frobenius, até 1939, quando a perseguição nazista o obrigou mudar para a Palestina.

pelo conjunto $\{1, \ldots, N\}$ e onde cada estado se associa aos demais por uma matriz de probabilidades de transição $P = [p_{ij}] \in \mathbb{B}(\mathbb{R}^n)$.

$\Gamma = (\Gamma_1, \ldots, \Gamma_N) \in \mathbb{H}^n$, $A = (A_1, \ldots, A_N) \in \mathbb{H}^n$, $B = (B_1, \ldots, B_N) \in \mathbb{H}^{n,m}$ e x_0 é uma variável aleatória de segunda ordem em $\mathbb{C}^n$. Obviamente, quando $u(k) = F_{\theta(k)}x(k)$ para algum $F = (F_1, \ldots, F_N) \in \mathbb{H}^{n,m}$, o sistema (2.3) é equivalente ao sistema (2.2) com $\Gamma_i = A_i + B_iF_i$ para $i = 1, \ldots, N$.

Para $\Gamma = (\Gamma_1, \ldots, \Gamma_N) \in \mathbb{H}^n$ e a matriz de probabilidades de transição $P = [p_{ij}] \in \mathbb{B}(\mathbb{R}^n)$ tal que $p_{ij} \geq 0$ para $i, j = 1, \ldots, N$, definem-se os operadores: $\mathcal{E}(.) = (\mathcal{E}_1(.), \ldots, \mathcal{E}_N(.)) \in \mathbb{B}(\mathbb{H}^n)$, $\mathcal{L}(.) = (\mathcal{L}_1(.), \ldots, \mathcal{L}_N(.)) \in \mathbb{B}(\mathbb{H}^n)$ e $\mathcal{T}(.) = (\mathcal{T}_1(.), \ldots, \mathcal{T}_N(.)) \in \mathbb{B}(\mathbb{H}^n)$ tais que, para $V = (V_1, \ldots, V_N) \in \mathbb{H}^n$ e $i = 1, \ldots, N$,

$$\mathcal{E}_i(V) := \sum_{j=1}^{N} p_{ij}V_j, \tag{2.4.a}$$

$$\mathcal{L}_i(V) := \Gamma_i^* \mathcal{E}_i(V)\Gamma_i, \tag{2.4.b}$$

$$\mathcal{T}_j(V) := \sum_{i=1}^{N} p_{ij}\Gamma_i V_i \Gamma_i^* \tag{2.4.c}$$

para $j = 1, \ldots, N$.

É fácil verificar, pelo produto interno dado por (2.1.c), que $\mathcal{T} = \mathcal{L}^*$. Além disso, os operadores $\mathcal{E}$, $\mathcal{L}$ e $\mathcal{T}$ mapeiam $\mathbb{H}^{n^*}$ em $\mathbb{H}^{n^*}$ e $\mathbb{H}^{n+}$ em $\mathbb{H}^{n+}$.

Mostra-se em [14] que o sistema (2.2) e $V(k) = (V_1(k), \ldots, V_N(k)) \in \mathbb{H}^{n+}$ com $V_i(k) = E(x(k)x(k)^* 1_{\{\theta(k) = i\}})$ para $i = 1, \ldots, N$, onde $1_{\{.\}}$ é a medida de Dirac, levam a

$$V(k+1) = \mathcal{T}(V(k)) \text{ para } k = 0, 1, \ldots$$

onde $\mathcal{T}$ é como em (2.4.c).

A seguir é definido o conceito de estabilidade utilizado neste trabalho. No decorrer do texto, se utilizará tanto a expressão "estável na média quadrática" como "estável no sentido da média quadrática" . Ambas as expressões têm o mesmo significado, dado pela definição abaixo.

Definição 2.3: O sistema dado por (2.2) é dito estável na média quadrática se $\|x(k)\|_2^2 = E(\|x(k)\|^2) \to 0$ quando $k \to \infty$ para quaisquer condição inicial x_0 e distribui-

ção inicial para θ_0.

Comentário 2.4: Pode-se mostrar que a estabilidade de cada modelo (isto é, o máximo autovalor de Γ_i em módulo para $i = 1, \ldots, N$ ser menor que 1) não é uma condição necessária nem suficiente para a estabilidade na média quadrática (veja [34]). Além disso, se (2.2) é estável na média quadrática, então $\|x(k)\| \to 0$ quando $k \to \infty$ com probabilidade 1 (veja [14]).

Definição 2.5: Diz-se que $F = (F_1, \ldots, F_N) \in \mathbb{H}^{n,m}$ estabiliza (A, B) no sentido da média quadrática se, ao fazer $u(k) = F_{\theta(k)}x(k)$ no sistema (2.3), $\|x(k)\|_2^2 = E(\|x(k)\|^2) \to 0$ quando $k \to \infty$ para quaisquer x_0 e θ_0.

O resultado a seguir, provado nos teoremas 1 e 2 de [14], entre outras coisas, relaciona a estabilidade do sistema com os operadores definidos acima.

