

Kontrolltheorie

R. Altmann
TU Berlin

Sommersemester 2017
(Version vom 17. Juli 2017)

Dieses Skript basiert auf der Vorlesung 'Kontrolltheorie' von Dr. Christian Schröder, die im Sommersemester 2015 an der TU Berlin gehalten wurde. Die ursprüngliche \LaTeX Mitschrift stammt von Dennis Wenzel und wurde nun vollständig überarbeitet.

Inhaltsverzeichnis

1	Einführung und Beispiele	1
1.1	Mathematische Grundlagen	3
1.1.1	Gewöhnliche Differentialgleichungen	3
1.1.2	Die Laplace Transformation	7
1.1.3	Die Singulärwertzerlegung	8
1.1.4	Invariante Unterräume	9
1.2	Beispiel: Invertiertes Pendel	10
1.3	Beispiel: Parabolantenne	12
2	Lineare Kontrollsysteme	13
2.1	Kontrollprobleme im Frequenzbereich	15
2.2	Steuerbarkeit	17
2.2.1	Zeitvarianter Fall	18
2.2.2	Zeitinvarianter Fall	22
2.3	Stabilisierbarkeit	29
2.4	Rekonstruierbarkeit und Beobachtbarkeit	32
2.5	Der Staircase Algorithmus	38
3	Stabilisierung und Lyapunov-Gleichungen	44
3.1	Sylvester- und Lyapunov-Gleichungen	45
3.2	Stabilisierung mit Lyapunov-Gleichungen	50
3.3	Numerische Lösung von Lyapunov-Gleichungen	53
3.3.1	Kronecker-Produkt	54
3.3.2	Bartels-Stewart	55
3.3.3	Matrix Signum Methode	57
3.3.4	Hammarling-Verfahren	60
3.3.5	ADI-Iteration	60
3.4	Teilstabilisierung	61
3.5	Polvorgabe	62
3.5.1	Der Polvorgabe-Algorithmus von Miminis und Paige	68
3.5.2	Die Ackermann Formel	71
4	LQ-Optimalsteuerung	71
4.1	Die algebraische Riccati-Gleichung (ARE)	76
4.2	Newton-Methode zur Lösung der ARE	84
4.3	Signum Methode zur Lösung der ARE	88

1 Einführung und Beispiele

Ziel der Kontrolltheorie ist die Steuerung eines dynamischen Systems, sodass eine gewollte oder sogar optimierte Verhaltensweise erreicht wird. Dabei kann man nach verschiedenen Strategien optimieren:

- Zeit-optimale Steuerung ([time-optimal](#)),
- Energie-optimale Steuerung ([energy-optimal](#)),
- folgen bzw. minimale Abweichung von einer Ziel-Trajektorie ([path-following](#)),
- Mischungen aller Strategien.

Wir betrachten das stark vereinfachte Beispiel der Kontrolle eines Zuges. Dabei wird der Zug als Massepunkt modelliert. Die zurückgelegte Strecke zur Zeit t bezeichnen wir mit $s(t)$, die Geschwindigkeit mit $v(t)$. Um das ganze als System schreiben zu können, definieren wir den Zustandsvektor x und den Anfangszustand x_0 zur Zeit t_0 durch

$$x(t) = \begin{bmatrix} s(t) \\ v(t) \end{bmatrix}, \quad x_0 = x(t_0) = \begin{bmatrix} s_0 \\ v_0 \end{bmatrix}.$$

Zudem kann man Einschränkungen festlegen wie eine maximale Geschwindigkeit auf einer Teilstrecke $v_{\max}(t)$. Dann soll also $v(t) \in [0, v_{\max}(t)]$ gelten für alle t . Aber auch andere Einschränkungen zum Beispiel durch den Fahrplan sind denkbar. Wir nehmen an, dass wir mit der Kontrolle den Zustand durch Bremsen oder Beschleunigen durch eine Kraft $F(t)$ beeinflussen können,

$$u(t) = F(t) = m\ddot{s}(t)$$

Zudem gibt es noch Kräfte, die nicht vom Zugführer beeinflusst werden können wie zum Beispiel Reibung durch Luftwiderstand oder Gravitationskräfte. Unter Vernachlässigung dieser Kräfte ergibt sich das dynamische System

$$\dot{x}(t) = \begin{bmatrix} \dot{s}(t) \\ \dot{v}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u(t) =: Ax(t) + Bu(t)$$

mit Anfangsbedingung $x(t_0) = x_0$. Weitere mögliche Beschränkungen sind maximale Brems- bzw. Beschleunigungskraft u_{\min} , u_{\max} oder eine maximale Moterleistung P_{\max} . In diesem Fal wäre $u_{\max}(v) = P_{\max}/v$.

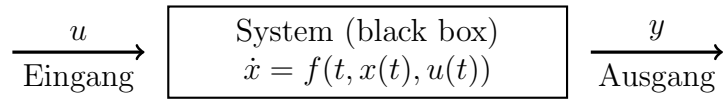


Abbildung 1.1: Schema eines Kontrollsystems mit Eingang/Kontrolle (input) u und Ausgang (output) y .

Da man gewöhnlich nicht alle Zustandsgrößen x messen kann, ist der Zustand meist nur teilweise bekannt. In diesem Beispiel wäre es denkbar, dass nur die Position des Zuges beim Passieren von Sensoren gemessen wird, d.h. der sogenannte Ausgang (output) ist gegeben durch

$$y(t) = [1 \quad 0] x(t) =: Cx(t).$$

Allgemein betrachten wir in dieser Vorlesung Kontrollsysteme der Form

$$\begin{aligned} \dot{x}(t) &= f(t, x(t), u(t)), & x(t_0) &= x_0, \\ y(t) &= g(t, x(t), u(t)). \end{aligned}$$

Die erste Gleichung heißt *Zustandsgleichung*, die zweite *Outputgleichung*. Dabei bezeichnen wir die einzelnen Variablen als

$$\begin{array}{ll} \text{Zustand} & x: [t_0, \infty) \rightarrow \mathfrak{X} \subseteq \mathbb{R}^n \quad (\text{state}), \\ \text{Kontrolle/Input} & u: [t_0, \infty) \rightarrow \mathfrak{U} \subseteq \mathbb{R}^m \quad (\text{input}), \\ \text{Ausgang/Output} & y: [t_0, \infty) \rightarrow \mathfrak{Y} \subseteq \mathbb{R}^p \quad (\text{output}). \end{array}$$

Entsprechend heißen \mathfrak{X} , \mathfrak{U} und \mathfrak{Y} Zustands-, Kontroll- und Outputraum. Ein schematische Darstellung so eines Systems ist in Abbildung 1 zu finden. In vielen Anwendungen sind x , \dot{x} nicht bekannt oder von sehr hoher Dimension aber die Relation von Input zu Output kann gemessen bzw. angegeben werden.

Es gibt auch diskrete Kontrollsysteme, die wir hier allerdings nicht weiter betrachten werden. Ein diskretes lineares Steuerungsproblem hat die Form

$$x_{k+1} = A_k x_k + B_k u_k, \quad x_0 \text{ gegeben}, \quad (1.1a)$$

$$y_k = C_k x_k + D_k u_k. \quad (1.1b)$$

Beispiel 1.1. *Die Inflationsrate eines Landes wird nur in monatlichen Raten bestimmt, also nur zu diskreten Zeitpunkten. Die Zentralbank versucht diese*

durch den Leitzins zu steuern und im allgemeinen niedrig zu halten. Wir bezeichnen mit x_k die Inflationsrate im Monat k und u_k den Leitzins. Dann ergibt sich ein System der Form

$$x_{k+1} = f(x_k, u_k).$$

Dies ist ein diskretes Kontrollproblem mit dem man die Inflationsrate steuern kann.

1.1 Mathematische Grundlagen

In diesem Abschnitt fassen wir einige Ergebnisse zusammen, die im Laufe der Vorlesung gebraucht werden.

1.1.1 Gewöhnliche Differentialgleichungen

Wir wiederholen zunächst die Begriffe der Stabilität und asymptotischen Stabilität.

Definition 1.2. Eine Lösung $x: [t_0, \infty) \rightarrow \mathbb{R}^n$ der gewöhnlichen DGL

$$\dot{x}(t) = A(t)x(t)$$

heißt stabil, falls es eine Konstante $M > 0$ gibt, sodass $\|x(t)\| \leq M$ für alle $t \in [t_0, \infty)$ gilt. Gilt zusätzlich noch

$$\lim_{t \rightarrow \infty} x(t) = 0,$$

so heißt x asymptotisch stabil.

Im Fall von linearen und zeitinvarianten Differentialgleichungen kann man die Stabilität direkt von den Eigenwerten der Systemmatrix A ablesen.

Satz 1.3. Sei $A \in \mathbb{R}^{n,n}$ und $\Lambda(A)$ die Menge der Eigenwerte. Dann ist jede Lösung der Differentialgleichung $\dot{x} = Ax$

- i) asymptotisch stabil, falls $\Re(\lambda) < 0$ für alle $\lambda \in \Lambda(A)$, und
- ii) stabil, falls $\Re(\lambda) \leq 0$ für alle $\lambda \in \Lambda(A)$ sowie für alle $\lambda \in \Lambda(A)$ mit $\Re(\lambda) = 0$ die algebraische mit der geometrischen Vielfachheit übereinstimmt.

Beweis. Wir geben nur eine Beweisskizze. Zu einem Anfangswert $x(0) = x_0$ ist die Lösung des Systems gegeben durch

$$x(t) = \exp(At)x_0.$$

Mit der Jordanschen Normalform $S^{-1}AS = J$ erhält man dann

$$x(t) = SS^{-1} \exp(At)SS^{-1}x_0 = S \exp(Jt)S^{-1}x_0.$$

Ohne Beschränkung der Allgemeinheit können wir also annehmen, dass A ein Jordan-Block ist und damit den Beweis fortsetzen. \square

Definition 1.4. Eine Matrix mit der Eigenschaft, dass alle Eigenwerte negativen Realteil haben, heißt Hurwitz. Manchmal nennt man so eine Matrix auch asymptotisch stabil.

Bemerkung 1.5. Man beachte, dass die Bedingung $\Lambda(A(t)) \subseteq \mathbb{C}^-$ im zeitvarianten Fall nicht hinreichend ist für asymptotische Stabilität. Für nichtautonome Systeme folgt dies nämlich nicht aus der punktweisen Stabilität von $A(t)$.

Bemerkung 1.6. In der linearen Kontrolltheorie werden wir Systeme der Form

$$\dot{x}(t) = Ax(t) + Bu(t)$$

betrachten. Gesucht ist dann zum Beispiel eine lineare Zustandsrückführung (*state-feedback*), d.h. ein u der Form $u(t) = -Fx(t)$, sodass das System asymptotisch stabil ist. Dementsprechend muss F so konstruiert werden, dass $A - BF$ Hurwitz ist. Es soll also gelten

$$\Lambda(A - BF) \subseteq \mathbb{C}^- = \{z \in \mathbb{C} \mid \Re(z) < 0\}.$$

Beispiel 1.7. Wir betrachten die Matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 - \alpha & -\beta \end{bmatrix}.$$

Im Fall $\alpha = \beta = 0$ hat die Matrix die Eigenwerte $\lambda_{1,2} = \pm 1$. Demnach ist das System instabil, da es einen positiven Eigenwert gibt. Für $\beta = 0$ hat die Matrix die Eigenwerte $\lambda_{1,2} = \pm\sqrt{1 - \alpha}$. Das System ist also instabil für

$\alpha \leq 1$ und stabil für $\alpha > 1$. Da der Realteil aber nicht negativ ist, ist das System nicht asymptotisch stabil. Im allgemeinen Fall hat A die Eigenwerte

$$\lambda_{1,2} = -\frac{\beta}{2} \pm \frac{1}{2}\sqrt{\beta^2 + 4(1 - \alpha)}.$$

Das System ist also asymptotisch stabil für $\alpha > 1$ und $\beta > 0$.

Im Folgenden bezeichne $\Phi(t, s)$ die *Fundamentallösung* des homogenen linearen Systems

$$\dot{x}(t) = A(t)x(t).$$

Dies bedeutet, dass Φ die eindeutige (Matrix-)Lösung des Anfangswertproblems

$$\frac{d}{dt}\Phi(t, s) = A(t)\Phi(t, s), \quad \Phi(s, s) = I_n$$

ist.

Lemma 1.8. Die Fundamentallösung $\Phi(t, s)$ erfüllt für alle $r, s, t \in \mathbb{R}$ die folgenden Eigenschaften:

- i) Es gilt die Halbgruppen-Eigenschaft $\Phi(t, s) = \Phi(t, r)\Phi(r, s)$.
- ii) $\Phi(t, s)$ ist invertierbar mit $\Phi(t, s)^{-1} = \Phi(s, t)$.
- iii) Es gilt $\frac{d}{ds}\Phi(t, s) = -\Phi(t, s)A(s)$.
- iv) $x(t) = \Phi(t, t_0)x_0$ ist die eindeutige Lösung des Anfangswertproblems

$$\dot{x}(t) = A(t)x(t), \quad x(t_0) = x_0.$$

Beweis. Übungsaufgabe 1.1a □

Satz 1.9. Gegeben sei das System

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t), & x(t_0) &= x_0, \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned}$$

und die Fundamentallösung Φ des homogenen Problems, d.h., $\frac{d}{dt}\Phi(t, s) = A(t)\Phi(t, s)$ mit $\Phi(s, s) = I_n$. Dann ist x gegeben durch

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^t \Phi(t, s)B(s)u(s) ds$$

und y durch

$$y(t) = C(t)\Phi(t, t_0)x_0 + C(t) \int_{t_0}^t \Phi(t, s)B(s)u(s) ds + D(t)u(t).$$

Beweis. Der Satz ist ein Spezialfall der Formel von Duhamel. Die Leibniz-Regel für die Ableitung von Integralen besagt

$$\frac{d}{dt} \int_{t_0}^t f(t, s) ds = f(t, t) + \int_{t_0}^t \frac{d}{dt} f(t, s) ds.$$

Damit leiten wir die gegebene Funktion x ab und zeigen, dass es die Differentialgleichung erfüllt.

$$\begin{aligned} \frac{d}{dt}x(t) &= \frac{d}{dt}\Phi(t, t_0)x_0 + \frac{d}{dt} \int_{t_0}^t \Phi(t, s)B(s)u(s) ds \\ &= A(t)\Phi(t, t_0)x_0 + \Phi(t, t)B(t)u(t) + \int_{t_0}^t A(t)\Phi(t, s)B(s)u(s) ds \\ &= A(t)x(t) + B(t)u(t). \end{aligned}$$

Offensichtlich erfüllt y dann auch $y(t) = C(t)x(t) + D(t)u(t)$. □

Korollar 1.10. *Im zeitinvarianten Fall, d.h. $A(t) \equiv A$, $B(t) \equiv B$, $C(t) \equiv C$ und $D(t) \equiv D$, gilt $\Phi(t, s) = e^{A(t-s)}$ und somit für $t_0 = 0$,*

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s) ds = e^{At} \left(x_0 + \int_0^t e^{-As}Bu(s) ds \right).$$

Bemerkung 1.11. *Sei $\mathcal{U}_{ad} \subseteq PC([t_0, \infty), \mathbb{R}^m)$ der Raum der zulässigen Kontrollen. Dann heißt die Abbildung*

$$\mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{U}_{ad} \rightarrow \mathbb{R}^n, \quad (t_0, x_0, t, u) \mapsto x(t)$$

Zustandsübergangsfunktion (*input-state mapping*). Die Abbildung auf den Ausgang

$$\mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \times \mathcal{U}_{ad} \rightarrow \mathbb{R}^p, \quad (t_0, x_0, t, u) \mapsto y(t)$$

heißt Ausgangsübergangsfunktion (*input-output mapping*) oder auch Transferfunktion (*transfer function*).

1.1.2 Die Laplace Transformation

Die Laplace Transformation ist eine Integraltransformation, die Systeme aus dem (reellen) Zeitbereich in den sogenannten (komplexen) *Spektral- oder Frequenzbereich* überführen. Die Laplace Transformation lässt sich wie folgt definieren.

Definition 1.12. Sei $f: [0, \infty) \rightarrow \mathbb{R}^n$. Existiert das Integral

$$\mathcal{L}\{f\}(s) := \int_0^{\infty} e^{-st} f(t) dt,$$

dann heißt $\mathcal{L}\{f\}$ die Laplace-Transformierte von f .

Ein Vorteil der Laplace Transformation ist, dass man eine lineare ODE in eine algebraische Gleichung umwandeln kann. Analog kann man bei partiellen Differentialgleichungen eine Variable eliminieren.

Lemma 1.13. Seien f und g zwei Funktionen, die eine Laplace-Transformierte besitzen und $\alpha, \beta \in \mathbb{R}$. Dann gelten folgende Eigenschaften:

- i) $\mathcal{L}\{\alpha f + \beta g\} = \alpha \mathcal{L}\{f\} + \beta \mathcal{L}\{g\}$,
- ii) $\mathcal{L}\{f'\}(s) = s \mathcal{L}\{f\}(s) - f(0)$,
- iii) $\mathcal{L}\{\int_0^t f(\eta) d\eta\}(s) = s^{-1} \mathcal{L}\{f\}(s)$,
- iv) $\mathcal{L}\{f^{(n)}\}(s) = s^n \mathcal{L}\{f\}(s) - s^{n-1} f(0) - \dots - f^{(n-1)}(0)$,
- v) $\mathcal{L}\{e^{at}\}(s) = \frac{1}{s-a}$ für $\Re(s) > a$,
- vi) $\mathcal{L}\{t^n\}(s) = \frac{n!}{s^{n+1}}$ für $\Re(s) > 0$.