Proposição 2.6: As seguintes afirmações são equivalentes:

i) O sistema dado por (2.2) é estável na média quadrática.

ii) $\|V(k)\|_2 \to 0$ quando $k \to \infty$ para quaisquer condição inicial x_0 e distribuição inicial θ_0.

iii) $r_\sigma(\mathcal{T}) < 1$.

iv) $r_\sigma(\mathcal{L}) < 1$.

v) Existem $\alpha \in (0, 1)$ e $a \in \mathbb{R}$, $a > 0$, tais que para $k = 0, 1, \ldots$

$$E(\|x(k)\|^2) \leq a\alpha^k.$$

vi) (equações acopladas de Lyapunov) Dado $S = (S_1, \ldots, S_N) > 0$ em $\mathbb{H}^{n+}$, existe $X = (X_1, \ldots, X_N) > 0$ em $\mathbb{H}^{n+}$ satisfazendo

$$X - \mathcal{T}(X) = S \text{ com } X = \sum_{k=0}^{\infty} \mathcal{T}^k(S).$$

vii) (equações acopladas de Lyapunov adjuntas) Dado $S = (S_1, \ldots, S_N) > 0$ em $\mathbb{H}^{n+}$, existe $X = (X_1, \ldots, X_N) > 0$ em $\mathbb{H}^{n+}$ satisfazendo

$$X - \mathcal{L}(X) = S \text{ com } X = \sum_{k=0}^{\infty} \mathcal{L}^k(S).$$

Além disso, se $r_\sigma(\mathcal{T}) < 1$, então para qualquer $S \in \mathbb{H}^n$ existe uma única $X \in \mathbb{H}^n$ tal que $X - \mathcal{T}(X) = S$. Se $S \geq T \geq 0$ (> 0 respectivamente) e $L - \mathcal{T}(L) = T$, então $X \geq L \geq 0$ (> 0).

Esses resultados também valem ao se trocar $\mathcal{T}$ por $\mathcal{L}$.

Na definição abaixo são introduzidos os conceitos de *estabilizabilidade na média quadrática* e *detectabilidade na média quadrática* para sistemas sujeitos a saltos markovianos (veja [15]).

Definição 2.7: Sejam $A = (A_1, \ldots, A_N) \in \mathbb{H}^n$, $B = (B_1, \ldots, B_N) \in \mathbb{H}^{m,n}$ e $M = (M_1, \ldots, M_N) \in \mathbb{H}^{n,s}$. Diz-se que o par (A, B) é estabilizável na média quadrática se existir $F = (F_1, \ldots, F_N) \in \mathbb{H}^{n,m}$ que estabilize (A, B) (veja a definição 2.5). De maneira análoga, diz-se que o par (M, A) é detectável na média quadrática se existir $H = (H_1, \ldots, H_N) \in \mathbb{H}^{s,n}$ tal que o sistema (2.2) com $\Gamma_i = A_i + H_i M_i$ para $i = 1, \ldots, N$ é estável na média quadrática. Diz-se nesse caso que H estabiliza (M, A) na média quadrática.

2.3.3 Testes de estabilizabilidade e detectabilidade

As duas proposições a seguir apresentam condições necessárias e suficientes para que um dado sistema markoviano seja estabilizável ou detectável. Em ambos os casos, as formulações utilizadas para essas condições conduzem a uma maneira factível de se testar a estabilizabilidade ou detectabilidade via programação convexa.

Proposição 2.8: O par (A, B) é estabilizável na média quadrática se e somente se existirem $W_1 = (W_{11}, \ldots, W_{N1}) \in \mathbb{H}^{n+}$, $W_2 = (W_{12}, \ldots, W_{N2}) \in \mathbb{H}^{m,n}$, $W_3 = (W_{13}, \ldots, W_{N3}) \in \mathbb{H}^{m+}$ tais que para $j = 1, \ldots, N$,

$$\sum_{i=1}^{N} p_{ij}(A_i W_{i1} A_i^* + B_i W_{i2}^* A_i^* + A_i W_{i2} B_i^* + B_i W_{i3} B_i^*) - W_{j1} < 0 \tag{2.5.a}$$

$$W_{j3} \geq W_{j2}^* W_{j1}^{-1} W_{j2} \tag{2.5.b}$$

$$W_{j1} > 0 \tag{2.5.c}$$

Prova: Inicialmente seja o seguinte conjunto:
$\Psi := \{ W_1 \in \mathbb{H}^{n+}, W_2 \in \mathbb{H}^{m,n}, W_3 \in \mathbb{H}^{m+};\ W_{i1}, W_{i2}, W_{i3}$ satisfazem as equações (2.5.a)-(2.5.c) para $j = 1, \ldots, N \}$.
(necessidade): Assumindo que (A, B) seja estabilizável na média quadrática, então existe $F = (F_1, \ldots, F_N) \in \mathbb{H}^{n,m}$ tal que o sistema (2.2) com $\Gamma_j = A_j + B_j F_j$ para $j = 1, \ldots, N$ é estável na média quadrática (veja a definição 2.7). Tem-se da proposição 2.6 - item (vi), que existe $X = (X_1, \ldots, X_N) > 0$ em $\mathbb{H}^{n+}$ tal que

$$-X_j + \sum_{i=1}^{N} p_{ij}(A_i + B_iF_i)X_i(A_i + B_iF_i)^* < 0 \text{ para } j = 1, \ldots, N. \qquad (2.6)$$

Tomando $W_{j1} = X_j$, $W_{j2} = W_{j1}F_j^*$ e $W_{j3} = W_{j2}^*W_{j1}^{-1}W_{j2}$ para $j = 1, \ldots, N$, verifica-se facilmente a partir de (2.6) que as equações (2.5.a)-(2.5.c) são satisfeitas e que $\Psi \neq \emptyset$.

(suficiência): Assuma que $\Psi \neq \emptyset$, ou seja, que existem $W_1 \in \mathbb{H}^{n+}$, $W_2 \in \mathbb{H}^{m,n}$ e $W_3 \in \mathbb{H}^{m+}$ satisfazendo (2.5.a)-(2.5.c). Seja também $F = (F_1, \ldots, F_N) \in \mathbb{H}^{n,m}$ dado por $F_j = W_{j2}^*W_{j1}^{-1}$ para $j = 1, \ldots, N$. Tem-se que para $j = 1, \ldots, N$,

$$-W_{j1} + \sum_{i=1}^{N} p_{ij}(A_i + B_iF_i)X_i(A_i + B_iF_i)^* =$$

$$= -W_{j1} + \sum_{i=1}^{N} p_{ij}(A_i + B_iW_{i2}^*W_{i1}^{-1})X_i(A_i + B_iW_{i2}^*W_{i1}^{-1})^*$$

$$= -W_{j1} + \sum_{i=1}^{N} p_{ij}(A_iW_{i1}A_i^* + B_iW_{i2}^*A_i^* + A_iW_{i2}B_i^* + B_iW_{j2}^*W_{j1}^{-1}W_{j2}B_i^*)$$

$$\leq -W_{j1} + \sum_{i=1}^{N} p_{ij}(A_iW_{i1}A_i^* + B_iW_{i2}^*A_i^* + A_iW_{i2}B_i^* + B_iW_{i3}B_i^*) < 0.$$

Da proposição 2.6 tem-se que F estabiliza (A, B) na média quadrática, e portanto (A, B) é claramente estabilizável na média quadrática.

□

Proposição 2.9: O par (M, A) é detectável na média quadrática se e somente se existirem $W_1 = (W_{11}, \ldots, W_{N1}) \in \mathbb{H}^{n+}$, $Z = (Z_1, \ldots, Z_N) \in \mathbb{H}^{n+}$, $W_2 = (W_{12}, \ldots, W_{N2}) \in \mathbb{H}^{s,n}$ $W_3 = (W_{13}, \ldots, W_{N3}) \in \mathbb{H}^{s+}$ tais que para $i = 1, \ldots, N$

$$A_i^*Z_iA_i + M_i^*W_{i2}^*A_i + A_i^*W_{i2}M_i + M_i^*W_{i3}M_i - W_{i1} < 0 \qquad (2.7.a)$$

$$W_{i3} \geq W_{i2}^*Z_i^{-1}W_{i2} \qquad (2.7.b)$$

$$Z_i \geq \mathcal{E}_i(W_1) \qquad (2.7.c)$$

$$W_{i1} > 0 \qquad (2.7.d)$$

$$Z_i > 0. \qquad (2.7.e)$$

Prova: Esta prova é feita de maneira análoga à da proposição anterior. Seja o conjunto

$\Phi := \{ W_1, Z \in \mathbb{H}^{n+}, W_2 \in \mathbb{H}^{s,n}, W_3 \in \mathbb{H}^{s+};\ W_{i1}, Z_i, W_{i2}, W_{i3}$ satisfazem as equações (2.7.a)-(2.7.e)para $i = 1, \ldots, N \}$.

(necessidade): Assumindo que o par (M, A) seja detectável na média quadrática, existe $H = (H_1, \ldots, H_N) \in \mathbb{H}^{s,n}$ tal que o sistema (2.2) com $\Gamma_i = A_i + H_i M_i$ para $i = 1, \ldots, N$ é estável na média quadrática. Portanto, da proposição 2.6 tem-se que existe $X = (X_1, \ldots, X_N) > 0$ em $\mathbb{H}^{n+}$ tal que

$$-X_i + (A_i + H_i M_i)^* \mathcal{E}_i(X)(A_i + H_i M_i) < 0 \text{ para } i = 1, \ldots, N. \tag{2.8}$$

Tomando $W_{i1} = X_i > 0$, $Z_i = \mathcal{E}_i(X)$, $W_{i2} = Z_i H_i$ e $W_{i3} = W_{i2}^* Z_i^{-1} W_{i2}$ para $i = 1, \ldots, N$, verifica-se facilmente a partir de (2.8) que as equações (2.7.a)-(2.7.e) são satisfeitas e que $\Phi \neq \emptyset$.