Beweis. Übungsaufgabe (optional) □

Bemerkung 1.14. Analog zur Laplace Transformation für Differentialgleichungen benutzt man für diskrete Systeme den Shift-Operator für Folgen,

$$\mathcal{S}(x_0, x_1, \dots) := (x_1, x_2, \dots).$$

Mit $x = (x_0, x_1, \dots)$, $u = (u_0, u_1, \dots)$ und $y = (y_0, y_1, \dots)$ wird dann aus dem diskreten linearen Kontrollproblem (1.1) mit konstanten Matrizen $A_k \equiv A$, $B_k \equiv B$, $C_k \equiv C$, $D_k \equiv D$ das System

$$\mathcal{S}x = Ax + Bu, \quad y = Cx + Du$$

und damit $x = (\mathcal{S} - A)^{-1}Bu$.

1.1.3 Die Singulärwertzerlegung

Ein wichtiges Hilfsmittel für spätere Normalformen ist die Singulärwertzerlegung (SVD).

Satz 1.15 (Singulärwertzerlegung, SVD). Sei $A \in \mathbb{C}^{m,n}$ eine Matrix mit Rang r . Dann existieren unitäre Matrizen $U \in \mathbb{C}^{m,m}$ und $V \in \mathbb{C}^{n,n}$, sodass

$$A = U\Sigma V^*, \quad \Sigma = \begin{bmatrix} \text{diag}(\sigma_1, \dots, \sigma_r) & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{m,n}$$

mit $\sigma_1 \geq \dots \geq \sigma_r > 0$. Die Werte $\sigma_1, \dots, \sigma_r$ sind eindeutig bestimmt.

Beweis. Beweis wird in der Vorlesung Numerik I gegeben. Alternativ findet man ihn im Buch Golub, Van Loan: *Matrix computations*, 3. Edition, 1996. \square

Definition 1.16. Seien A , $U = [u_1, \dots, u_m]$, $V = [v_1, \dots, v_n]$ und Σ wie in Satz 1.15. Zudem setzen wir $\sigma_i := 0$ für $i = r + 1, \dots, \min(m, n)$. Dann nennen wir $\sigma_1, \dots, \sigma_{\min(m,n)}$ die Singulärwerte (*singular values*) von A . Die Vektoren u_1, \dots, u_m heißen linke Singulärvektoren (*left singular vectors*) von A und v_1, \dots, v_n rechte Singulärvektoren (*right singular vectors*) von A .

Bemerkung 1.17. Mithilfe der SVD lassen sich die Frobenius- und die 2-Norm berechnen,

$$\|A\|_F^2 = \sigma_1^2 + \dots + \sigma_r^2, \quad \|A\|_2 = \sigma_1.$$

Bemerkung 1.18. Die Werte $\sigma_1^2, \dots, \sigma_r^2$ sind die positiven Eigenwerte von A^*A und AA^* . Da $AV = U\Sigma$, gilt

$$\text{Ker } A = \text{span}\{v_{r+1}, \dots, v_n\}, \quad \text{Rang } A = \text{span}\{u_1, \dots, u_r\}.$$

Zudem ist $A_k = \sum_{i=1}^k \sigma_i u_i v_i^*$ die beste Approximation von A mit Rang k , d.h.

$$\|A - A_k\|_2 = \inf_{B \in \mathbb{C}^{m,n}, \text{Rang } B=k} \|A - B\| = \sigma_{k+1}.$$

Im reellen Fall $A \in \mathbb{R}^{m,n}$ sind V und U orthogonal.

Bemerkung 1.19. Die SVD ist die beste Methode, um numerisch den Rang einer Matrix zu bestimmen.

Beispiel 1.20 (numerische Rang-Entscheidungen). *Wir betrachten die Matrix*

$$T_n(c) := \begin{bmatrix} 1 & & & & \\ & s & & & \\ & & s^2 & & \\ & & & \ddots & \\ & & & & s^{n-1} \end{bmatrix} \cdot \begin{bmatrix} 1 & -c & -c & \dots & -c \\ & 1 & -c & \dots & -c \\ & & 1 & \ddots & \vdots \\ & & & \ddots & -c \\ & & & & 1 \end{bmatrix} \in \mathbb{R}^{n,n}$$

mit $c^2 + s^2 = 1$ und $s, c > 0$. Ziel ist es, den Rang der Matrix $T_{100}(\sqrt{0.2})$ zu berechnen. Wir zeigen

$$\sigma_n(T_n(c)) = \min_{x \neq 0} \frac{\|T_n(c)x\|_2}{\|x\|_2} \leq \frac{s^{n-1}}{c(c+1)^{n-2}}.$$

Damit berechnen wir

n	10	25	50	75	100
$\sigma_n \leq$	$4.26 e^{-2}$	$3.12 e^{-5}$	$1.86 e^{-10}$	$1.11 e^{-15}$	$6.61 e^{-21}$

d.h., die Matrix $T_n(c)$ wirkt singular obwohl die Diagonalelemente der einzelnen Matrizen weit weg von der 0 sind, $s^{99} \approx 1.6 e^{-5}$. *Übungsaufgabe (optional)*

1.1.4 Invariante Unterräume

Eine wichtige Eigenschaft von Unterräumen ist die Invarianz bezüglich einer Matrix.

Definition 1.21. Sei $\mathcal{V} \subseteq \mathbb{R}^n$ ein Unterraum von \mathbb{R}^n und $A \in \mathbb{R}^{n,n}$ eine Matrix. Der Unterraum \mathcal{V} heißt A -invariant, wenn $A\mathcal{V} \subseteq \mathcal{V}$.

Eine Charakterisierung A -invarianter Unterräume ist durch das folgende Lemma gegeben.

Lemma 1.22. Seien $A \in \mathbb{R}^{n,n}$ und $X \in \mathbb{R}^{n,k}$ Matrizen und $\mathcal{V} := \text{Bild } X$. Dann ist \mathcal{V} A -invariant genau dann, wenn eine Matrix $B \in \mathbb{R}^{k,k}$ existiert mit $AX = XB$. Hat X vollen Spaltenrang, dann gilt zudem $\Lambda(B) \subseteq \Lambda(A)$.

Beweis. Übungsaufgabe 1.2a

□

Beispiel 1.23. Sei $A \in \mathbb{R}^{n,n}$ eine Matrix. Zu einer Menge von Eigenvektoren x_1, \dots, x_k definieren wir die Matrix $X := [x_1, \dots, x_k] \in \mathbb{R}^{n,k}$. Dann ist $\text{Bild } X = \text{span}\{x_1, \dots, x_k\}$ ein A -invarianter Unterraum von \mathbb{R}^n .

Um dies zu sehen, seien $\lambda_1, \dots, \lambda_k$ die entsprechenden Eigenwerte von A . Dann gilt

$$AX = [Ax_1 \ \dots \ Ax_k] = [\lambda_1 x_1 \ \dots \ \lambda_k x_k] = X \text{diag}(\lambda_1, \dots, \lambda_k).$$

Nach Lemma 1.22 ist $\text{Bild } X$ dann A -invariant.

Eigenräume bilden also invariante Unterräume. Andersrum gilt dies aber auch wie das folgende Lemma zeigt.

Lemma 1.24. Sei $\mathcal{V} = \text{Bild } X \subseteq \mathbb{R}^n$ ein A -invarianter Unterraum. Dann ist \mathcal{V} der Spann von gewissen Eigen- und Hauptvektoren von A .

Beweis. Übungsaufgabe 1.2b □

Definition 1.25. Sei wieder $A \in \mathbb{R}^{n,n}$ und $\mathcal{P} \subseteq \mathbb{C}$. Dann bezeichnet der A -invariante Unterraum zu den Eigenwerten in \mathcal{P} den Spann der Eigen- und Hauptvektoren von A zu den Eigenwerten $\Lambda(A) \cap \mathcal{P}$.

1.2 Beispiel: Invertiertes Pendel

Eine wichtige Fragestellung in der Robotik ist die Kontrolle der Position eines Gelenks mittels eines Motors. Das mathematische Modell dazu ist das invertierte Pendel. Hier ist das Ziel, das Pendel im höchsten Punkt zu stabilisieren, indem ein Drehmoment im Lager angelegt wird.

Bei der Modellierung nehmen wir an, dass die Masse am Ende des Stabes liegt. Zudem vernachlässigen wir Reibung.

Beispiel 1.26. Folgt man Newtons Gesetz erhalten wir die DGL

$$m\ddot{\theta}(t) + mg \sin \theta(t) = u(t).$$

Dabei ist die externe Kraft durch das Drehmoment u , welches wir steuern, gegeben. Die Variable θ beschreibt die Bewegung des Winkels zwischen der negativen y -Achse und dem Pendel gegen den Uhrzeigersinn, siehe Abbildung 1.2. Zur Vereinfachung betrachten wir normierte Länge und Masse, d.h., $l = 1$, $m = 1$ und $g = 1$.

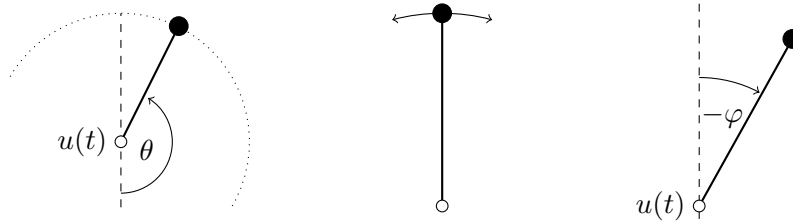


Abbildung 1.2: Invertiertes Pendel: Modell mit Winkel θ (links), instabiles Gleichgewicht (mitte), linearisiertes Modell mit Winkel φ .

Für $u(t) \equiv 0$ ist der Zustand $[\theta, \dot{\theta}]^T = [\pi, 0]^T$ ein instabiles Gleichgewicht. Wir linearisieren die Differentialgleichung um $\theta = \pi$. Dazu setzen wir $\varphi(t) := \theta(t) - \pi$ und erhalten

$$\sin \theta(t) = \sin(\pi + \varphi(t)) = \sin(\pi) + \cos(\pi)\varphi(t) + \mathcal{O}(\varphi^2) = -\varphi(t) + \mathcal{O}(\varphi^2).$$

Für kleine Auslenkungen $\varphi(t)$ gilt also $\sin \theta(t) \approx -\varphi(t)$. Eingesetzt in die Differentialgleichung ergibt das

$$\ddot{\varphi}(t) - \varphi(t) = u(t).$$

Wir definieren den Zustand $x(t) = [\varphi(t), \dot{\varphi}(t)]^T$ und erhalten das zeitinvariante System

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \quad y(t) = Cx(t).$$

Das Ziel ist nun für gegebenes x_0 ein $u(t)$ zu finden, sodass das Pendel in die obere Position gebracht (und gehalten) wird. Eine Daumenregel besagt, dass Regelungen, die auf einer Linearisierung basieren, zumindest lokal um den Linearisierungspunkt gut funktionieren. Wir untersuchen zwei Ansätze.

Proportionale Zustandsrückführung (naiver Ansatz)

Wir nehmen an, dass φ gemessen werden kann, also $C = [1 \ 0]$. Im Fall $\varphi < 0$ gibt es eine Abweichung von der Nulllage nach rechts und damit ist ein positives Drehmoment nötig. Im anderen Fall analog ein negatives Drehmoment. Mit dem Ansatz $u(t) = -\alpha\varphi(t)$, $\alpha > 0$, ergibt sich

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [-\alpha \ 0] x(t) = \begin{bmatrix} 0 & 1 \\ 1 - \alpha & 0 \end{bmatrix} x(t) =: \tilde{A}x(t).$$

Aus Beispiel 1.7 wissen wir, dass die Eigenwerte von \tilde{A} durch $\lambda_{1,2} = \pm\sqrt{1-\alpha}$ gegeben sind. Damit ergibt sich (im Fall $\lambda_1 \neq \lambda_2$) die Lösung

$$\varphi(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

Es treten die folgenden Fälle auf:

- i) $0 < \alpha < 1$: Beide Eigenwerte sind reell und $\lambda_1 > 0$, $\lambda_2 < 0$. Das System ist also instabil und $\varphi(t) \rightarrow \infty$.
- ii) $\alpha = 1$: Doppelter Eigenwert $\lambda_{1,2} = 0$. Bei \tilde{A} handelt es sich um einen Jordanblock. Die algebraische und geometrische Vielfachheit vom Eigenwert 0 stimmen also nicht überein. Das System ist instabil.
- iii) $\alpha > 1$: Es gilt $\Re(\lambda_{1,2}) = 0$ und $\varphi(t)$ schwingt zwischen $\varphi(0)$ und $-\varphi(0)$.

Das System erfüllt also für beliebige Wahl von α nicht die Voraussetzungen.

2. Versuch: Proportional + Ableitung

Wir nehmen an, dass zusätzlich auch $\dot{\varphi}$ gemessen werden kann, also $C = [1 \ 1]$. Für die Kontrolle nehmen wir den Ansatz

$$u(t) = -\alpha\varphi(t) - \beta\dot{\varphi}(t)$$

Dabei sei entsprechend den Ergebnissen aus Beispiel 1.7 $\alpha > 1$ und $\beta > 0$. Dann gilt

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) - \begin{bmatrix} 0 \\ 1 \end{bmatrix} [\alpha \ \beta] x(t) = \begin{bmatrix} 0 & 1 \\ 1 - \alpha & -\beta \end{bmatrix} x(t).$$

Die entsprechenden Eigenwerte sind gegeben durch

$$\lambda_{1,2} = \frac{1}{2} \left(-\beta \pm \sqrt{\beta^2 + 4(1-\alpha)} \right).$$

In dem Fall gilt also $\lim_{t \rightarrow \infty} \varphi(t) = 0$, wobei je nach Imaginärteil möglicherweise noch eine Schwingung auftritt.

1.3 Beispiel: Parabolantenne

Das Steuerungsproblem einer Parabolantenne, die kontinuierlich auf einen Satellit oder ein Raumfahrzeug ausgerichtet wird, führt auf die linearisierten

Differentialgleichungen

$$\dot{\varphi}(t) = \omega(t), \quad (1.2a)$$

$$j\dot{\omega}(t) = -r\omega(t) + ku(t). \quad (1.2b)$$

Dabei bezeichnen φ den Rotationswinkel und ω die Winkelgeschwindigkeit. Die Konstanten beinhalten das Massenträgheitsmoment des Drehmotors (**moment of inertia**) j , einen Reibungskoeffizienten r sowie einen Verstärkungsfaktor k . Gesteuert wird die Eingangsspannung des Motors $u(t)$. Wir wollen die Lösung des Systems (1.2) bestimmen zu Anfangswerten $\varphi(0) = 0$ und $\omega(0) = \omega_0$. Dazu betrachten wir drei Fälle für die Steuerfunktion:

i) $u(t) = 0$ (freies System),

ii) $u(t) = \alpha\omega(t)$, $\alpha \in \mathbb{R}$,

iii) $u(t) = e^{-t}$.

Außerdem soll ein u gefunden werden, sodass $\varphi(1) = \pi$ gilt. Dafür nutze man den Ansatz $\omega(t) = \omega_0 + \gamma t$. [Übungsaufgabe 1.3](#)

2 Lineare Kontrollsysteme

Zunächst wollen wir definieren was ein lineares Kontrollsysteme überhaupt ist.

Definition 2.1. *Ein lineares Kontrollproblem besteht aus einem Anfangswertproblem, d.h. einer linearen DGL mit Anfangswert,*

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0 \quad (2.1a)$$

auf dem Intervall $[t_0, \infty)$ und einer Gleichung

$$y(t) = C(t)x(t) + D(t)u(t). \quad (2.1b)$$

Dabei ist $x_0 \in \mathbb{R}^n$. Ferner seien A , B , C und D stetige Funktionen der Form

$$\begin{aligned} A: [t_0, \infty) &\rightarrow \mathbb{R}^{n,n}, & B: [t_0, \infty) &\rightarrow \mathbb{R}^{n,m} \\ C: [t_0, \infty) &\rightarrow \mathbb{R}^{p,n}, & D: [t_0, \infty) &\rightarrow \mathbb{R}^{p,m}. \end{aligned}$$

Mit $\mathcal{U}_{ad} \subseteq PC([t_0, \infty), \mathbb{R}^m)$ bezeichnen wir den Raum der zulässigen Kontrollen und $u \in \mathcal{U}_{ad}$ ist dementsprechend eine (zulässige) Kontrolle (*admissible control*). Eine Lösung x der Zustandsgleichung heißt Zustand (*state*), eine Lösung y der zugehörigen Output-Gleichung heißt Output des Systems. Für eine zulässige Kontrolle $\hat{u} \in \mathcal{U}_{ad}$ bezeichnet $x(t; \hat{u})$ die dazugehörige Lösung der Gleichung $\dot{x} = Ax + B\hat{u}$.

Definition 2.2. Ein lineares Kontrollproblem heißt zeitinvariant, kurz LTI (*linear time-invariant*), falls A, B, C und D konstant sind. Andernfalls heißt es zeitvariant, kurz LTV (*linear time-variant*).