(suficiência): Assuma que $\Phi \neq \emptyset$, isto é, que existem $W_1, Z \in \mathbb{H}^{n+}$, $W_2 \in \mathbb{H}^{s,n}$, $W_3 \in \mathbb{H}^{s+}$ satisfazendo (2.7.a)-(2.7.e). Seja então $H = (H_1, \ldots, H_N) \in \mathbb{H}^{s,n}$ dado por $H_i = Z_i^{-1} W_{i2}$ para $i = 1, \ldots, N$. Tem-se que para $i = 1, \ldots, N$,

$$\begin{aligned}
&(A_i + H_i M_i)^* \mathcal{E}_i(W_1)(A_i + H_i M_i) - W_i = \\
&= (A_i + Z_i^{-1} W_{i2} M_i)^* \mathcal{E}_i(W_1)(A_i + Z_i^{-1} W_{i2} M_i) - W_i \\
&\leq A_i^* \mathcal{E}_i(W_1) A_i + M_i^* W_{i2}^* Z_i^{-1} \mathcal{E}_i(W_1) A_i + A_i^* \mathcal{E}_i(W_1) Z_i^{-1} W_{i2} M_i \\
&\qquad + M_i^* W_{i2}^* Z_i^{-1} \mathcal{E}_i(W_1) Z_i^{-1} W_{i2} M_i - W_{i1} \\
&\leq A_i^* Z_i A_i + M_i^* W_{i2}^* A_i + A_i^* W_{i2} M_i + M_i^* W_{i3} M_i < 0.
\end{aligned}$$

Portanto, da proposição 2.6, tem-se que H estabiliza (M, A) na média quadrática e que (M, A) é detectável na média quadrática.

□

Comentário 2.10: A proposição 2.8 apresenta condições para a estabilizabilidade de um dado sistema que podem ser testadas com relativa facilidade, através de métodos numéricos. O teste consiste apenas em encontrar $W_1, Z \in \mathbb{G}^{n+}$, $W_2 \in \mathbb{G}^{m,n}$, $W_3 \in \mathbb{G}^{m+}$ tais que

$$\sum_{i=1}^{N} p_{ij}(A_i W_{i1} A_i^* + B_i W_{i2}^* A_i^* + A_i W_{i2} B_i^* + B_i W_{i3} B_i^*) - W_{j1} < 0 \tag{2.9.a}$$

$$\begin{bmatrix} W_{j1} & W_{j2} \\ W_{j2}^* & W_{j3} \end{bmatrix} \geq 0 \tag{2.9.b}$$

$$W_{j1} > 0 \tag{2.9.c}$$

para $j = 1, \ldots, N$. Caso exista solução para o problema de programação convexa dado por (2.9.a)-(2.9.c), o sistama será estabilizável, caso contrário o sistema será não-estabilizável. Note que a equação (2.9.b) é equivalente por complemento de Schur (veja o comentário 2.1) a (2.5.b).

Comentário 2.11: Como no caso anterior, para se estabelecer a detectabilidade de um dado sistema, basta resolver o seguinte problema convexo: encontrar $W_1, Z \in \mathbb{G}^{n+}$, $W_2 \in \mathbb{G}^{s,n}$, $W_3 \in \mathbb{G}^{s+}$ tais que

$$A_i^* Z_i A_i + M_i^* W_{i2}^* A_i + A_i^* W_{i2} M_i + M_i^* W_{i3} M_i - W_{i1} < 0 \tag{2.10.a}$$

$$\begin{bmatrix} Z_i & W_{i2} \\ W_{i2}^* & W_{i3} \end{bmatrix} \geq 0 \tag{2.10.b}$$

$$Z_i \geq \mathcal{E}_i(W_1) \tag{2.10.c}$$

$$W_{i1} > 0 \tag{2.10.d}$$

$$Z_i > 0\,. \tag{2.10.e}$$

para $i = 1, \ldots, N$. Caso exista solução o sistema será detectável. Caso contrário será não-detectável.

Comentário 2.12: Para os problemas de programação convexa descritos acima, as matrizes envolvidas foram assumidas reais, apesar do desenvolvimento todo ter sido feito para matrizes complexas. Isso é feito basicamente porque os métodos de resolução de problemas de programação convexa tratam normalmente apenas matrizes reais. Note porém que isso não significa perda de generalidade, pois uma desigualdade complexa pode ser transformada em uma desigualdade real com o uso do seguinte fato (veja [27]):

$$W \geq 0 \Leftrightarrow \begin{bmatrix} \mathrm{Re}\{W\} & \mathrm{Im}\{W\} \\ -\mathrm{Im}\{W\} & \mathrm{Re}\{W\} \end{bmatrix} \geq 0$$

qualquer que seja $W \in \mathbb{H}^{n+}$.

Por uma questão de simplicidade, todos os problemas de programação convexa apresentados neste trabalho estarão restritos a variáveis reais. Para se lidar com problemas complexos, basta ter em mente o comentário acima.

2.3.4 Um importante resultado auxiliar

O lema 2.13, apresentado a seguir, é de crucial importância para o desenvolvimento dos capítulos 3, 4 e 5 deste trabalho.

Sejam $A = (A_1, \ldots, A_N) \in \mathbb{H}^n$, $B = (B_1, \ldots, B_N) \in \mathbb{H}^{m,n}$ e $F = (F_1, \ldots, F_N) \in \mathbb{H}^{n,m}$ tais que $\Gamma_i = A_i + B_i F_i$ para $i = 1, \ldots, N$. Define-se o operador $\overline{\mathcal{L}}(.) = (\overline{\mathcal{L}}_1(.), \ldots, \overline{\mathcal{L}}_N(.)) \in \mathbb{B}(\mathbb{H}^n)$ como

$$\overline{\mathcal{L}}_i(.) := \Lambda_i^* \mathcal{E}_i(.) \Lambda_i \text{ para } i = 1, \ldots, N, \tag{2.11}$$

onde $\Lambda_i = A_i + B_i G_i$, $i = 1, \ldots, N$, para alguma $G = (G_1, \ldots, G_N) \in \mathbb{H}^{n,m}$.

Lema 2.13: Sejam $\mathcal{L}$ e $\overline{\mathcal{L}}$ como definidos em (2.4.b) e (2.11) acima, com $\Gamma_i = A_i + B_i F_i$ e $\Lambda_i = A_i + B_i G_i$ para $i = 1, \ldots, N$. Suponha que $r_\sigma(\mathcal{L}) < 1$ e que para alguma $X = (X_1, \ldots, X_N) \in \mathbb{H}^{n+}$ e algum $\delta > 0$,

$$X_i - \Lambda_i^* \mathcal{E}_i(X) \Lambda_i \geq \delta (G_i - F_i)^* (G_i - F_i) \text{ para } i = 1, \ldots, N. \tag{2.12}$$

Então $r_\sigma(\overline{\mathcal{L}}) < 1$.

Prova: Seja $\overline{\mathcal{T}} = \overline{\mathcal{L}}^*$, de forma que para qualquer $V = (V_1, \ldots, V_N) \in \mathbb{H}^n$,

$$\overline{\mathcal{T}}_j(V) := \sum_{i=1}^{N} p_{ij} \Lambda_i V_i \Lambda_i^* \text{ para } j = 1, \ldots, N. \tag{2.13}$$

Note que para $\varepsilon > 0$ arbitrário e $V = (V_1, \ldots, V_N) \geq 0$,

$$\begin{aligned} 0 \leq & \left[\varepsilon(A_i + B_i F_i) - \frac{1}{\varepsilon} B_i (G_i - F_i)\right] V_i \left[\varepsilon(A_i + B_i F_i) - \frac{1}{\varepsilon} B_i (G_i - F_i)\right]^* \\ = & \ \varepsilon^2 (A_i + B_i F_i) V_i (A_i + B_i F_i)^* + \frac{1}{\varepsilon^2} B_i (G_i - F_i) V_i (G_i - F_i)^* B_i^* - \\ & (A_i + B_i F_i) V_i (G_i - F_i)^* B_i^* - B_i (G_i - F_i) V_i (A_i + B_i F_i)^* \end{aligned}$$

para $i = 1, \ldots, N$, de forma que

$$(A_i + B_iF_i)V_i(G_i - F_i)^*B_i^* + B_i(G_i - F_i)V_i(A_i + B_iF_i)^*$$
$$\leq \varepsilon^2(A_i + B_iF_i)V_i(A_i + B_iF_i)^* + \frac{1}{\varepsilon^2}B_i(G_i - F_i)V_i(G_i - F_i)^*B_i^* \qquad (2.14)$$

para $i = 1, \ldots, N$. Combinando (2.13) e (2.14), tem-se que

$$0 \leq \overline{\mathcal{T}}_j(V) = \sum_{i=1}^{N} p_{ij}\Lambda_i V_i \Lambda_i^*$$

$$= \sum_{i=1}^{N} p_{ij}(A_i + B_iF_i + B_i(G_i - F_i))V_i(A_i + B_iF_i + B_i(G_i - F_i))^*$$

$$= \sum_{i=1}^{N} p_{ij}[(A_i + B_iF_i)V_i(A_i + B_iF_i)^* + B_i(G_i - F_i)V_i(A_i + B_iF_i)^* +$$

$$(A_i + B_iF_i)V_i(G_i - F_i)^*B_i^* + B_i(G_i - F_i)V_i(G_i - F_i)^*B_i^*]$$

$$\leq (1 + \varepsilon^2)\sum_{i=1}^{N} p_{ij}(A_i + B_iF_i)V_i(A_i + B_iF_i)^* +$$