Bemerkung 2.3. Nach dem Satz von Picard-Lindelöf existiert die Lösung des Anfangswertproblems

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0$$

und ist eindeutig für beliebige $u \in \mathcal{U}_{ad}$. Oft wählt man $\mathcal{U}_{ad} = PC([t_0, \infty), \mathbb{R}^m)$ oder $\mathcal{U}_{ad} = C([t_0, \infty), \mathbb{R}^m)$.

Bemerkung 2.4. Für LTI Systeme wird meist $t_0 = 0$ gewählt, da jedes System mit $t_0 \neq 0$ entsprechend transformiert werden kann. Dafür ersetzt man t durch $t + t_0$. In dem Fall hat das lineare Kontrollproblem die Form

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad (2.2a)$$

$$y(t) = Cx(t) + Du(t). \quad (2.2b)$$

Zudem ist oft $D = 0$.

In Hinblick auf lineare Kontrollprobleme ergeben sich folgende Fragen:

- i) Steuerbarkeit (*controllability*): Gibt es zu gegebenen $x_0, x_1 \in \mathbb{R}^n$ ein $\hat{u} \in \mathcal{U}_{ad}$, sodass ein $t_1 \geq t_0$ existiert, welches $x(t_1; \hat{u}) = x_1$ erfüllt?
- ii) Stabilisierbarkeit (*stabilizability*): Gibt es zu einer gegebenen Referenz-Trajektorie $\hat{x}(t)$ ein $\hat{u} \in \mathcal{U}_{ad}$, sodass

$$\lim_{t \rightarrow \infty} \|x(t; \hat{u}) - \hat{x}(t)\| = 0?$$

Häufig ist hier $\hat{x}(t) \equiv c \in \mathbb{R}$ oder sogar $\hat{x}(t) \equiv 0$.

- iii) Optimalsteuerung (*optimal control*): Gibt es ein $\hat{u} \in \mathcal{U}_{ad}$, welches ein gegebenes Kostenfunktional $J(x(t; u), u)$ minimiert? Solche Problemstellungen betrachten wir in Abschnitt 4.

2.1 Kontrollprobleme im Frequenzbereich

Wir benutzen die Laplace Transformation aus Abschnitt 1.1.2, um Kontrollprobleme in den Frequenzbereich zu überführen. Da Produkte von Funktionen dadurch zu Faltungen werden, wird dies hauptsächlich für LTI Systeme benutzt. Wir betrachten also das System (2.2). Die Laplace Transformation mit $\hat{x}(s) := \mathcal{L}\{x\}(s)$ führt dann auf das System

$$\begin{aligned} s\hat{x}(s) &= A\hat{x}(s) + B\hat{u}(s) + x_0, \\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s). \end{aligned}$$

Hierbei handelt es sich um algebraische Gleichungen mit der Lösung

$$\hat{x}(s) = (sI - A)^{-1}B\hat{u}(s) + (sI - A)^{-1}x_0$$

und

$$\begin{aligned} \hat{y}(s) &= C(sI - A)^{-1}B\hat{u}(s) + D\hat{u}(s) + C(sI - A)^{-1}x_0 \\ &= \underbrace{(C(sI - A)^{-1}B + D)}_{=: H(s)} \hat{u}(s) + C(sI - A)^{-1}x_0. \end{aligned}$$

Die so definierte Funktion H heißt *Transfer- oder auch Übertragungsfunktion* (**transfer function**). Plottet man H über die imaginäre Achse (*Frequenzgang*) - das sogenannte *Bode-Diagramm* - lässt sich daraus ablesen was mit verschiedenen Frequenzen im System passiert. Solch ein Diagramm zeigt den Zusammenhang zwischen harmonischen Anregungen im Eingang u und dem Ausgang y .

Man beachte, dass die Transferfunktion H im Fall $x_0 = 0$ die gesamte Information des Systems beinhaltet.

Bemerkung 2.5. *Wir werden vor allem mit der Formulierung im Zeitbereich (2.2) arbeiten, da der Frequenzbereich nur für LTI Systeme geeignet ist. Dafür existieren bessere numerische Algorithmen in Bezug auf Effizienz und numerischer Stabilität. Der Frequenzbereich ist aber zum Beispiel in der Modellreduktion sehr nützlich. Dabei versucht man, Systeme hoher Dimension durch kleine Systeme möglichst gut zu approximieren.*

Beispiel 2.6 (RLC Schaltkreis). *Wir betrachten eine Reihenschaltung aus Widerstand (R), Induktionsspule (L) und Kondensator (C), vgl. Abbildung 2.1.*

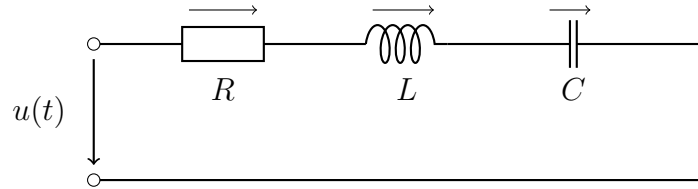


Abbildung 2.1: RCL Schaltkreis aus Beispiel 2.6.

Der Eingang des Systems ist gegeben durch die Spannung $u(t)$ an den Kontakten. Mit v_R , v_L und v_C bezeichnen wir die Spannung an R , L bzw. C . Analog ist der Strom am entsprechenden Element durch i_R , i_L und i_C gegeben. Um das lineare Kontrollsystem herzuleiten, nutzen wir folgende lineare Gesetze mit Konstanten R , L und C :

$$\begin{aligned}
 \text{linearer Widerstand:} & & v_R(t) &= R i_R(t), \\
 \text{lineare Induktivität:} & & u(t) &= L \frac{d}{dt} i_L(t), \\
 \text{linearer Kondensator:} & & i_C(t) &= C \frac{d}{dt} v_C(t).
 \end{aligned}$$

Zudem haben wir die Kirchhoff'schen Gesetze

$$i_R(t) = i_C(t) = i_L(t) =: i(t), \quad u(t) = v_R(t) + v_C(t) + v_L(t).$$

Als Zustandsvariablen wählen wir $x_1 = i_L(t) = i(t)$ (Strom durch die Induktionsspule) und $x_2 = v_C(t)$ (Spannung am Kondensator). Als Ausgang messen wir die Spannung am Widerstand, also $y(t) = v_R(t) = R x_1(t)$. Die Gleichung für den Kondensator impliziert $x_1 = C \dot{x}_2$. Außerdem erhalten wir

$$u(t) = R i_R(t) + v_C(t) + L \frac{d}{dt} i(t) = R x_1(t) + x_2(t) + L \dot{x}_1(t).$$

Somit erhalten wir für $x := [x_1, x_2]^T$ das lineare Kontrollproblem

$$\begin{aligned}
 \dot{x} &= \begin{bmatrix} -\frac{R}{L} & -\frac{1}{L} \\ \frac{1}{C} & 0 \end{bmatrix} x + \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} u, \\
 y &= [R \quad 0] x.
 \end{aligned}$$

Für die Transferfunktion ergibt sich also

$$\begin{aligned}
 H(s) &= (C(sI - A)^{-1}B) \\
 &= [R \ 0] \begin{bmatrix} s + \frac{R}{L} & \frac{1}{L} \\ -\frac{1}{C} & s \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} \\
 &= \left(s^2 + \frac{R}{L}s + \frac{1}{LC}\right)^{-1} [R \ 0] \begin{bmatrix} s & -\frac{1}{L} \\ \frac{1}{C} & s + \frac{R}{L} \end{bmatrix} \begin{bmatrix} \frac{1}{L} \\ 0 \end{bmatrix} \\
 &= \frac{Rs}{L\left(s^2 + \frac{Rs}{L} + \frac{1}{LC}\right)} \\
 &= \frac{Rs}{s^2L + Rs + \frac{1}{C}}.
 \end{aligned}$$

Für einen Eingang u mit Laplace Transformation \hat{u} berechnet man dann den Ausgang y im Fall $x_0 = 0$ mittels $\hat{y}(s) = H(s)\hat{u}(s)$.

Schließlich betrachten wir noch die Stabilität des Systems, also die Eigenwerte der Matrix A . Diese sind gegeben durch

$$\lambda_{1,2} = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}.$$

Beide Eigenwerte haben negativen Realteil, sodass das System asymptotisch stabil ist.

2.2 Steuerbarkeit

Wir kommen nun zur Frage der Steuerbarkeit ([controllability](#)). Gesucht ist also eine Kontrolle $u \in \mathcal{U}_{\text{ad}}$, die das LTV System von (t_0, x_0) zu (t_1, x_1) bringt. Zunächst geben wir einige Definitionen.

Definition 2.7. Gegeben sei ein LTV System der Form (2.1).

- i) Das System heißt steuerbar nach x_1 zur Zeit t_1 mit $t_1 > t_0$, falls es ein $u \in \mathcal{U}_{\text{ad}}$ gibt, sodass $x(t_1; u) = x_1$ gilt.
- ii) Es heißt steuerbar von x_0 nach x_1 , falls ein t_1 existiert, sodass es nach x_1 zur Zeit t_1 steuerbar ist.
- iii) Es heißt (vollständig) steuerbar ([controllable](#)), falls es von jedem x_0 zu jedem x_1 steuerbar ist.

iv) Seien $t_0, t_1 \in \mathbb{R}$ mit $t_1 > t_0$ und $x_1 \in \mathbb{R}^n$. Wir definieren die Mengen

(a) $\mathcal{C}(t_0, t_1, x_1) := \{x_0 \in \mathbb{R}^n \mid (t_0, x_0) \text{ ist steuerbar nach } (t_1, x_1)\}$,

(b) $\mathcal{C}(t_0, x_1) := \bigcup_{t_1 > t_0} \mathcal{C}(t_0, t_1, x_1)$, die Steuerbarkeitsmenge von x_1
(controllability set with respect to x_1),

(c) $\mathcal{C} := \mathcal{C}(0, 0)$ im LTI Fall.

2.2.1 Zeitvarianter Fall

Wir betrachten zunächst den LTV Fall aber beschränken uns auf $x_1 = 0$.

Lemma 2.8. Sei $x_0 \in \mathbb{R}^n$. Dann sind die zwei folgenden Aussagen äquivalent.

i) Es gilt $x_0 \in \mathcal{C}(t_0, t_1, 0)$.

ii) Es existiert ein $u \in \mathcal{U}_{ad}$, sodass $x_0 = - \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s) ds$.

Beweis. Die Aussage $x_0 \in \mathcal{C}(t_0, t_1, 0)$ bedeutet, dass man x_0 zur 0 steuern kann. Es gibt also eine Kontrolle $u \in \mathcal{U}_{ad}$ und ein t_1 , sodass

$$\begin{aligned} 0 = x(t_1; u) &= \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, s)B(s)u(s) ds \\ &= \Phi(t_1, t_0) \left(x_0 + \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s) ds \right). \end{aligned}$$

Da $\Phi(t_1, t_0)$ regulär ist, ist dies äquivalent zu

$$x_0 = - \int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s) ds. \quad \square$$

Definition 2.9. Sei $G: [t_0, t_1] \rightarrow \mathbb{R}^{n,m}$ stückweise stetig. Dann heißt

$$V(t_0, t_1) := \int_{t_0}^{t_1} G(s)G(s)^T ds \in \mathbb{R}^{n,n}$$

die Gramsche Matrix (*Gramian*) von G auf $[t_0, t_1]$.

Lemma 2.10. Sei $V(t_0, t_1)$ die Gramsche Matrix von der stückweise stetigen Funktion $G: [t_0, t_1] \rightarrow \mathbb{R}^{n,m}$ auf $[t_0, t_1]$. Dann hat V die folgenden Eigenschaften:

- i) $V(t_0, t_1)$ ist symmetrisch und positiv semidefinit,
- ii) $\text{Ker } V(t_0, t_1) = \{z \in \mathbb{R}^n \mid G(t)^T z \equiv 0 \text{ für alle } t \in [t_0, t_1]\}$,
- iii) die folgenden Aussagen äquivalent:
- (a) Es gibt ein $u \in \mathcal{U}_{ad}$, sodass $x_0 = \int_{t_0}^{t_1} G(s)u(s) ds$,
- (b) $x_0 \in \text{Bild } V(t_0, t_1)$, d.h. es gibt ein $z \in \mathbb{R}^n$ mit $x_0 = V(t_0, t_1)z$.

Beweis.

- i) Offensichtlich ist $V(t_0, t_1)$ symmetrisch. Des Weiteren gilt für alle $z \in \mathbb{R}^n$, dass

$$z^T V(t_0, t_1)z = \int_{t_0}^{t_1} z^T G(s)G(s)^T z ds = \int_{t_0}^{t_1} \|G(s)^T z\|_2^2 ds \geq 0.$$

- ii) Die Rechnung zu i) zeigt, dass $V(t_0, t_1)z = 0$ genau dann gilt, wenn $\|G(s)^T z\|_2 \equiv 0$, also $G(t)^T z \equiv 0$ auf $[t_0, t_1]$.
- iii) Wir zeigen zunächst die Rückrichtung. Sei also $x_0 \in \text{Bild } V(t_0, t_1)$. Dann gibt es ein $z \in \mathbb{R}^n$ mit

$$x_0 = V(t_0, t_1)z = \int_{t_0}^{t_1} G(s)G(s)^T z ds.$$

Wählen wir nun als Kontrolle $u(s) = G(s)^T z$, dann hat x_0 die gewünschte Form. Nun definieren wir den Unterraum

$$L := \{x_0 \in \mathbb{R}^n \mid \text{es gibt ein } u \in \mathcal{U}_{ad} \text{ mit } x_0 = \int_{t_0}^{t_1} G(s)u(s) ds\}.$$

Wir haben bereits $\text{Bild } V(t_0, t_1) \subseteq L$ gezeigt und beweisen nun die Gleichheit der beiden Unterräume und somit die Äquivalenz der beiden Aussagen. Dazu zeigen wir $\text{Ker } V(t_0, t_1) \cap L = \{0\}$. Sei also $x_0 \in \text{Ker } V(t_0, t_1) \cap L$. Dann gilt wegen ii) $G(t)^T x_0 \equiv 0$ und somit

$$\|x_0\|_2^2 = x_0^T x_0 = x_0^T \int_{t_0}^{t_1} G(s)u(s) ds = \int_{t_0}^{t_1} \underbrace{x_0^T G(s)}_{\equiv 0} u(s) ds = 0.$$

Damit folgt $x_0 = 0$ und $\text{Ker } V(t_0, t_1) \cap L = \{0\}$. Die Gleichheit der beiden Unterräume folgt dann nach einem Dimensionsargument. Die Abschätzung

$$\begin{aligned} n &\geq \dim(\text{Ker } V(t_0, t_1) + L) \\ &= \dim(\text{Ker } V(t_0, t_1)) + \dim L - \dim(\text{Ker } V(t_0, t_1) \cap L) \\ &\geq \dim(\text{Ker } V(t_0, t_1)) + \dim(\text{Bild } V(t_0, t_1)) = n \end{aligned}$$

liefert $\dim L = \dim(\text{Bild } V(t_0, t_1))$ und damit die Behauptung. \square

Bemerkung 2.11. *Lässt sich x_0 in der Form $x_0 = \int_{t_0}^{t_1} G(s)u(s) ds$ mit $G(s) = -\Phi(t_0, s)B(s)$ darstellen, dann kann man den Zustand $x_1 = 0$ erreichen. Zusammen mit Lemma 2.10 erhält man dann eine gewünschte Kontrollfunktion mit dem Ansatz $u(t) = -B^T(s)\Phi^T(t_0, s)z_0$ für ein $z_0 \in \mathbb{R}^n$.*

Die spezielle Wahl von G führt auf die folgende Definition.

Definition 2.12. *Die Matrix*

$$W(t_0, t_1) := \int_{t_0}^{t_1} \Phi(t_0, s)B(s)B(s)^T\Phi(t_0, s)^T ds$$

heißt Steuerbarkeits-Gramsche (*controllability Gramian*) des LTV Systems.

Die bisherigen Ergebnisse implizieren das folgende Resultat.

Satz 2.13. *Für die Steuerbarkeits-Gramsche $W(t_0, t_1)$ und jedes $z \in \mathbb{R}^n$ gilt*

- i) $\mathcal{C}(t_0, t_1, 0) = \text{Bild } W(t_0, t_1)$
 $= \{x_0 \in \mathbb{R}^n \mid \text{es gibt ein } u \in \mathcal{U}_{ad} \text{ mit } x_0 = -\int_{t_0}^{t_1} \Phi(t_0, s)B(s)u(s) ds\},$
- ii) $W(t_0, t_1)z = 0 \Leftrightarrow B(t)^T\Phi(t_0, t)^T z \equiv 0 \text{ auf } [t_0, t_1].$
Also ist $\text{Ker } W(t_0, t_1) = \{z_0 \in \mathbb{R}^n \mid B(t)^T\Phi(t_0, t)^T z_0 \equiv 0 \text{ auf } [t_0, t_1]\}.$

Beweis. [Übungsaufgabe 2.1](#) \square

Definition 2.14. *Zu einer Differentialgleichung $\dot{x}(t) = A(t)x(t)$ definieren wir die adjungierte Gleichung (*adjoint equation*) mittels $\dot{z}(t) = -A(t)^T z(t)$.*

Lemma 2.15. *Sei $\Phi(t, t_0)$ die Fundamentallösung von $\dot{x}(t) = A(t)x(t)$. Dann ist $\Phi(t_0, t)^T$ die Fundamentallösung der adjungierten Gleichung.*

Beweis. Übungsaufgabe 1.1b □

Bemerkung 2.16. Sei x eine Lösung von $\dot{x}(t) = A(t)x(t)$ und z eine Lösung der adjungierten Gleichung. Dann gilt

$$\begin{aligned}\frac{d}{dt}(z(t)^T x(t)) &= \dot{z}(t)^T x(t) + z(t)^T \dot{x}(t) \\ &= -z(t)^T A(t)x(t) + z(t)^T A(t)x(t) = 0.\end{aligned}$$

Dies bedeutet, dass $z(t)^T x(t) \equiv c \in \mathbb{R}$.

Satz 2.17. Die drei folgenden Aussagen sind äquivalent:

- i) Das LTV System ist steuerbar, d.h. für beliebigen Start (t_0, x_0) wird jeder Endzustand erreicht.
- ii) Für jede Lösung z der adjungierten Gleichung $\dot{z}(t) = -A(t)^T z(t)$ gilt die folgende Implikation:

$$B(t)^T z(t) \equiv 0 \text{ auf } [t_0, \infty) \text{ für ein } t_0 \in \mathbb{R} \quad \Rightarrow \quad z(t) \equiv 0 \text{ auf } \mathbb{R}.$$

- iii) Für jedes $t_0 \in \mathbb{R}$ existiert ein $t_1 > t_0$, sodass $W(t_0, t_1)$ positiv definit ist.

Beweis. $i) \Rightarrow ii)$ (Widerspruch) Es existiere eine Lösung z von $\dot{z} = -A(t)^T z(t)$ mit $B(t)^T z(t) \equiv 0$ auf $[t_0, \infty)$ für ein $t_0 \in \mathbb{R}$, aber $z \not\equiv 0$ auf \mathbb{R} . Es existiert also ein \bar{t} mit $z(\bar{t}) \neq 0$. Da $z(\bar{t}) = \Phi(t_0, \bar{t})^T z(t_0)$, muss also auch $z(t_0) \neq 0$ gelten. Wähle nun $x_0 \in \mathbb{R}^n$ mit $x_0^T z(t_0) \neq 0$. Angenommen, das System sei steuerbar. Dann gibt es ein $t_1 > t_0$ und ein $u \in \mathcal{U}_{\text{ad}}$, sodass $x(t_0) = x_0$ und $x(t_1; u) = 0$. Definiere $f(t) := z(t)^T x(t)$. Dann gilt $f(t_0) \neq 0$ und $f(t_1) = 0$. Weiter gilt

$$\begin{aligned}f'(t) &= (z(t)^T x(t))' \\ &= \dot{z}(t)^T x(t) + z(t)^T (A(t)x(t) + B(t)u(t)) \\ &= -z(t)^T A(t)x(t) + z(t)^T A(t)x(t) + z(t)^T B(t)u(t) = 0\end{aligned}$$

für $t \geq t_0$. Dies ist nun aber ein Widerspruch zu $f(t_0) \neq 0 = f(t_1)$.

$ii) \Rightarrow iii)$ Der Beweis besteht aus zwei Schritten. Zunächst zeigen wir, dass zu jedem $t_0 \in \mathbb{R}$ ein $t_1 > t_0$ existiert, sodass jede Lösung $z \not\equiv 0$ der adjungierten Gleichung die Bedingung $B(t)^T z(t) \neq 0$ auf $[t_0, t_1]$ erfüllt. Wir

führen einen Widerspruchsbeweis und nehmen an, dass die Aussage falsch ist. In dem Fall existiert eine Folge $(t_k)_{k=1}^\infty$ mit $t_k \rightarrow \infty$ und eine Folge $(z_k(t))_{k=1}^\infty$ von Lösungen der adjungierten Gleichung $\dot{z} = -A(t)^T z(t)$. Durch Normierung erhalten wir $\|z_k(t_0)\| = 1$ und $B(t)^T z_k(t) = 0$ für alle $t \in [t_0, t_k]$. Da die Einheitskugel kompakt ist, existiert der Grenzwert (zumindest einer Teilfolge nach Bolzano-Weierstraß) $\hat{z}_0 := \lim_{k \rightarrow \infty} z_k(t_0)$. Es bezeichne \hat{z} die Lösung der adjungierten Gleichungen mit Anfangswert \hat{z}_0 . Dann gilt $\hat{z}(t_0) = \hat{z}_0 \neq 0$, da $\|\hat{z}_0\| = 1$, und somit auch $\hat{z} \not\equiv 0$. Punkt ii) impliziert nun $B(t)^T \hat{z}(t) \equiv 0$ auf $[t_0, \infty)$. Es existiert also ein $t_* > t_0$ mit $B(t_*)^T \hat{z}(t_*) \neq 0$. Da die Lösung der DGL stetig von den Anfangswerten abhängt, konvergiert auch $z_k(t_*) \rightarrow \hat{z}(t_*)$. Somit folgt $B(t_*)^T z_k(t_*) \neq 0$ für genügend großes k . Dies ist ein Widerspruch zur Wahl von z_k .

Im zweiten Schritt zeigen wir nun, dass das eben gezeigte $W(t_0, t_1) > 0$ impliziert. Da die Steuerbarkeits-Gramsche sowieso positiv semidefinit ist, reicht zu zeigen, dass der Kern von $W(t_0, t_1)$ nur aus dem Nullvektor besteht. Sei $z_0 \in \text{Ker } W(t_0, t_1)$. Dann gilt für alle $t \in [t_0, t_1]$ nach Satz 2.13 und Lemma 2.15,

$$0 = B(t)^T \Phi(t_0, t)^T z_0 = B(t)^T z(t),$$

da $\Phi(t_0, t)^T$ die Fundamentallösung der adjungierten Gleichung ist. Mit dem ersten Teil des Beweises folgt nun $z \equiv 0$ und somit auch $z_0 = z(t_0) = 0$.

iii) \Rightarrow i) [Übungsaufgabe 2.2a](#) □

Bemerkung 2.18. Aus dem Beweis von Satz 2.17 geht eine Formel für die „passende“ Kontrolle hervor, nämlich

$$u(t) = B(t)^T \Phi(t_0, t)^T W(t_0, t_1)^{-1} (\Phi(t_0, t_1) x_1 - x_0).$$

Bemerkung 2.19. Die Steuerbarkeits-Gramsche $W(t_0, t_1)$ eines LTV Systems ist genau dann positiv definit, wenn

$$\hat{W}(t_0, t_1) := \int_{t_0}^{t_1} \Phi(t_1, s) B(s) B(s)^T \Phi(t_1, s)^T ds$$

positiv definit ist. ([Übungsaufgabe 2.2b](#))

2.2.2 Zeitinvarianter Fall

Im Folgenden betrachten wir den Spezialfall eines LTI Systems. In diesem Fall erhalten wir eine rein algebraische Charakterisierung von Steuerbarkeit. Genauer gilt

- i) $\Phi(t, 0) = e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k$ und $\Phi(0, t) = e^{-At}$,
- ii) $W(0, t_1) = \int_0^{t_1} e^{-As} B B^T e^{-A^T s} ds$,
- iii) Bild $W(0, t_1) = \mathcal{C}(0, t_1, 0) =: \mathcal{C}_0(t_1)$ und
 $\text{Ker } W(0, t_1) = \{z_0 \in \mathbb{R}^n \mid z_0^T e^{-At} B \equiv 0 \text{ auf } [0, t_1]\}$.

Für asymptotisch stabile A gilt zudem das folgende Lemma.

Lemma 2.20. *Ein LTI System mit asymptotisch stabiler Matrix $A \in \mathbb{R}^{n,n}$ ist genau dann steuerbar, wenn*

$$W := \int_0^{\infty} e^{As} B B^T e^{A^T s} ds$$

positiv definit ist.

Beweis. Übungsaufgabe 2.2c □

Definition 2.21. *Die Matrix*

$$K = K(A, B) := [B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B] \in \mathbb{R}^{n,nm}$$

heißt Steuerbarkeitsmatrix (*controllability matrix*) des LTI Systems (2.2).

Satz 2.22. *Für LTI Systeme gilt $\mathcal{C}_0(t_1) = \text{Bild } K$. Insbesondere ist $\mathcal{C}_0(t_1)$ unabhängig von t_1 und $\mathcal{C} = \bigcup_{t_1 > 0} \mathcal{C}_0(t_1) = \text{Bild } K$.*

Beweis. Beide Mengen sind Teilräume von \mathbb{R}^n . Wir zeigen $\mathcal{C}_0(t_1)^\perp = (\text{Bild } K)^\perp$. Sei

$$z_0 \in \mathcal{C}_0(t_1)^\perp = (\text{Bild } W(0, t_1))^\perp = \text{Ker } W(0, t_1).$$

Die letzte Gleichheit gilt, da $W(0, t_1)$ symmetrisch ist. Für eine symmetrische Matrix S gilt nämlich

$$Sz = 0 \Leftrightarrow z^T S = 0 \Leftrightarrow z^T S w = 0 \quad \text{für alle } w.$$

Es gilt also $W(0, t_1)z_0 = 0$, was für LTI Systeme wiederum bedeutet, dass $z_0^T e^{-At} B \equiv 0$ auf dem Intervall $[0, t_1]$ gilt. Das ist äquivalent zu $z_0^T A^k B = 0$ für alle $k \in \mathbb{N}$. Der Satz von Cayley-Hamilton besagt, dass A^n eine Linearkombination von I, A, \dots, A^{n-1} ist. Damit gilt

$$\begin{aligned} z_0^T A^k B = 0 \quad \text{für alle } k &\Leftrightarrow z_0^T A^k B = 0, \quad k = 1 \dots n-1 \\ &\Leftrightarrow z_0^T K = 0 \\ &\Leftrightarrow z_0^T K w = 0 \quad \text{für alle } w \in \mathbb{R}^{nm} \\ &\Leftrightarrow z_0 \in (\text{Bild } K)^\perp. \end{aligned}$$

Da wir in einem endlich-dimensionalen Vektorraum arbeiten, folgt aus $\mathcal{C}_0(t_1)^\perp = (\text{Bild } K)^\perp$ die Behauptung. \square

Korollar 2.23. *Ein LTI System ist steuerbar genau dann, wenn $\text{Rang } K = n$ gilt.*

Beweis. Mit Satz 2.17 erhalten wir:

$$\begin{aligned} \text{Das System ist steuerbar.} &\Leftrightarrow W(0, t_1) \text{ ist positiv definit f\u00fcr ein } t_1. \\ &\Leftrightarrow \text{Ker } W(0, t_1) = \{0\} \text{ f\u00fcr ein } t_1. \\ &\Leftrightarrow \mathcal{C}_0(t_1) = \text{Bild } W(0, t_1) = \mathbb{R}^n \text{ f\u00fcr ein } t_1. \end{aligned}$$

Die Steuerbarkeit ist also \u00e4quivalent zu $\text{Bild } K = \mathcal{C}_0(t_1) = \mathbb{R}^n$ und damit auch zu $\text{Rang } K = n$. \square

Definition 2.24. *Das Paar $(A, B) \in \mathbb{R}^{n,n} \times \mathbb{R}^{n,m}$ hei\u00dft steuerbar (*controllable*), falls $\text{Rang } K(A, B) = n$ gilt.*

Eine Klassifizierung des Bildes der Steuerbarkeitsmatrix $K(A, B)$ ergibt sich durch seine A -Invarianz, vgl. Abschnitt 1.1.4.

Lemma 2.25. *Seien $A \in \mathbb{R}^{n,n}$ und $B \in \mathbb{R}^{n,m}$. Dann ist $\mathcal{K} := \text{Bild } K(A, B)$ der kleinste A -invariante Unterraum, der das Bild von B enth\u00e4lt.*

Beweis. Offensichtlich gilt $\text{Bild } B \subseteq \mathcal{K}$. Wir zeigen, dass \mathcal{K} A -invariant ist:

$$\begin{aligned} z_0 \in \mathcal{K} &\Rightarrow z_0 = Bw_1 + ABw_2 + \cdots + A^{n-1}Bw_n \\ &\Rightarrow Az_0 = ABw_1 + \cdots + A^nBw_n. \end{aligned}$$

Nach Cayley-Hamilton gilt wiederum $A^n = -\alpha_0 I - \alpha_1 A - \cdots - \alpha_{n-1} A^{n-1}$ und damit $Az_0 \in \mathcal{K}$. Schlie\u00dflich zeigen wir noch die Minimalit\u00e4t. F\u00fcr einen A -invarianten Unterraum \mathcal{U} mit $\text{Bild } B \subseteq \mathcal{U}$ m\u00fcssen wir also zeigen $\mathcal{K} \subseteq \mathcal{U}$. Zun\u00e4chst sei bemerkt, dass $A \text{Bild } B = \text{Bild}(AB)$ gilt. Damit folgt

$$\text{Bild}(AB) = A \text{Bild } B \subseteq A\mathcal{U} \subseteq \mathcal{U}.$$

Eine Wiederholung dieses Arguments liefert $\text{Bild}(A^{k-1}B) \subseteq \mathcal{U}$ f\u00fcr $k = 1, \dots, n$ und somit $\mathcal{K} \subseteq \mathcal{U}$. \square

Satz 2.26 (Kalman-Zerlegung). Seien $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$ und $\text{Rang } K(A, B) = r \leq n$. Dann existiert eine invertierbare Matrix $V \in \mathbb{R}^{n,n}$, sodass

$$V^{-1}AV = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}, \quad V^{-1}B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$$

und (A_1, B_1) steuerbar ist. Dabei ist $A_1 \in \mathbb{R}^{r,r}$. Zudem kann V orthogonal gewählt werden.

Beweis. Seien v_1, \dots, v_r sowie v_{r+1}, \dots, v_n jeweils eine Orthonormalbasis von $\mathcal{K} = \text{Bild } K(A, B)$ und \mathcal{K}^\perp . Dann ist $V = [v_1 \cdots v_n] \in \mathbb{R}^{n,n}$ offensichtlich orthogonal ($V^{-1} = V^T$) und es gilt für alle $j = r+1, \dots, n$ die Orthogonalität $v_j^T K(A, B) = 0$. Folglich gilt für eine Matrix $B_1 \in \mathbb{R}^{r,m}$,

$$V^{-1}B = V^T B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

Da $\mathcal{K} = \text{span}\{v_1, \dots, v_r\}$ A -invariant ist nach Lemma 2.25, gilt $Av_j \in \mathcal{K}$ für $j = 1, \dots, r$ und daher

$$AV = V \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} =: V\tilde{A}.$$

Das zeigt die spezielle Form von $V^{-1}AV$. Für die Steuerbarkeit von (A_1, B_1) müssen wir noch zeigen, dass $\text{Rang } K(A_1, B_1) = r$ (Korollar 2.23). Aus

$$V^{-1}A^k B = V^{-1}A^k V V^{-1}B = (V^{-1}AV)^k V^{-1}B = (V^{-1}V\tilde{A})^k V^{-1}B$$

folgt

$$V^{-1}A^k B = \tilde{A}^k V^{-1}B = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}^k \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = \begin{bmatrix} A_1^k & * \\ 0 & A_3^k \end{bmatrix} \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = \begin{bmatrix} A_1^k B_1 \\ 0 \end{bmatrix}.$$

Das ergibt dann schließlich

$$\text{Rang } K(A_1, B_1) = \text{Rang} \left(V \begin{bmatrix} K(A_1, B_1) \\ 0 \end{bmatrix} \right) = \text{Rang } K(A, B) = r. \quad \square$$

Bemerkung 2.27. Die Kalman-Zerlegung entspricht einem Basiswechsel $\tilde{x} := V^{-1}x = [\tilde{x}_1^T, \tilde{x}_2^T]^T$ mit

$$\begin{aligned} \dot{\tilde{x}}_1 &= A_1 \tilde{x}_1 + A_2 \tilde{x}_2 + B_1 u, \\ \dot{\tilde{x}}_2 &= A_3 \tilde{x}_2. \end{aligned}$$

Die Komponenten von \tilde{x}_2 sind durch den Anfangswert $\tilde{x}_2(0)$ eindeutig bestimmt und können nicht durch u beeinflusst werden. Die Zustände \tilde{x}_2 heißen nicht-steuerbar (*uncontrollable*). Die Eigenwerte von A_3 heißen nicht-steuerbare Moden (*uncontrollable modes*) von (A, B) .

Beispiel 2.28. Wir betrachten das LTI System

$$\dot{x} = \begin{bmatrix} 2 & 1 & 1 \\ 5 & -3 & -6 \\ -5 & 1 & 4 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u.$$

Es gilt

$$K(A, B) = [B \quad AB \quad A^2B] = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 5 & 25 \\ 0 & -5 & -25 \end{bmatrix},$$

wobei $\text{Rang } K(A, B) = 2 < n = 3$. Für die Kalman-Zerlegung wählen wir

$$V = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 5 & 0 \\ 0 & -5 & 1 \end{bmatrix} \quad \text{mit} \quad V^{-1} = \begin{bmatrix} 1 & -\frac{2}{5} & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Damit erhalten wir

$$V^{-1}AV = \left[\begin{array}{cc|c} 0 & -6 & \frac{17}{5} \\ 1 & 5 & -\frac{6}{5} \\ 0 & 0 & -2 \end{array} \right] \quad \text{und} \quad V^{-1}B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Somit ist B_1 gegeben durch $B_1 = [1, 0]^T$. Das steuerbare System ist gegeben durch (A_1, B_1) und wir überprüfen dies mittels

$$\text{Rang } K(A_1, B_1) = \text{Rang} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 2.$$

Das transformierte System hat somit die Form

$$\begin{aligned} \dot{\tilde{x}}_1 &= -6\tilde{x}_2 + \frac{17}{5}\tilde{x}_3 + u, & (\text{steuerbar}) \\ \dot{\tilde{x}}_2 &= \tilde{x}_1 + 5\tilde{x}_2 - \frac{6}{5}\tilde{x}_3, & (\text{steuerbar}) \\ \dot{\tilde{x}}_3 &= -2\tilde{x}_3. & (\text{nicht steuerbar}) \end{aligned}$$

Satz 2.29 (Hautus-Popov-Lemma). Für zeitinvariante Matrizen $A \in \mathbb{R}^{n,n}$ und $B \in \mathbb{R}^{n,m}$ sind folgende Aussagen äquivalent:

- i) (A, B) ist steuerbar, d.h. $\text{Rang } K(A, B) = n$.
- ii) Wenn $z \in \mathbb{R}^n$ ein Eigenvektor von A^T ist, dann gilt $z^T B \neq 0$.
- iii) $\text{Rang } [\lambda I - A \quad B] = n$ für alle $\lambda \in \mathbb{C}$ (Hautus-Test).