$$\left(1 + \frac{1}{\varepsilon^2}\right)\sum_{i=1}^{N} p_{ij}B_i(G_i - F_i)V_i(G_i - F_i)^*B_i^*$$

$$= (1 + \varepsilon^2)\mathcal{T}_j(V) + \left(1 + \frac{1}{\varepsilon^2}\right)\mathcal{Q}_j(V), \qquad (2.15)$$

para $j = 1, \ldots, N$, onde $\mathcal{Q}(V) = (\mathcal{Q}_1(V), \ldots, \mathcal{Q}_N(V)) \in \mathbb{B}(\mathbb{H}^n)$ é definido como

$$\mathcal{Q}_j(V) := \sum_{i=1}^{N} p_{ij}B_i(G_i - F_i)V_i(G_i - F_i)^*B_i^* \qquad (2.16)$$

e $\mathcal{T} = \mathcal{L}^*$. Como $r_\sigma(\mathcal{T}) = r_\sigma(\mathcal{L}) < 1$ por hipótese, é possível escolher $\varepsilon > 0$ tal que $r_\sigma(\tilde{\mathcal{T}}) < 1$, onde $\tilde{\mathcal{T}}(.) := (1 + \varepsilon^2)\mathcal{T}(.)$

Defina para $t = 0, 1, \ldots$ as sequências

$$Y(t+1) = \overline{\mathcal{T}}(Y(t)), \qquad Y(0) \in \mathbb{H}^{n+}$$
$$Z(t+1) = \tilde{\mathcal{T}}(Z(t)) + \tilde{\mathcal{Q}}(Y(t)), \qquad Z(0) = Y(0) \qquad (2.17)$$

onde $\tilde{\mathcal{Q}}(.) := \left(1 + \frac{1}{\varepsilon^2}\right)\mathcal{Q}(.)$. Então para $t = 0, 1, 2, \ldots$

$$Z(t) \geq Y(t) \geq 0\,. \tag{2.18}$$

(2.18) é imediata de (2.17) para $t = 0$. Suponha por indução que (2.18) seja valida para t. Então, de (2.15)

$$\begin{aligned} Z(t+1) &= \tilde{\mathcal{T}}(Z(t)) + \tilde{\mathcal{Q}}(Y(t)) \\ &= (1+\varepsilon^2)\mathcal{T}(Z(t)) + \left(1+\frac{1}{\varepsilon^2}\right)\mathcal{Q}(Y(t)) \\ &\geq (1+\varepsilon^2)\mathcal{T}(Y(t)) + \left(1+\frac{1}{\varepsilon^2}\right)\mathcal{Q}(Y(t)) \\ &\geq \bar{\mathcal{T}}(Y(t)) \\ &= Y(t+1), \end{aligned}$$

o que mostra o resultado para $t+1$. De (2.17) tem-se que

$$Z(t) = \tilde{\mathcal{T}}^t(Y(0)) + \sum_{s=0}^{t-1} \tilde{\mathcal{T}}^{t-1-s}(\tilde{\mathcal{Q}}(Y(s)))\,.$$

Tomando-se a norma 1 (veja equação (2.1.a)) da expressão acima, tem-se que

$$\|Z(t)\|_1 = \left\|\tilde{\mathcal{T}}^t\right\| \|Y(0)\|_1 + \sum_{s=0}^{t-1} \left\|\tilde{\mathcal{T}}^{t-1-s}\right\| \left\|\tilde{\mathcal{Q}}(Y(s))\right\|_1\,.$$

Como $r_\sigma(\tilde{\mathcal{T}}) < 1$, é possível encontrar $a > 0$, $b < 1$ tais que

$$\left\|\tilde{\mathcal{T}}^s\right\| \leq ab^s \text{ para } s = 0, 1, \ldots$$

e portanto

$$\|Z(t)\|_1 \leq ab^t \|Y(0)\|_1 + a\sum_{s=0}^{t-1} b^{t-1-s} \left\|\tilde{\mathcal{Q}}(Y(s))\right\|_1\,.$$

Suponha por agora que $\sum_{s=0}^{\infty} \left\|\tilde{\mathcal{Q}}(Y(s))\right\|_1 < \infty$. Então

$$\begin{aligned} \sum_{t=0}^{\infty} \|Z(t)\|_1 &\leq \frac{a}{1-b}\|Y(0)\|_1 + a\sum_{t=0}^{\infty}\sum_{s=0}^{t-1} b^{t-1-s}\left\|\tilde{\mathcal{Q}}(Y(s))\right\|_1 \\ &= \frac{a}{1-b}\|Y(0)\|_1 + \frac{a}{1-b}\sum_{s=0}^{\infty}\left\|\tilde{\mathcal{Q}}(Y(s))\right\|_1 < \infty. \end{aligned}$$

e portanto, de (2.17) e (2.18), para qualquer $Y(0) = (Y_1(0), \ldots, Y_N(0)) \geq 0$,

$$0 \leq \sum_{t=0}^{\infty} \left\| \bar{\mathcal{T}}^t(Y(0)) \right\|_1 = \sum_{t=0}^{\infty} \|Y(t)\|_1 \leq \sum_{t=0}^{\infty} \|Z(t)\|_1 < \infty .$$