Beweis. i) \Rightarrow ii) Wir führen einen Widerspruchsbeweis. Sei $z \neq 0$ ein Eigenvektor von A^T zum Eigenwert λ , also $A^T z = \lambda z$ bzw. $z^T A = \lambda z^T$. Nehmen wir an, dass $z^T B = 0$, dann folgt

$$\begin{aligned} z^T A^k B &= \lambda^k z^T B = 0 \quad \text{für alle } k = 1, \dots, n-1 \\ \Rightarrow z^T K(A, B) &= 0 \\ \Rightarrow \text{Rang } K(A, B) &< n. \end{aligned}$$

Dies ist aber ein Widerspruch zur Steuerbarkeit von (A, B) .

ii) \Rightarrow iii) Wir nehmen an es gebe ein $\lambda \in \mathbb{C}$ mit $\text{Rang } [\lambda I - A \quad B] < n$. Dann existiert ein Vektor $z \neq 0$, sodass

$$z^T [\lambda I - A \quad B] = 0$$

und somit $\lambda z^T = A z^T$ und $z^T B = 0$. Also ist z ein Eigenvektor von A^T mit $z^T B = 0$, was ein Widerspruch ist.

iii) \Rightarrow i). Es sei $\text{Rang } K(A, B) < n$. Dann existiert mit der Kalman-Zerlegung eine orthogonale Matrix V , sodass

$$V^T A V = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}, \quad V^T B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

Sei nun $z_2 \neq 0$ ein Eigenvektor von A_3^T , d.h. $z_2^T A_3 = \lambda z_2^T$. Dann erfüllt der Vektor $z := V \begin{bmatrix} 0 \\ z_2 \end{bmatrix}$ die Gleichung

$$z^T B = [0 \quad z_2^T] V^T B = [0 \quad z_2^T] \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = 0.$$

Des Weiteren gilt

$$\begin{aligned} z^T A &= [0 \quad z_2^T] V^T A V V^T = [0 \quad z_2^T] \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T \\ &= [0 \quad z_2^T A_3] V^T \\ &= \lambda [0 \quad z_2^T] V^T = \lambda z^T. \end{aligned}$$

Daraus folgt nun $z^T [\lambda I - A \quad B] = 0$, also $\text{Rang } [\lambda I - A \quad B] < n$. □

Beispiel 2.30 (Invertiertes Pendel). Wir betrachten noch einmal das invertierte Pendel aus Abschnitt 1.2 mit den Matrizen

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Die Steuerbarkeitsmatrix $K(A, B)$ hat offensichtlich Rang 2. Das System ist also steuerbar. Wir führen auch noch den Hautus-Test durch. Ist λ kein Eigenwert von A , dann hat $[\lambda I - A \quad B]$ Rang zwei. Sei nun λ ein Eigenwert. Für $\lambda = 1$ erhalten wir

$$\text{Rang} [\lambda I - A \quad B] = \text{Rang} \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix} = 2.$$

Und auch für $\lambda = -1$ gilt

$$\text{Rang} [\lambda I - A \quad B] = \text{Rang} \begin{bmatrix} -1 & -1 & 0 \\ -1 & -1 & 1 \end{bmatrix} = 2.$$

Beispiel 2.31. Wir betrachten noch einmal das LTI System mit den Matrizen

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 5 & -3 & -6 \\ -5 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Die Eigenwerte sind gegeben durch $\Lambda(A) = \{-2, 2, 3\}$. Der Hautus-Test für $\lambda = -2$ ergibt

$$\text{Rang} [\lambda I - A \quad B] = \text{Rang} \begin{bmatrix} -4 & -1 & -1 & 1 \\ -5 & 1 & 6 & 0 \\ 5 & -1 & -6 & 0 \end{bmatrix} = 2 < 3.$$

Das System ist also nicht steuerbar.

Beispiel 2.32. Das Kontrollproblem

$$\dot{x} = x + [0 \quad 0 \quad \dots \quad 0 \quad 1]^T u$$

mit $x(t) \in \mathbb{R}^n$ und $u(t) \in \mathbb{R}$ ist nicht steuerbar für $n > 1$.

Übungsaufgabe 2.3a

Beispiel 2.33. *Das Kontrollproblem*

$$\dot{x} = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \dots & \dots & -\alpha_{n-1} \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u$$

ist steuerbar.

Übungsaufgabe 2.3b

2.3 Stabilisierbarkeit

Anstelle der Forderung nach Steuerbarkeit kann auch schwächer gefordert werden, dass sich ein System für einen beliebigen Startwert der Nulllage annähern lässt (mit einer geeigneten Kontrolle u).

Definition 2.34. *Ein System*

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0 \quad (2.3)$$

heißt stabilisierbar (*stabilizable*), falls für alle Paare $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$ ein $u \in \mathcal{U}_{ad}$ existiert, sodass für die Lösung x gilt

$$\lim_{t \rightarrow \infty} x(t; u) = 0.$$

Bemerkung 2.35. *Ist die konstante Matrix A Hurwitz, das System also für $u \equiv 0$ asymptotisch stabil, dann ist es trivialerweise stabilisierbar.*

Bemerkung 2.36. *Falls das System steuerbar ist, dann ist es auch stabilisierbar. Stabilisierbarkeit ist also ein schwächeres Konzept als Steuerbarkeit. Betrachte dazu $x_1 = 0$ und eine Kontrolle u_* mit $u_*(t_1) = 0$, sodass $x(t_1) = 0$ für ein t_1 (entsprechend der Steuerbarkeit). Dann definiert man*

$$u(t) = \begin{cases} u_*(t) & t \in [t_0, t_1] \\ 0 & t \geq t_1 \end{cases}.$$

Ohne die Bedingung $u_(t_1) = 0$ ist u zumindest stückweise stetig.*

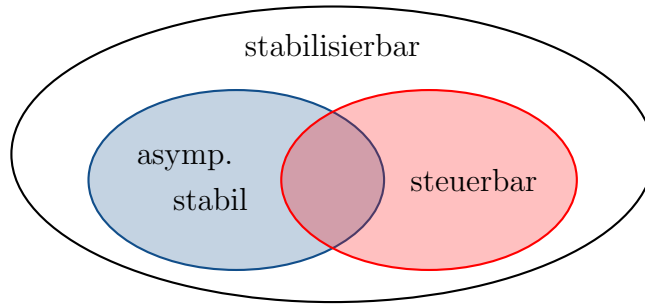


Abbildung 2.2: Zusammenhang zwischen Steuerbarkeit, Stabilisierbarkeit und asymptotischer Stabilität.

Beispiel 2.37. *Tatsächlich sind die stabilisierbaren Systeme eine (echte) Obermenge der asymptotisch stabilen und auch der steuerbaren Systeme, vergleiche Abbildung 2.2. Dazu betrachten wir das System*

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u.$$

Das System ist stabilisierbar, da x_1 steuerbar und x_2 asymptotisch stabil ist. Genauer erhalten wir mit $u(t) = -3x_1(t)$ ein asymptotisch stabiles System. Allerdings ist das Ausgangssystem weder stabil (A hat den Eigenwert 1) noch steuerbar, da

$$\text{Rang } K(A, B) = \text{Rang} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 1.$$

Satz 2.38 (Notwendige Bedingung für LTV Systeme). *Sei das LTV System (2.3) stabilisierbar und $z \not\equiv 0$ eine Lösung der adjungierten Gleichung $\dot{z} = -A(t)^T z$, die für $t \rightarrow \infty$ beschränkt ist. Dann gilt*

$$z(t)^T B(t) \not\equiv 0 \quad \text{auf } [t_0, \infty).$$

Beweis. Wir führen wieder einen Widerspruchsbeweis. Dazu sei $z \not\equiv 0$ eine Lösung der adjungierten Gleichung mit $\lim_{t \rightarrow \infty} \|z(t)\| < \infty$ und $z(t)^T B(t) = 0$ für alle $t \in [t_0, \infty)$. Da z nicht konstant null ist, muss auch $z(t_0) \neq 0$ gelten. Somit existiert ein $x_0 \in \mathbb{R}^n$ mit $x_0^T z(t_0) =: c \neq 0$. Analog zu Bemerkung 2.16 zeigt man, dass $x(t)^T z(t)$ konstant ist, wobei x die Lösung von $\dot{x} = Ax + Bu$

mit Startwert x_0 bezeichnet. Sei nun $u \in \mathcal{U}_{\text{ad}}$ eine stabilisierende Kontrolle für das LTV System mit $x(t_0) = x_0$. Es gilt also

$$\lim_{t \rightarrow \infty} x(t; u) = 0.$$

Da aber $x(t)^T z(t) \equiv c$, müsste $\lim_{t \rightarrow \infty} \|z(t)\| = \infty$ gelten. Dies ist ein Widerspruch zur Konstruktion von z . \square

Eine allgemeine hinreichende Bedingung für LTV Systeme ist nicht bekannt. Im Folgenden soll sich deshalb wieder auf LTI Systeme beschränkt werden.

Satz 2.39. *Das LTI System $\dot{x} = Ax + Bu$ ist stabilisierbar genau dann, wenn eine Matrix $F \in \mathbb{R}^{m,n}$ existiert, sodass $\Re(\lambda) < 0$ für alle $\lambda \in \Lambda(A - BF)$.*

Beweis. Existiert so eine Matrix F , dann wählt man $u(t) = -Fx(t)$. Zum Beweis der anderen Richtung fehlen uns noch die Mittel. Der Beweis wird daher vertagt. (Vergleiche Satz 3.14) \square

Bemerkung 2.40. *Manchmal definiert man Stabilisierbarkeit auch mit der Existenz einer Feedback-Matrix F , sodass $\Lambda(A - BF) \subseteq \mathbb{C}^-$.*

Satz 2.41. *Für ein LTI System der Form $\dot{x} = Ax + Bu$ sind die folgenden Aussagen äquivalent:*

- i) *Das LTI System ist stabilisierbar.*
- ii) *Wenn $A^T z = \lambda z$ für $z \neq 0$ und $\Re(\lambda) \geq 0$ gilt, dann gilt $z^T B \neq 0$.*
- iii) *$\text{Rang} [\lambda I - A \quad B] = n$ für alle $\lambda \in \mathbb{C}$ mit $\Re(\lambda) \geq 0$ (Hautus-Test).*
- iv) *In der dazugehörigen Kalman-Zerlegung gilt $\Lambda(A_3) \subseteq \mathbb{C}^-$.*

Beweis. Wir zeigen i) \Rightarrow ii) \Rightarrow iv) \Rightarrow i) sowie die Äquivalenz ii) \Leftrightarrow iii). [Übungsaufgabe 3.1](#) \square

Beispiel 2.42. *Wir betrachten noch einmal das System aus Beispiel 2.31, welches nicht steuerbar war. Wir zeigen nun, dass es aber trotzdem stabilisierbar ist. Dazu untersuchen wir $[\lambda I - A \quad B]$ für die positiven Eigenwerte von A , also für $\lambda_1 = 2$ und $\lambda_2 = 3$. Dies ergibt*

$$\text{Rang} [\lambda_1 I - A \quad B] = \text{Rang} \begin{bmatrix} 0 & -1 & -1 & 1 \\ -5 & 5 & 6 & 0 \\ 5 & -1 & -2 & 0 \end{bmatrix} = 3$$

und

$$\text{Rang} [\lambda_2 I - A \quad B] = \text{Rang} \begin{bmatrix} 1 & -1 & -1 & 1 \\ -5 & 6 & 6 & 0 \\ 5 & -1 & -1 & 0 \end{bmatrix} = 3.$$

Zudem erhalten wir in der Kalman-Zerlegung $A_3 = [-2]$, also $\Lambda(A_3) \subseteq \mathbb{C}^-$.

Beispiel 2.43. Das Kontrollproblem

$$\dot{x} = x + [0 \quad 0 \quad \dots \quad 0 \quad 1]^T u$$

aus Beispiel 2.32 ist für $n > 1$ auch nicht stabilisierbar.

Übungsaufgabe (optional)

Beispiel 2.44. Wir betrachten noch einmal das Beispiel der Parabolantenne aus Abschnitt 1.3. Das dazugehörige Kontrollproblem ist gegeben in (1.2) und hat die Form $\dot{x} = Ax + Bu$ mit den Matrizen

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{r}{j} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{k}{j} \end{bmatrix}.$$

Die Parameter erfüllen $k, j, r > 0$. Gesucht sind alle stabilisierenden Feedback-Matrizen $F \in \mathbb{R}^{1,2}$. *Übungsaufgabe 3.2*

2.4 Rekonstruierbarkeit und Beobachtbarkeit

Wir betrachten wieder das LTV System (2.1), d.h.

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0, \quad (2.4a)$$

$$y(t) = C(t)x(t) + D(t)u(t). \quad (2.4b)$$

Definition 2.45. Das System heißt rekonstruierbar (*reconstructable*) bzw. beobachtbar (*observable*), falls für alle $t_0 \in \mathbb{R}$ folgendes gilt: Sind $x(t)$ und $\tilde{x}(t)$ zwei Lösungen des Systems für dasselbe $u \in \mathcal{U}_{ad}$ und gilt

$$C(t)x(t) + D(t)u(t) = y(t) = \tilde{y}(t) = C(t)\tilde{x}(t) + D(t)u(t)$$

für alle $t \leq t_0$ bzw. $t \geq t_0$, dann sind die beiden Lösungen gleich, d.h. $x(t) = \tilde{x}(t)$ für alle $t \leq t_0$ bzw. $t \geq t_0$.

Bemerkung 2.46. *Rekonstruierbarkeit ist die Restriktion zur Vergangenheit: Zwei Systeme mit gleichem Ein- und Ausgang in der Vergangenheit haben auch den gleichen Zustand in der Vergangenheit. Analog ist Beobachtbarkeit die Einschränkung auf die Zukunft. Für LTI Systeme sind Rekonstruierbarkeit und Beobachtbarkeit äquivalent wie wir später noch sehen werden.*

Satz 2.47 (Dualität von Steuerbarkeit und Rekonstruierbarkeit). *Das LTV System ist rekonstruierbar genau dann, wenn das duale System*

$$\dot{x} = A(-t)^T x + C(-t)^T \tilde{u}$$

steuerbar ist.

Beweis. Seien $x(t)$ und $\tilde{x}(t)$ Lösungen des LTV Systems zur selben Kontrolle u . Die Differenz $z(t) := x(t) - \tilde{x}(t)$ löst dann die Gleichung $\dot{z} = A(t)z$. Wir stellen eine Reihe von Äquivalenzen auf.

Das System ist rekonstruierbar.

\Leftrightarrow Für alle t_0 gilt: Ist $C(t)z(t) \equiv 0$ auf $(-\infty, t_0]$, dann ist $z \equiv 0$ für $t \leq t_0$.

\Leftrightarrow Für alle t_0 gilt: Ist $z \neq 0$ Lösung von $\dot{z} = A(t)z$ auf $(-\infty, t_0]$,
dann gilt $C(t)z(t) \neq 0$ auf $(-\infty, t_0]$.

\Leftrightarrow Für alle t_0 gilt: Ist $z \neq 0$ Lösung von $\dot{z} = -A(-t)z$ auf $[t_0, \infty)$,
dann gilt $C(-t)z(t) \neq 0$ auf $[t_0, \infty)$.

\Leftrightarrow Das System $\dot{x} = A(-t)^T x + C(-t)^T \tilde{u}$ ist steuerbar.

Im letzten Schritt haben wir die Charakterisierung von Steuerbarkeit aus Satz 2.17 benutzt. Diese besagt, dass $\dot{x} = A(t)x + B(t)u$ genau dann steuerbar ist, wenn jede Lösung z der adjungierten Gleichung folgende Implikation erfüllt: Gilt $z \neq 0$, dann auch $B(t)^T z(t) \neq 0$. \square

Wir wollen einige Kriterien für die verschiedenen Begriffe in einem Satz zusammenfassen. Davor geben wir noch eine Definition.

Definition 2.48. *Die Matrix*

$$Q(t_0, t_1) := \int_{t_0}^{t_1} \Phi(s, t_1)^T C(s)^T C(s) \Phi(s, t_1) ds$$

heißt die Beobachtbarkeits-Gramsche (observability Gramian) des LTV Systems.

Satz 2.49. *Das LTV System ist*

- i) steuerbar $\Leftrightarrow \forall t_0 \exists t_1: \int_{t_0}^{t_1} \Phi(t_0, s)B(s)B(s)^T\Phi(t_0, s)^T ds > 0,$
- ii) rekonstruierbar $\Leftrightarrow \dot{z} = A(-t)^T z + C(-t)^T \tilde{u}$ ist steuerbar,
 $\Leftrightarrow \forall t_1 \exists t_0: Q(t_0, t_1) > 0,$
- iii) beobachtbar $\Leftrightarrow \dot{z} = -A(t)^T z + C(t)^T \tilde{u}$ ist steuerbar,
 $\Leftrightarrow \forall t_0 \exists t_1: Q(t_0, t_1) > 0.$

Beweis. i) Diese Äquivalenz ist Teil von Satz 2.17.

ii) Die erste Äquivalenz ist die Dualität aus Satz 2.47. Nun nutzen wir die Charakterisierung mithilfe der Steuerbarkeits-Gramschen aus i). Das System $\dot{x} = A^T(-t)x$ hat die Fundamentallösung $\Phi(-s, -t)^T$, denn

$$\frac{d}{dt} [\Phi(-s, -t)^T] = A(-t)^T \Phi(-s, -t)^T, \quad \Phi(-s, -s)^T = I.$$

Zu jedem t_0 existiert also ein t_1 , sodass $W(t_0, t_1) > 0$ mit

$$\begin{aligned} W(t_0, t_1) &= \int_{t_0}^{t_1} \Phi(-s, -t_0)^T C(-s)^T C(-s) \Phi(-s, -t_0) ds \\ &= \int_{-t_1}^{-t_0} \Phi(s, -t_0)^T C(s)^T C(s) \Phi(s, -t_0) ds. \end{aligned}$$

Mit $\tilde{t}_0 := -t_1$ und $\tilde{t}_1 := -t_0$ existiert nun also zu jedem \tilde{t}_0 ein \tilde{t}_1 , sodass $Q(\tilde{t}_0, \tilde{t}_1) > 0$.

iii) **Übungsaufgabe (optional)** □

Der Satz oben zeigt, dass Beobachtbarkeit und Rekonstruierbarkeit sehr eng miteinander verbunden sind. Unterschiedlich sind lediglich die Quantoren.

Bemerkung 2.50. *Zwischen dem Eingang u und dem Ausgang y besteht im Fall $D = 0$ mit $x(t_1) = x_1$ und $t_0 \leq t_1$ die Beziehung*

$$\begin{aligned} &\int_{t_0}^{t_1} \Phi(t, t_1)^T C(t)^T y(t) dt \\ &= Q(t_0, t_1)x_1 + \int_{t_0}^{t_1} \left(\Phi(t, t_1)^T C(t)^T C(t) \int_{t_1}^t \Phi(t, s)B(s)u(s) ds \right) dt. \end{aligned}$$

Dies impliziert die Eindeutigkeit von x_1 für rekonstruierbare Systeme, da dann $Q(t_0, t_1) > 0$ gilt. Übungsaufgabe (optional)

Im Folgenden wird wieder das LTI System betrachtet. In diesem Fall vereinfachen sich viele Dinge. Zunächst geben wir aber noch eine Definition.

Definition 2.51. *Die Matrix*

$$O(A, C) := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

heißt Beobachtbarkeits-Matrix (*observability matrix*).

Satz 2.52. *Im zeitinvarianten Fall sind die folgenden Aussagen äquivalent:*

- i) *Das LTI System ist rekonstruierbar.*
- ii) *Das LTI System ist beobachtbar.*
- iii) *Das duale System $\dot{z} = A^T z + C^T \tilde{u}$ ist steuerbar.*
- iv) *Die Beobachtbarkeits-Matrix $O(A, C)$ hat Rang n .*
- v) *Wenn $v \neq 0$ ein Eigenvektor von A ist, dann gilt $Cv \neq 0$.*
- vi) *Der Rang von $\begin{bmatrix} \lambda I - A \\ C \end{bmatrix}$ ist gleich n für alle $\lambda \in \mathbb{C}$.*

Beweis. Folgt aus der Dualität. Für die Äquivalenz mit ii) wäre im dualen System $A(t)$ durch $A(-t)$ ersetzt. Dies ändert jedoch nichts. \square

Beispiel 2.53. *Wir betrachten noch einmal das invertierte Pendel aus Beispiel 2.30 bzw. Abschnitt 1.2 mit*

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x = \begin{bmatrix} \phi \\ \dot{\phi} \end{bmatrix}.$$

Angenommen, wir messen nur den Winkel ϕ , d.h. $C = [1 \ 0]$. Ist das System dann beobachtbar bzw. rekonstruierbar? Dies sollte so sein, da, wenn man ϕ für alle t kennt, dann kennt man auch die Ableitung $\dot{\phi}$. Für die Beobachtbarkeits-Matrix gilt

$$\text{Rang } O(A, C) = \text{Rang} \begin{bmatrix} C \\ CA \end{bmatrix} = \text{Rang} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 2.$$

Somit ist das System beobachtbar und rekonstruierbar.

Bemerkung 2.54. Aus dem Kapitel über die Steuerbarkeit kennen wir die Kalman-Zerlegung. Wir wenden diese auf (A^T, C^T) an und erhalten

$$W^T A^T W = \begin{bmatrix} A_1^T & A_2^T \\ 0 & A_3^T \end{bmatrix}, \quad W^T C^T = \begin{bmatrix} C_1^T \\ 0 \end{bmatrix}.$$

Dabei ist die Matrix W orthogonal und $\text{Rang } K(A^T, C^T) = \text{Rang } O(A, C) = \text{Rang } O(A_1, C_1)$. Damit ist gezeigt, dass es ein orthogonales W gibt, sodass

$$W^T A W = \begin{bmatrix} A_1 & 0 \\ A_2 & A_3 \end{bmatrix}, \quad C W = [C_1 \ 0].$$

Dabei ist (A_1, C_1) beobachtbar. Ein Basiswechsel $\tilde{x} = W^T x = [\tilde{x}_1^T, \tilde{x}_2^T]^T$ liefert das System

$$\begin{aligned} \dot{\tilde{x}}_1 &= A_1 \tilde{x}_1 + B_1 u, \\ \dot{\tilde{x}}_2 &= A_2 \tilde{x}_1 + A_3 \tilde{x}_2 + B_2 u, \\ y &= C_1 \tilde{x}_1 + D u. \end{aligned}$$

Somit hat \tilde{x}_2 also keinen Einfluss auf den Ausgang y . Die Zustände \tilde{x}_2 heißen unbeobachtbar (*unobservable*), die Eigenwerte von A_3 heißen unbeobachtbare Moden (*unobservable modes*).

Korollar 2.55. Seien $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$ und $C \in \mathbb{R}^{p,n}$. Dann existiert eine orthogonale Matrix $V \in \mathbb{R}^{n,n}$, sodass

$$V^T A V = \begin{bmatrix} A_{c\bar{o}} & A_{12} & A_{13} & A_{14} \\ 0 & A_{co} & A_{23} & A_{24} \\ 0 & 0 & A_{\bar{c}\bar{o}} & A_{34} \\ 0 & 0 & 0 & A_{\bar{c}o} \end{bmatrix}, \quad V^T B = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \quad C V = [0 \ C_2 \ 0 \ C_4],$$

wobei (A_{co}, B_2, C_2) steuerbar und beobachtbar ist.

Beweis. Übungsaufgabe 3.3 □

Definition 2.56 (Duales Konzept für Stabilisierbarkeit). Ein LTI System heißt entdeckbar (*detectable*), falls für jede Lösung z von $\dot{z} = Az$ mit $Cz(t) \equiv 0$ für $t \geq t_0$ gilt, dass

$$\lim_{t \rightarrow \infty} z(t) = 0.$$

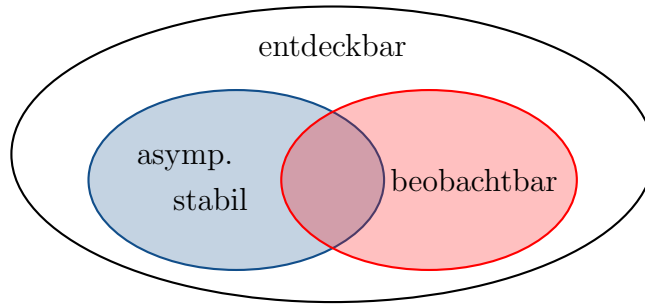


Abbildung 2.3: Zusammenhang zwischen Beobachtbarkeit, Entdeckbarkeit und asymptotischer Stabilität.

Wie kann man diese Definition interpretieren? Statt wie bei der Beobachtbarkeit zu fordern, dass gleiche Eingänge und $Cx \equiv C\tilde{x}$ schon $x \equiv \tilde{x}$ impliziert, wird jetzt nur gefordert, dass

$$\lim_{t \rightarrow \infty} x(t) = \lim_{t \rightarrow \infty} \tilde{x}(t)$$

gilt. Das System ist also nicht zwangsläufig beobachtbar, aber der unbeobachtbare Teil ist asymptotisch stabil und klingt ab. Analog zur Steuerbarkeit ergibt sich ein Schema, vgl. Abbildung 2.3.

Satz 2.57. *Für LTI Systeme sind die folgenden Aussagen äquivalent:*

- i) *Das System ist entdeckbar.*
- ii) *Das duale System $\dot{z} = A^T z + C^T \tilde{u}$ ist stabilisierbar.*
- iii) *Wenn $x \neq 0$ ein Eigenvektor von A zum Eigenwert λ mit $\Re(\lambda) \geq 0$ ist, dann gilt $Cx \neq 0$.*
- iv) *$\text{Rang} \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n$ für alle $\lambda \in \mathbb{C}$ mit $\Re(\lambda) \geq 0$.*
- v) *In der Kalman-Zerlegung von (A^T, C^T) gilt $\Lambda(A_3) \subseteq \mathbb{C}^-$.*

Beweis. Die Punkte ii) - v) sind äquivalent nach Satz 2.41. Wir zeigen nun i) \Rightarrow iii). Sei $x \neq 0$ ein Eigenvektor von A zum Eigenwert λ mit $\Re(\lambda) \geq 0$. Dann löst $z(t) := e^{\lambda t} x$ die Gleichung $\dot{z} = Az$. Angenommen es wäre $Cx = 0$. Dann wäre auch $Cz(t) \equiv 0$ und daher $\lim_{t \rightarrow \infty} z(t) = 0$. Dies ist nun aber ein Widerspruch zu $\Re(\lambda) \geq 0$.

Schließlich zeigen wir noch die Rückrichtung iii) \Rightarrow i). Sei z eine Lösung von $\dot{z} = Az$ mit $Cz \equiv 0$ für alle $t \geq t_0$. Angenommen das System ist nicht entdeckbar, d.h., $\lim_{t \rightarrow \infty} z(t) \neq 0$. Dann muss die Matrix A mindestens einen Eigenwert $\lambda \in \mathbb{C}$ enthalten mit $\Re(\lambda) \geq 0$. Diese Mode ist auch Teil der Lösung z , also enthält z einen Teil $c(t)e^{\lambda t}x$, wobei $x \neq 0$ die Gleichung $Ax = \lambda x$ erfüllt. Die Annahme $Cz \equiv 0$ impliziert $Cx = 0$, was ein Widerspruch zu iii) ist. Es folgt $\lim_{t \rightarrow \infty} z(t) = 0$. \square

2.5 Der Staircase Algorithmus

Die meisten System-Eigenschaften wie zum Beispiel Steuerbarkeit, Beobachtbarkeit, etc. sind schwierig (d.h. numerisch aufwendig) zu prüfen. Ziel ist nun, eine Transformationen zu finden, die die System-Eigenschaften erhalten, aber deren Überprüfung erleichtern.

Definition 2.58. *Zwei LTI-Systeme*

$$\begin{aligned} \dot{x} &= Ax + Bu, & \dot{\tilde{x}} &= \tilde{A}\tilde{x} + \tilde{B}\tilde{u}, \\ y &= Cx + Du, & \tilde{y} &= \tilde{C}\tilde{x} + \tilde{D}\tilde{u} \end{aligned}$$

heißen äquivalent (*feedback equivalent*), falls das zweite aus dem ersten hervorgeht durch eine endliche Zahl von Transformationen der folgenden Typen:

i) Basiswechsel mit invertierbaren Matrizen $P \in \mathbb{R}^{n,n}$, $Q \in \mathbb{R}^{m,m}$, $R \in \mathbb{R}^{p,p}$,

$$x \mapsto Px, \quad u \mapsto Qu, \quad y \mapsto Ry,$$

ii) lineare Zustandsrückführung (*linear state feedback*)

$$u \mapsto -Fx + v, \quad F \in \mathbb{R}^{m,n}, \quad v \in \mathcal{U}_{ad}.$$

Werden nur Basiswechsel verwendet, dann heißen die Systeme stark äquivalent.

Beispiel 2.59. *Ein Basiswechsel $x \mapsto Px =: \tilde{x}$ führt auf das System*

$$\begin{aligned} P\dot{x} &= (PAP^{-1})Px + PBu & \Leftrightarrow & \dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}\tilde{u}, \\ y &= (CP^{-1})Px + Du & \Leftrightarrow & y = \tilde{C}\tilde{x} + Du \end{aligned}$$

mit $\tilde{A} = PAP^{-1}$, $\tilde{B} = PB$, $\tilde{C} = CP^{-1}$. Eine lineare Zustandsrückführung $u \mapsto -Fx + \tilde{u}$ führt auf das äquivalente System

$$\begin{aligned} \dot{x} &= Ax + B(-Fx + \tilde{u}) & \Leftrightarrow & \dot{x} = \tilde{A}x + B\tilde{u}, \\ y &= Cx + D(-Fx + \tilde{u}) & \Leftrightarrow & y = \tilde{C}x + D\tilde{u} \end{aligned}$$

mit $\tilde{A} = A - BF$ und $\tilde{C} = C - DF$.

Die Transformationen kann man auch als lineare Operation auf einer Blockmatrix betrachten. Dazu definieren wir

$$\mathcal{A} := \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad \tilde{\mathcal{A}} := \begin{bmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{bmatrix} \in \mathbb{R}^{n+p, n+m}.$$

Zwei Systeme sind dann äquivalent mit $\tilde{x} = Px$, $\tilde{y} = Ry$ und $\tilde{u} = Qu + Fx$, wenn es invertierbare Matrizen P , Q und R sowie eine Matrix F gibt, sodass

$$\tilde{\mathcal{A}} = \begin{bmatrix} P & 0 \\ 0 & R \end{bmatrix} \mathcal{A} \begin{bmatrix} P^{-1} & 0 \\ -Q^{-1}FP^{-1} & Q^{-1} \end{bmatrix} \quad \text{bzw.} \quad \mathcal{A} = \begin{bmatrix} P^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix} \tilde{\mathcal{A}} \begin{bmatrix} P & 0 \\ F & Q \end{bmatrix}.$$

Zudem sind diese stark äquivalent, wenn $F = 0$ gilt.

Bemerkung 2.60. Die hier definierte System-Äquivalenz ist tatsächlich eine Äquivalenzrelationen auf dem Raum der LTI Systeme.

Satz 2.61. Sei das LTI-System steuerbar (stabilisierbar). Dann ist auch jedes äquivalente System steuerbar (stabilisierbar).

Beweis. Übungsaufgabe 4.1a □

Satz 2.62. Sei das LTI-System beobachtbar (rekonstruierbar, entdeckbar, asymptotisch stabil). Dann ist auch jedes stark äquivalente System beobachtbar (rekonstruierbar, entdeckbar, asymptotisch stabil).

Beweis. Übungsaufgabe 4.1b □

Beispiel 2.63. Zustandsrückführungen können die Beobachtbarkeit (bzw. Rekonstruierbarkeit) und Entdeckbarkeit ändern. Wir betrachten das System mit den Matrizen

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = [0 \quad 1], \quad D = [0 \quad 0].$$

Das System ist beobachtbar (und damit auch entdeckbar), da

$$O(A, C) = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

vollen Rang hat. Mit $F = A$ erhalten wir $\tilde{A} = A - BF = 0$ und $\tilde{C} = C - DF = C$. Für das transformierte System gilt nun

$$O(\tilde{A}, \tilde{C}) = \begin{bmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Diese Matrix hat keinen vollen Rang und damit ist das transformierte System nicht beobachtbar. Das transformierte System ist auch nicht entdeckbar, da $v := [1, 0]^T$ ein Eigenvektor von \tilde{A} zu $\lambda = 0$ ist aber $Cv \equiv 0$.

Neben Basiswechsel und linearer Zustandsrückführung sind auch andere Transformationen möglich wie die folgenden Bemerkungen zeigen.

Bemerkung 2.64. Analog zur linearen Zustandsrückführung kann man auch die lineare Ausgangsrückführung (*linear output feedback*) betrachten, d.h., $u \mapsto -Gy + v$ für eine Matrix $G \in \mathbb{R}^{m,p}$ und eine Kontrolle $v \in \mathcal{U}_{ad}$. Im Fall $D = 0$, also $y = Cx$, kann man dies aber wieder schreiben als $u \mapsto -GCx + v$.