Como mencionado no comentário 2.2, qualquer $S = (S_1, \ldots, S_N) \in \mathbb{H}^n$ pode ser decomposto como $S = (S^1 - S^2) + \sqrt{-1}(S^3 - S^4)$, portanto

$$\sum_{t=0}^{\infty} \left\| \bar{\mathcal{T}}^t(S) \right\|_1 \leq \sum_{i=1}^{4} \sum_{t=0}^{\infty} \left\| \bar{\mathcal{T}}^t(S^i) \right\|_1 < \infty ,$$

o que mostra (veja [38]) que $r_\sigma(\bar{\mathcal{T}}) = r_\sigma(\bar{\mathcal{L}}) < 1$.
Falta ainda provar que $\sum_{s=0}^{\infty} \left\| \tilde{\mathcal{Q}}(Y(s)) \right\|_1 < \infty$. Porém de (2.16), fazendo

$$c_0 = \left[\left(1 + \frac{1}{\varepsilon^2} \right) \max_i \|B_i\|^2 \right] \frac{N}{\delta} \text{ para } i = 1, \ldots, N ,$$

tem-se que

$$\begin{aligned}
\left\| \tilde{\mathcal{Q}}(Y(s)) \right\|_1 &= \sum_{j=1}^{N} \left\| \tilde{\mathcal{Q}}_j(Y(s)) \right\|_1 \\
&\leq \left(1 + \frac{1}{\varepsilon^2} \right) \sum_{i=1}^{N} \sum_{j=1}^{N} p_{ij} \left\| B_i (G_i - F_i) Y_i(s) (G_i - F_i)^* B_i^* \right\| \\
&\leq \left(1 + \frac{1}{\varepsilon^2} \right) \sum_{i=1}^{N} \sum_{j=1}^{N} p_{ij} \|B_i\|^2 \operatorname{tr}\{ (G_i - F_i) Y_i(s) (G_i - F_i)^* \} \\
&\leq \left[\left(1 + \frac{1}{\varepsilon^2} \right) \max_i \|B_i\|^2 \right] N \sum_{i=1}^{N} \operatorname{tr}\{ Y_i(s)^{1/2} (G_i - F_i)^* (G_i - F_i) Y_i(s)^{1/2} \} \\
&\leq c_0 \sum_{i=1}^{N} \operatorname{tr}\{ Y_i(s) (X_i - \mathcal{L}_i(X)) \} \\
&= c_0 \langle Y(s); (X - \mathcal{L}(X)) \rangle = c_0 \{ \langle Y(s); X \rangle - \langle Y(s); \mathcal{L}(X) \rangle \} \\
&= c_0 \{ \langle Y(s); X \rangle - \langle \mathcal{T}(Y(s)); X \rangle \} = c_0 \{ \langle Y(s); X \rangle - \langle Y(s+1); X \rangle \} .
\end{aligned}$$

Tomando a soma de $s = 0$ até r, tem-se que

$$\sum_{s=0}^{r} \left\| \tilde{\mathcal{Q}}(Y(s)) \right\|_1 \leq c_0 \{ \langle Y(0); X \rangle - \langle Y(r+1); X \rangle \} \leq c_0 \langle Y(0); X \rangle$$

pois $X = (X_1, \ldots, X_N) \geq 0$ e portanto $\langle Y(r+1); X \rangle \geq 0$. Tomando o limite quando $r \to \infty$, obtém-se o resultado desejado.

□

2.4 Normas especiais

Considere novamente em $(\Omega, \{\Im_k\}, \Im, \mathcal{P})$ o seguinte sistema:

$$\Sigma = \begin{cases} x(k+1) = \Gamma_{\theta(k)} x(k) + D_{\theta(k)} \omega(k) \\ z(k) = \Phi_{\theta(k)} x(k) \\ x(0) = x_0 \qquad \theta(0) = \theta_0 \end{cases} \tag{2.19}$$

com $D = (D_1, \ldots, D_N) \in \mathbb{H}^{r,n}$ $\Phi = (\Phi_1, \ldots, \Phi_N) \in \mathbb{H}^{n,s}$ e $D_i D_i^* > 0$ para $i = 1, \ldots, N$.

Suponha que $r_\sigma(\mathcal{T}) < 1$, ou seja, que o sistema dado por (2.2) seja estável na média quadrática, e que a sequência $\omega = (\omega(0), \ldots)$ seja um impulso unitário. Da proposição 2.6, item (v), tem-se que $z = (z(0), z(1), \ldots) \in l_2^s$.