Bemerkung 2.65. Eine Rückführung der Zustandsableitung (*linear derivative feedback*) $u \mapsto -F\dot{x} + v$ kann allerdings die LTI-Struktur zerstören. Dies ist wie folgt zu sehen:

$$\dot{x} = Ax + B(-F\dot{x} + v) \quad \Leftrightarrow \quad (I + BF)\dot{x} = Ax + Bv.$$

Dies ist kein LTI System, falls $E := I + BF$ singulär ist. In dem Fall erhalten wir eine Differential-algebraische Gleichung (DAE) bzw. ein Deskriptor-System.

Wir wollen nun eine Transformation

$$(A, B) \mapsto (PAP^T, PBQ) = (\tilde{A}, \tilde{B})$$

bestimmen mit orthogonalen Matrizen P und Q , sodass die Steuerbarkeit des Systems leicht ablesbar ist. Wir fordern hier orthogonale Matrizen, damit die Transformationen numerisch rückwärts stabil (*backward stable*) sind. Daher werden wir die SVD aus Kapitel 1.1.3 anwenden. Zudem nutzen wir die Charakterisierung der Steuerbarkeit, die besagt, dass (A, B) steuerbar ist genau dann, wenn $\text{Rang} \begin{bmatrix} \lambda I - A & B \end{bmatrix} = n$ für alle $\lambda \in \mathbb{C}$.

Satz 2.66 (Staircase-Form). Seien $A \in \mathbb{R}^{n,n}$ und $B \in \mathbb{R}^{n,m}$ gegeben. Dann existieren orthogonale Matrizen $P \in \mathbb{R}^{n,n}$, $Q \in \mathbb{R}^{m,m}$ und Zahlen n_i mit

$$m \geq n_1 \geq n_2 \geq \dots \geq n_{s-1} > 0, \quad n_s \geq 0,$$

sodass

$$PAP^T = \left[\begin{array}{cccc|cc} A_{11} & A_{12} & A_{13} & \dots & A_{1s-1} & A_{1s} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2s-1} & A_{2s} \\ 0 & A_{32} & A_{33} & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & A_{s-1s-2} & A_{s-1s-1} & A_{s-1s} \\ \hline 0 & 0 & 0 & 0 & 0 & A_{ss} \end{array} \right], \quad PBQ = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}.$$

Dabei gilt $A_{ij} \in \mathbb{R}^{n_i, n_j}$ und die Subdiagonalblöcke haben die Form

$$A_{i+1,i} = \begin{bmatrix} \Sigma_{i+1,i} & 0 \end{bmatrix} \in \mathbb{R}^{n_{i+1}, n_i},$$

wobei die $\Sigma_{i+1,i}$ invertierbar sind für $i = 1, \dots, s-2$. Die Matrix $\Sigma_{s-1, s-2}$ ist sogar diagonal. Zudem ist auch die Matrix $B_1 \in \mathbb{R}^{n_1, n_1}$ invertierbar.

Beweis. Der Beweis ist konstruktiv und folgt dem Staircase-Algorithmus (Van Dooren, 1979).

- 1.) Berechne die SVD von B ,

$$B = U_B \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix} V_B^T, \quad \Sigma_B \in \mathbb{R}^{n_1, n_1} \text{ invertierbar.}$$

Setze $P := U_B^T$, $Q := V_B$, $\tilde{n} := n - n_1$ und überschreibe

$$B := PBQ = \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix}, \quad A := PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

- 2.) Ist $\tilde{n} = 0$ oder $A_{21} = 0$, dann STOP. Andernfalls berechne die SVD von $A_{21} \in \mathbb{R}^{\tilde{n}, n_1}$, d.h., finde $U_{21} \in \mathbb{R}^{\tilde{n}, \tilde{n}}$ und $V_{21} \in \mathbb{R}^{n_1, n_1}$ orthogonal, sodass

$$A_{21} = U_{21} \begin{bmatrix} \Sigma_{21} & 0 \\ 0 & 0 \end{bmatrix} V_{21}^T \quad \text{mit } \Sigma_{21} \in \mathbb{R}^{n_2, n_2} \text{ invertierbar.}$$

Da $A_{21} \neq 0$ ist, gilt $n_2 \geq 1$. Offensichtlich gilt zudem $n_2 \leq n_1$. Überschreibe nun

$$P_2 := \begin{bmatrix} V_{21}^T & 0 \\ 0 & U_{21}^T \end{bmatrix}, \quad \tilde{n} := \tilde{n} - n_2, \quad P := P_2 P$$

sowie

$$A := P_2 A P_2^T = \begin{bmatrix} V_{21}^T A_{11} V_{21} & V_{21}^T A_{12} U_{21} \\ U_{21}^T A_{21} V_{21} & U_{21}^T A_{22} U_{21} \end{bmatrix} = \left[\begin{array}{c|cc} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{array} \right]$$

mit $A_{11} \in \mathbb{R}^{n_1, n_1}$, $A_{22} \in \mathbb{R}^{n_2, n_2}$, $A_{33} \in \mathbb{R}^{\tilde{n}, \tilde{n}}$ und $A_{21} = [\Sigma_{21} \ 0]$. Man beachte, dass auch die Teilmatrizen A_{ij} überschrieben wurden. Überschreibe nun auch

$$B := P_2 B = \begin{bmatrix} V_{21}^T \Sigma_B & 0 \\ 0 & 0 \end{bmatrix} =: \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Man beachte, dass $\text{Rang } B_1 = \text{Rang } B = n_1$.

- 3.) Ist $\tilde{n} = 0$ oder $A_{32} = 0$, dann STOP. Andernfalls berechne die SVD von A_{32} , d.h., finde $U_{32} \in \mathbb{R}^{\tilde{n}, \tilde{n}}$ und $V_{32} \in \mathbb{R}^{n_2, n_2}$ orthogonal, sodass

$$A_{32} = U_{32} \begin{bmatrix} \Sigma_{32} & 0 \\ 0 & 0 \end{bmatrix} V_{32}^T \quad \text{mit } \Sigma_{32} \in \mathbb{R}^{n_3, n_3} \text{ invertierbar.}$$

Es gilt also $n_3 \leq n_2$. Setze

$$P_3 := \begin{bmatrix} I_{n_1} & 0 & 0 \\ 0 & V_{32}^T & 0 \\ 0 & 0 & U_{32}^T \end{bmatrix}, \quad P := P_3 P, \quad \tilde{n} := \tilde{n} - n_3$$

und

$$A := P_3 A P_3^T = \left[\begin{array}{cc|cc} A_{11} & A_{12} & A_{13} & A_{14} \\ \hline A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & A_{32} & A_{33} & A_{34} \\ 0 & 0 & A_{43} & A_{44} \end{array} \right], \quad A_{32} = [\Sigma_{32} \ 0].$$

Die Matrix B verändert sich nicht mehr, da

$$B := P_3 B = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Zudem hat $A_{21} = [V_{32}^T \Sigma_{21} \ 0]$ weiterhin vollen Rang.

- 4.) Weitere Schritte laufen nach dem gleichen Schema ab: Ist $\tilde{n} = 0$ oder $A_{i,i-1} = 0$, dann STOP. Andernfalls berechne die SVD von A_{ii-1} und aktualisiere die Matrizen P und A .

Der Algorithmus stoppt nach einer endlichen Anzahl von Schritten (s), da \tilde{n} in jedem Schritt kleiner wird. Maximal werden n Schritte benötigt. \square

Bemerkung 2.67. *Der Algorithmus ist numerisch stabil, aber Rang-Entscheidungen für die Matrizen $A_{i+1,i}$ müssen getroffen werden und sind kritisch. Akkumulierende Rundungsfehler können zu falschen Rang-Entscheidungen führen.*

Bemerkung 2.68. *Es gibt maximal n Schritte mit je $\mathcal{O}(n^3)$ Flops, also insgesamt $\mathcal{O}(n^4)$ Flops. Zudem müssen die Matrizen A, B, P, Q gespeichert werden (A und B werden überschrieben).*

Korollar 2.69. *Ein LTI System ist genau dann*

- i) steuerbar, wenn $n_s = 0$ in der Staircase-Form von (A, B) ,*
- ii) stabilisierbar, wenn $\Lambda(A_{ss}) \subseteq \mathbb{C}^-$ in der Staircase-Form von (A, B) ,*
- iii) beobachtbar, wenn $n_s = 0$ in der Staircase-Form von (A^T, C^T) und*
- iv) entdeckbar, wenn $\Lambda(A_{ss}) \subseteq \mathbb{C}^-$ in der Staircase-Form von (A^T, C^T) .*

Beweis. Für i) und ii) sei (A, B) o.B.d.A. in Staircase-Form. Dann gilt

$$\begin{aligned}
 & [B \quad \lambda I - A] \\
 &= \begin{bmatrix} B_1 & 0 & \lambda I - A_{11} & \dots & \dots & \dots & -A_{1s} \\ 0 & 0 & -A_{21} & \ddots & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & & 0 & -A_{s-1s-2} & \lambda I - A_{s-1s-1} & \vdots \\ 0 & 0 & \dots & \dots & 0 & 0 & \lambda I - A_{ss} \end{bmatrix},
 \end{aligned}$$

wobei $B_1, A_{21}, \dots, A_{s-1,s-2}$ vollen Rang haben. Also gilt

$$\text{Rang} [\lambda I - A \quad B] = n \quad \Leftrightarrow \quad \text{Rang}(\lambda I - A_{ss}) = n_s.$$

Daraus folgen die Aussagen mit den bekannten Tests: Die Matrix $\lambda I - A_{ss}$ hat vollen Rang für alle $\lambda \in \mathbb{C}$ genau dann, wenn es sich um die Nullmatrix

handelt, also wenn $n_s = 0$. Betrachtet man nur λ mit $\Re(\lambda) \geq 0$, dann hat $\lambda I - A_{ss}$ genau dann vollen Rang, wenn $\Lambda(A_{ss}) \subseteq \mathbb{C}^-$. Die Aussagen iii) und iv) folgen durch Dualität. \square

Definition 2.70. *Spezialfälle sind gegeben durch $m = 1$, also nur ein Eingang (single input), oder $p = 1$, nur ein Ausgang (single output). Den Fall $m = p = 1$ bezeichnet man als SISO (single-input-single-output). Sind $m > 1$ und $p > 1$, so nennt man das System MIMO (multi-input-multi-output).*

Definition 2.71. *Eine Matrix $A = [a_{ij}] \in \mathbb{R}^{n,n}$ heißt Hessenberg, wenn für alle $i > j + 1$ gilt, dass $a_{ij} = 0$ ist. Unterhalb der Subdiagonalen sind also alle Einträge gleich 0. Die Matrix heißt unreduziert Hessenberg (unreduced Hessenberg matrix), wenn $a_{ii-1} \neq 0$ für alle $i = 2, \dots, n$ gilt.*

Satz 2.72 (System-Hessenberg-Form). *Für Matrizen $A \in \mathbb{R}^{n,n}$ und $B \in \mathbb{R}^n$, also $m = 1$, existiert eine orthogonale Matrix P , sodass*

$$PAP^T = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ a_{21} & \cdots & \cdots & a_{2n} \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots & a_{nn-1} & a_{nn} \end{bmatrix}, \quad PB = \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Beweis. Wie in der Staircase-Form mit $n_1 = n_2 = \dots = n_{s-1} = 1$. Im Fall $n_s \neq 0$ muss man sich noch um die Matrix A_{ss} kümmern, d.h., man muss diese in Hessenberg-Form bringen. Wir nutzen, dass es eine orthogonale Matrix U_s gibt, sodass $U_s A_{ss} U_s^T$ in Hessenberg-Form ist. \square

Korollar 2.73. *Ein single-input LTI System ist genau dann steuerbar, wenn die System-Hessenberg-Form unreduziert ist sowie $b_1 \neq 0$ gilt.*

3 Stabilisierung und Lyapunov-Gleichungen

Gegeben sei ein LTI-System

$$\dot{x} = Ax + Bu, \quad x(0) = x_0.$$

Das Ziel ist, eine stabilisierende Feedback-Matrix F zu finden, d.h. ein $F \in \mathbb{R}^{m,n}$, sodass $\Lambda(A - BF) \subseteq \mathbb{C}^-$ gilt. Mit $u := -Fx$ sind dann alle Lösungen von

$$\dot{x} = Ax + Bu = (A - BF)x$$

asymptotisch stabil. Dazu betrachten wir 2 Methoden:

- i) direkte Methode mittels Lyapunov-Gleichungen:
Es müssen keine Lösungen von DGLn berechnet werden.
- ii) Polvorgabe (**pole placement**):
Zu gegebenem $\mathcal{P} = \{\mu_1, \dots, \mu_n\}$ wird $F \in \mathbb{R}^{m,n}$ konstruiert, sodass $\Lambda(A - BF) = \mathcal{P}$ gilt. Eine Wahl $\mathcal{P} \subseteq \mathbb{C}^-$ löst dann das Problem.

3.1 Sylvester- und Lyapunov-Gleichungen

Wir betrachten zunächst die *Sylvester-Gleichung*

$$AX + XB = W \tag{3.1}$$

mit $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{m,m}$, $W \in \mathbb{R}^{n,m}$ gegeben und $X \in \mathbb{R}^{n,m}$ gesucht. Dies ist eine lineare Gleichung in den nm Unbekannten x_{ij} , d.h. es existiert eine äquivalente Repräsentation

$$Mx = b$$

mit $M \in \mathbb{R}^{nm,nm}$ und $b \in \mathbb{R}^{nm}$.

Definition 3.1. Für $A \in \mathbb{R}^{m,n}$ und $B \in \mathbb{R}^{k,\ell}$ ist durch

$$A \otimes B := \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{bmatrix} \in \mathbb{R}^{mk,n\ell}$$

das Kronecker-Produkt von A und B definiert.

Lemma 3.2. Das Kronecker-Produkt erfüllt für entsprechende Dimensionen folgende Eigenschaften:

- i) $(\alpha A) \otimes B = \alpha(A \otimes B) = A \otimes (\alpha B)$,
- ii) $(A + B) \otimes C = A \otimes C + B \otimes C$,
- iii) $A \otimes (B + C) = A \otimes B + A \otimes C$,
- iv) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$,
- v) $(A \otimes B)^T = A^T \otimes B^T$,
- vi) $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$,
- vii) $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,

$$viii) \Lambda(A \otimes B) = \{\lambda_i \mu_j \mid \lambda_i \in \Lambda(A), \mu_j \in \Lambda(B)\},$$

$$ix) \det(A \otimes B) = (\det A)^m (\det B)^n \text{ für } A \in \mathbb{R}^{n,n}, B \in \mathbb{R}^{m,m},$$

$$x) \text{Rang}(A \otimes B) = \text{Rang } A \cdot \text{Rang } B.$$

Beweis. Übungsaufgabe 4.2a □

Definition 3.3. Für $A = [a_{ij}] \in \mathbb{R}^{m,n}$ wird durch $\text{vec}: \mathbb{R}^{m,n} \rightarrow \mathbb{R}^{mn}$,

$$\text{vec}(A) := [a_{11}, \dots, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{1n}, \dots, a_{mn}]^T \in \mathbb{R}^{mn}$$

die Vektorisierung von A definiert. Diese „klebt“ die Spalten von A zu einem Vektor zusammen.

Lemma 3.4. Das Kronecker-Produkt und die Vektorisierung erfüllen für Matrizen A, B, X und einem Vektor b entsprechender Dimensionen

$$i) \text{vec}(A + B) = \text{vec}(A) + \text{vec}(B),$$

$$ii) \text{vec}(AXB) = (B^T \otimes A) \text{vec}(X),$$

$$iii) \text{vec}(AX) = (I \otimes A) \text{vec}(X),$$

$$iv) \text{vec}(Xb) = (b^T \otimes I) \text{vec}(X),$$

$$v) \|(I \otimes A) \text{vec}(I)\|_2 = \|(A \otimes I) \text{vec}(I)\|_2 = \|\text{vec}(A)\|_2 = \|A\|_F,$$

$$vi) \|A \otimes B\|_2 = \|A\|_2 \|B\|_2.$$

Beweis. Übungsaufgabe 4.2b □

Wendet man diese Vektorisierung auf die Sylvester-Gleichung $AX + XB = W$ an, so ergibt sich das lineare Gleichungssystem

$$\underbrace{(I \otimes A + B^T \otimes I)}_{=:M} \underbrace{\text{vec}(X)}_{=:x} = \text{vec}(AXI_m) + \text{vec}(I_n XB) \\ = \text{vec}(AX + XB) = \text{vec}(W) =: b.$$

Diese Formulierung ist sehr nützlich, um die Lösbarkeit der Sylvester-Gleichung zu untersuchen.