Segue-se na próxima subseção uma generalização da norma H_2 para sistemas discretos determinísticos para o caso Markoviano.

2.4.1 A norma H_2

Definição 2.14: Define-se a norma H_2 do sistema Σ (dado por (2.19)) como

$$\|\Sigma\|_2^2 := \sum_{s=1}^{r} \sum_{j=1}^{N} \|z_{s,j}\|_2^2$$

onde $z_{s,j}$ representa a sequência de saída $(z(0), z(1), \ldots)$ quando

i) a sequência de entrada é dada por $\omega = (\omega(0), \omega(1), \ldots)$, $\omega(0) = e_s$, $\omega(k) = 0$ para $k > 0$ com $e_s \in \mathbb{R}^r$ o vetor unitário formado por 1 na s-ésima posição e zero nas componentes restantes, e

ii) $\theta(0) = \theta(1) = j$.

Uma forma alternativa para se definir a norma H_2 para sistemas markovianos (e talvez mais intuitiva do que a condição (ii) da definição acima) seria considerar, para o sistema livre de entradas,

condições iniciais dadas por $x(0) = D_j e_s$ e $\theta(0) = j$ para $j = 1, \ldots, N$, $s = 1, \ldots, r$.

Para o caso determinístico (com $N = 1$ e $P = 1$), a definição acima se reduz à norma H_2 usual. Como no caso determinístico, tem-se que a norma H_2 como definida acima pode ser calculada como a solução dos gramianos[1] discretos acoplados de observabilidade e controlabilidade. Para tanto define-se $\mathcal{O} = (\Phi_1^* \Phi_1, \ldots, \Phi_N^* \Phi_N) \in \mathbb{H}^{n+}$ e $\mathcal{C} = (D_1 D_1^*, \ldots, D_N D_N^*) \in \mathbb{H}^{n+}$, bem como $So = (So_1, \ldots, So_N) \in \mathbb{H}^{n+}$ e $Sc = (Sc_1, \ldots, Sc_N) \in \mathbb{H}^{n+}$, as únicas soluções das equações

$$So = \mathcal{L}(So) + \mathcal{O} \qquad \text{(gramiano de observabilidade)} \tag{2.20}$$

$$Sc = \mathcal{T}(Sc) + \mathcal{C} \qquad \text{(gramiano de controlabilidade).} \tag{2.21}$$

A norma H_2 será abordada novamente num diferente contexto no capítulo 4.

O próximo resultado apresenta uma caracterização da norma H_2 em termos da solução dos gramianos de observabilidade e controlabilidade.

Proposição 2.15: (provada em [19]) $\|\Sigma\|_2^2 = \sum_{j=1}^{N} \text{tr}\,\{D_j^* So_j D_j\} = \sum_{j=1}^{N} \text{tr}\,\{\Gamma_j Sc_j \Gamma_j^*\}$.

2.4.2 A norma H_∞

Considere novamente o sistema Σ como em (2.19) com $\omega = (\omega(0), \ldots) \in l_2^r$.

Proposição 2.16: (provada em [17]) $r_\sigma(\mathcal{T}) < 1$ se e somente se $x = (0, x(1), \ldots) \in l_2^n$ para qualquer $\omega = (\omega(0), \omega(1), \ldots) \in l_2^r$.

O resultado acima mostra que, se $r_\sigma(\mathcal{T}) < 1$, os operadores $\mathcal{X}(\theta_0, \omega) = (0, x(1), \ldots)$ e $\mathcal{Z}(\theta_0, \omega) = (0, z(1), \ldots)$ para $\omega \in l_2^r$, onde $x(k)$ e $z(k)$ são dados por (2.19), são operadores lineares limitados, isto é, $\mathcal{X}(\theta_0, .) \in \mathbb{B}(l_2^r, l_2^n)$ e $\mathcal{Z}(\theta_0, .) \in \mathbb{B}(l_2^r, l_2^s)$. Dessa forma pode-se definir a norma induzida como

$$\left\|\mathcal{Z}(\theta_0, .)\right\| := \sup_{\omega \in l_2^r} \frac{\left\|\mathcal{Z}(\theta_0, \omega)\right\|_2}{\|\omega\|_2}.$$

Suponha que $r_\sigma(\mathcal{T}) < 1$. Daí $x = (0, x(1), \ldots) \in l_2^n$ e portanto $z = (0, z(1), \ldots) \in l_2^s$. Define-se então a norma H_∞ para o sistema Σ.

Definição 2.17: Define-se a norma H_∞ do sistema Σ (dado por (2.19)) como

$$\|\Sigma\|_\infty^2 := \sup_{\theta_0} \sup_{\omega \in l_2^r} \left\|\mathcal{Z}(\theta_0, .)\right\|.$$

Como anteriormente, para o caso determinístico ($N = 1$ e $P = 1$) a definição acima se resume

[1] Em homenagem a **Jorgen Pedersen Gram** (1850-1916). Matemático dinamarquês.

à norma H_∞ usual no domínio do tempo.