Satz 3.5. Die Sylvester-Gleichung (3.1) besitzt eine eindeutige Lösung genau dann, wenn

$$\Lambda(A) \cap \Lambda(-B) = \emptyset.$$

Beweis. Durch die Vektorisierung der Sylvester-Gleichung ist klar, dass es eindeutige Lösung gibt genau dann, wenn M invertierbar ist, also $0 \notin \Lambda(M)$. Seien $A = TJ_AT^{-1}$ und $B^T = SJ_BS^{-1}$ mit

$$J_A = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad J_B = \begin{bmatrix} \mu_1 & & * \\ & \ddots & \\ & & \mu_m \end{bmatrix}$$

die Jordan-Normalformen von A und B . Dann gilt mit Lemma 3.2,

$$\begin{aligned} M &= I \otimes A + B^T \otimes I \\ &= I \otimes (TJ_AT^{-1}) + (SJ_BS^{-1}) \otimes I \\ &= (SS^{-1}) \otimes (TJ_AT^{-1}) + (SJ_BS^{-1}) \otimes (TT^{-1}) \\ &= (S \otimes T)(I \otimes J_A)(S^{-1} \otimes T^{-1}) + (S \otimes T)(J_B \otimes I)(S^{-1} \otimes T^{-1}) \\ &= (S \otimes T)(I \otimes J_A + J_B \otimes I)(S^{-1} \otimes T^{-1}). \end{aligned}$$

Somit ist M ähnlich zu $I \otimes J_A + J_B \otimes I$. Weiter gilt

$$I \otimes J_A + J_B \otimes I = \text{diag}(J_A, J_A, \dots, J_A) + \begin{bmatrix} \mu_1 I & & * \\ & \ddots & \\ & & \mu_m I \end{bmatrix}.$$

Diese Matrix hat die Diagonaleinträge $\lambda_k + \mu_\ell$ für $k = 1, \dots, n$, $\ell = 1, \dots, m$, dh.,

$$\Lambda(M) = \{\lambda_k + \mu_\ell \mid \lambda_k \in \Lambda(A), \mu_\ell \in \Lambda(B)\}.$$

Also ist M invertierbar genau dann, wenn A und $-B$ keine gemeinsamen Eigenwerte besitzen, also $\Lambda(A) \cap \Lambda(-B) = \emptyset$. \square

Definition 3.6. Für eine Matrix $X \in \mathbb{R}^{n,n}$ schreiben wir $X > 0$, wenn diese positiv definit ist und $X \geq 0$, wenn sie positiv semidefinit ist.

Satz 3.7. Seien $A \in \mathbb{R}^{n,n}$ und $B \in \mathbb{R}^{m,m}$ asymptotisch stabil, d.h., $\Lambda(A), \Lambda(B) \subseteq \mathbb{C}^-$. Dann hat die Sylvester-Gleichung (3.1) die eindeutige Lösung

$$X = - \int_0^\infty e^{At} W e^{Bt} dt.$$

Beweis. Da $\Lambda(A), \Lambda(B) \subseteq \mathbb{C}^-$ gilt, folgt $\Lambda(A) \cap \Lambda(-B) = \emptyset$. Damit existiert X und ist eindeutig. Sei $Z: [0, \infty) \rightarrow \mathbb{R}^{n,m}$ die eindeutige Lösung von

$$\dot{Z}(t) = AZ(t) + Z(t)B, \quad Z(0) = W.$$

Wir zeigen, dass

$$Z(t) = e^{At}W e^{Bt}.$$

Offensichtlich gilt $Z(0) = IWI = W$. Da B und e^{Bt} kommutieren, gilt zudem

$$\dot{Z}(t) = Ae^{At}W e^{Bt} + e^{At}W B e^{Bt} = AZ(t) + Z(t)B.$$

Da A und B asymptotisch stabil sind, erfüllt der Grenzwert

$$Z_\infty := \lim_{t \rightarrow \infty} Z(t) = \lim_{t \rightarrow \infty} e^{At}W e^{Bt} = 0.$$

Integration über $[0, \infty)$ liefert schließlich

$$-W = Z_\infty - Z(0) = \int_0^\infty \dot{Z}(t) dt = A \int_0^\infty Z(t) dt + \int_0^\infty Z(t) dt B. \quad \square$$

Ein Spezialfall der Sylvester-Gleichung (3.1) ist die *Lyapunov-Gleichung*. Diese ist gegeben für $m = n$, $B = A^T$ und $W = W^T$, d.h.,

$$AX + XA^T = W. \quad (3.2)$$

Aus Satz 3.5 wissen wir, dass diese eindeutig lösbar ist genau dann, wenn $\Lambda(A) \cap \Lambda(-A) = \emptyset$. Dies ist zum Beispiel erfüllt, wenn $\Lambda(A) \subseteq \mathbb{C}^-$. Grundsätzlich nicht erfüllt ist diese Bedingung, wenn A einen rein-imaginären Eigenwert hat. Ist A Hurwitz, dann gilt für die Lösung

$$X = - \int_0^\infty e^{At}W e^{A^T t} dt.$$

Aufgrund der Symmetrie der Gleichung (3.2) ist auch X^T eine Lösung. Da die Lösung aber eindeutig ist, muss als $X^T = X$ gelten.

Korollar 3.8. *Sei $A \in \mathbb{R}^{n,n}$ asymptotisch stabil. Dann ist (A, B) genau dann steuerbar, wenn die Lösung der Lyapunov-Gleichung*

$$AX + XA^T = -BB^T$$

positiv definit ist.

Beweis. Da A asymptotisch stabil ist, wissen wir aus Bemerkung 2.20, dass (A, B) genau dann steuerbar ist, wenn $W = \int_0^\infty e^{At} B B^T e^{A^T t} dt > 0$. Zudem wissen wir, dass W die Gleichung $AX + X A^T = -B B^T$ löst. \square

Satz 3.9 (Lyapunov, 1897). *Seien $W = W^T \in \mathbb{R}^{n,n}$ negativ definit und $A \in \mathbb{R}^{n,n}$.*

- i) *Ist A asymptotisch stabil, dann hat die Lyapunov-Gleichung $AX + X A^T = W$ eine eindeutige Lösung X , wobei $X = X^T > 0$.*
- ii) *Besitzt $AX + X A^T = W$ eine Lösung X mit $X = X^T > 0$, dann ist A asymptotisch stabil.*

Beweis. i) Es ist nur die positiv Definitheit der Lösung zu zeigen. Diese ist gegeben durch

$$X = - \int_0^\infty e^{At} W e^{A^T t} dt = \int_0^\infty e^{At} (-W) e^{A^T t} dt.$$

Da W negativ definit ist und diese kongruent transformiert wird, folgt $X > 0$.

ii) Sei λ ein Eigenwert von A^T , d.h., $A^T x = \lambda x$. Adjungieren der Gleichung liefert dann $x^* A = \bar{\lambda} x^*$. Mit der Lyapunov-Gleichung folgt dann

$$0 > x^* W x = x^* (AX + X A^T) x = x^* X x (\bar{\lambda} + \lambda).$$

Da nach Voraussetzung $x^* X x > 0$ gilt, muss $2\Re(\lambda) = \bar{\lambda} + \lambda < 0$ sein, also A asymptotisch stabil. \square

Mithilfe der zweiten Aussage des Satzes kann die (asymptotische) Stabilität eines Systems untersucht werden, ohne Eigenwerte zu berechnen. Dies führt auf den *Lyapunov-Test* für asymptotische Stabilität von $\dot{x} = Ax$:

- i) Wähle ein $\alpha > 0$ und löse die Lyapunov-Gleichung $AX + X A^T = -\alpha I$.
- ii) Prüfe, ob $X = X^T > 0$ ist (z.B. mit Cholesky-Zerlegung). In dem Fall sind alle Lösungen von $\dot{x} = Ax$ asymptotisch stabil. Daher nennt man asymptotische Stabilität auch *Lyapunov-Stabilität*.

Die negative Definitheit von W im Satz von Lyapunov kann für steuerbare LTI Systeme abgeschwächt werden.

Satz 3.10 (Chen 1973, Wimmer 1974). *Sei $(A, B) \in \mathbb{R}^{n,n} \times \mathbb{R}^{n,m}$ steuerbar.*

- i) Ist A asymptotisch stabil, dann hat $AX + XA^T = -BB^T$ eine eindeutige Lösung X mit $X = X^T > 0$.
- ii) Hat $AX + XA^T = -BB^T$ eine Lösung X mit $X = X^T > 0$, dann ist A asymptotisch stabil.

Beweis. Der erste Punkt folgt aus Korollar 3.8. Für den zweiten Punkt verweisen wir auch die Artikel von Chen und Wimmer. \square

Dieser Satz bildet eine Verallgemeinerung von Satz 3.9, da $-BB^T$ nur negativ semidefinit ist.

3.2 Stabilisierung mit Lyapunov-Gleichungen

Wir erinnern uns an das Ziel diesen Abschnitts. Zu gegebenen Matrizen A, B suchen wir eine Feedback-Matrix F , sodass $A - BF$ asymptotisch stabil ist. Diese nennen wir dann *stabilisierende Feedback-Matrix* ([stabilizing feedback matrix](#)). Dazu brauchen wir noch eine Verallgemeinerung der Inversen einer Matrix.

Definition 3.11. Sei $A \in \mathbb{R}^{n,m}$ eine Matrix. Mit A^+ bezeichnen wir die Moore-Penrose-Inverse. Diese ist definiert als die eindeutige Matrix, welche die folgenden Eigenschaften erfüllt:

- i) $AA^+A = A$,
- ii) $A^+AA^+ = A^+$,
- iii) $(AA^+)^T = AA^+$,
- iv) $(A^+A)^T = A^+A$.

Der folgende Satz zeigt, dass sich die Moore-Penrose-Inverse mithilfe der SVD aus Kapitel 1.1.3 berechnen lässt.

Satz 3.12. Sei $A \in \mathbb{R}^{n,m}$ eine Matrix. Dann existiert genau eine Moore-Penrose-Inverse. Zu einer SVD

$$A = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^T$$

ist A^+ eindeutig gegeben durch

$$A^+ = V \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T.$$

Im Fall $m = n$ und A regulär ist, gilt $A^+ = A^{-1}$.

Beweis. Übungsaufgabe (optional) □

Definition 3.13. Für eine quadratische Matrix $A \in \mathbb{R}^{n,n}$ bezeichnet man mit $\rho(A)$ den Stabilitätsradius, d.h. $\rho(A) = \max_{\lambda \in \Lambda(A)} |\lambda|$.

Satz 3.14. Sei $(A, B) \in \mathbb{R}^{n,n} \times \mathbb{R}^{n,m}$ stabilisierbar und $\beta \in \mathbb{R}$ mit $\beta > \rho(A)$. Sei zudem X eine Lösung der Lyapunov-Gleichung

$$-(A + \beta I)X - X(A + \beta I)^T = -2BB^T. \quad (3.3)$$

Dann ist $F = B^T X^+$ eine stabilisierende Feedback-Matrix.

Beweis. Wir betrachten zwei Fälle.

1. Fall: Sei (A, B) steuerbar. Dann ist nach dem Hautus-Test auch $(-(A + \beta I), \sqrt{2}B)$ steuerbar, denn

$$\begin{aligned} & \text{Rang} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n \quad \text{für alle } \lambda \in \mathbb{C} \\ \Leftrightarrow & \text{Rang} \begin{bmatrix} -A + \lambda I & \sqrt{2}B \end{bmatrix} = n \quad \text{für alle } \lambda \in \mathbb{C} \\ \Leftrightarrow & \text{Rang} \begin{bmatrix} -(A + \beta I) - \tilde{\lambda} I & \sqrt{2}B \end{bmatrix} = n \quad \text{für alle } \tilde{\lambda} \in \mathbb{C}. \end{aligned}$$

Da $-(A + \beta I)$ asymptotisch stabil ist, folgt mit Satz 3.10, dass eine Lösung X der Lyapunov-Gleichung (3.3) existiert mit $X > 0$ gilt. Also gilt,

$$X^{-1}(A + \beta I) + (A + \beta I)^T X^{-1} = 2X^{-1}BB^T X^{-1}.$$

Durch Umformen erhalten wir

$$X^{-1} \underbrace{(A - BB^T X^{-1})}_{\tilde{A}^T} + \underbrace{(A^T - X^{-1}BB^T)}_{\tilde{A}} X^{-1} = \underbrace{-2\beta X^{-1}}_{\tilde{W}}$$

und somit eine weitere Lyapunov-Gleichung der Form $\tilde{A}\tilde{X} + \tilde{X}\tilde{A}^T = \tilde{W} < 0$. Diese Gleichung hat offensichtlich die Lösung $\tilde{X} = X^{-1} > 0$. Damit folgt mit Satz 3.9, dass \tilde{A} und somit auch $\tilde{A}^T = A - BB^T X^{-1}$ asymptotisch stabil ist. Da $X^{-1} = X^+$ gilt, folgt die Behauptung.

2. Fall: Sei (A, B) nicht steuerbar, aber stabilisierbar. Wie im ersten Fall ist $-(A + \beta I)$ asymptotisch stabil und somit besitzt (3.3) nach Satz 3.7 die eindeutige Lösung

$$X = 2 \int_0^\infty e^{-(A+\beta I)t} BB^T e^{-(A+\beta I)^T t} dt \geq 0.$$

Man sieht, dass X symmetrisch und positiv semidefinit ist. Die Invertierbarkeit ist damit jedoch noch nicht gegeben. Wir betrachten die Kalman-Zerlegung von (A, B) ,

$$A = V \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T, \quad B = V \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

Dabei ist V orthogonal, (A_1, B_1) steuerbar und A_3 asymptotisch stabil. Wir wenden diesen Basiswechsel auf (3.3) an,

$$V^T(A + \beta I)V V^T X V + V^T X V V^T(A + \beta I)^T V = 2V^T B B^T V.$$

Ausgeschrieben ergibt das

$$\begin{bmatrix} A_1 + \beta I & A_2 \\ 0 & A_3 + \beta I \end{bmatrix} V^T X V + V^T X V \begin{bmatrix} A_1^T + \beta I & 0 \\ A_2^T & A_3^T + \beta I \end{bmatrix} = 2 \begin{bmatrix} B_1 B_1^T & 0 \\ 0 & 0 \end{bmatrix}.$$

Mit der Notation

$$V^T X V = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix}, \quad X_1 \in \mathbb{R}^{r,r}, \quad X_3 \in \mathbb{R}^{n-r,n-r}$$

ergibt eine blockweise Betrachtung der Gleichung dann

$$\begin{aligned} (A_1 + \beta I)X_1 + X_1(A_1 + \beta I)^T + A_2 X_2^T + X_2 A_2^T &= 2B_1 B_1^T, \\ (A_3 + \beta I)X_3 + X_3(A_3 + \beta I)^T &= 0. \end{aligned}$$

Da $\Lambda(A_3 + \beta I) \subseteq \mathbb{C}^+$, hat die zweite Gleichung die eindeutige Lösung $X_3 = 0$. Da $X \geq 0$, ist auch $V^T X V \geq 0$. Damit muss $X_2 = 0$ gelten ([Übungsaufgabe \(optional\)](#)). Übrig bleibt die Lyapunov-Gleichung

$$(A_1 + \beta I)X_1 + X_1(A_1 + \beta I)^T = 2B_1 B_1^T.$$

Dabei ist (A_1, B_1) steuerbar. Im ersten Fall haben wir gezeigt, dass $F_1 := B_1^T X_1^{-1}$ eine stabilisierende Feedback-Matrix ist. Es gilt also $\Lambda(A_1 - B_1 F_1) \subseteq \mathbb{C}^-$. Zudem erhalten wir für die Moore-Penrose-Inverse von X durch die spezielle Struktur,

$$X = V \begin{bmatrix} X_1 & 0 \\ 0 & 0 \end{bmatrix} V^T \quad \Rightarrow \quad X^+ = V \begin{bmatrix} X_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^T.$$

Setzt man $F := B^T X^+ = [B_1^T X_1^{-1} \ 0] V^T$, so ergibt sich

$$\begin{aligned} A - BF &= V \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T - V \begin{bmatrix} B_1 \\ 0 \end{bmatrix} [B_1^T X_1^{-1} \ 0] V^T \\ &= V \begin{bmatrix} A_1 - B_1 F_1 & A_2 \\ 0 & A_3 \end{bmatrix} V^T. \end{aligned}$$

Somit gilt $\Lambda(A - BF) = \Lambda(A_1 - B_1 F_1) \cup \Lambda(A_3) \subseteq \mathbb{C}^-$. □

Wir haben nun also gezeigt, dass für jedes stabilisierbare System (A, B) eine stabilisierende Feedback-Matrix existiert. Damit ist nun auch Satz 2.39 bewiesen.

Aus den Überlegungen zum Beweis von Satz 3.14 resultiert der folgende Bass-Algorithmus zur Berechnung der stabilisierenden Feedback-Matrix.

Algorithmus 3.1 (Bass)

Input: (A, B) stabilisierbar

Output: F , sodass $\Lambda(A - BF) \subseteq \mathbb{C}^-$ gilt

- 1: Setze β so, dass $-(A + \beta I)$ asymptotisch stabil ist, z.B. durch $\beta := 2\|A\|_p$, wobei $p \in \{1, \infty, F\}$. Das impliziert $\beta > \rho(A)$.
 - 2: Löse $(A + \beta I)X + X(A + \beta I)^T = 2BB^T$ nach X .
 - 3: Berechne X^+ .
 - 4: Setze $F := B^T X^+$.
-

Für den Algorithmus ist die Lösung einer Lyapunov-Gleichung notwendig. Deshalb betrachten wir im folgenden Abschnitt numerische Methoden zu diesem Zweck.

3.3 Numerische Lösung von Lyapunov-Gleichungen

Ziel ist es, die Lyapunov-Gleichung

$$AX + XA^T = W$$

mit $W = W^T$ und $A \in \mathbb{R}^{n,n}$ numerisch zu lösen. Die gesuchte Lösung ist eine symmetrische Matrix $X = X^T$. Um die Eindeutigkeit zu garantieren, nehmen wir an, dass $\Lambda(A) \cap \Lambda(-A) = \emptyset$. Im Folgenden betrachten wir verschiedene Methoden zur Berechnung von X .

3.3.1 Kronecker-Produkt

Wir haben bereits gesehen, dass die Lyapunov-Gleichung äquivalent ist zum linearen Gleichungssystem

$$M \operatorname{vec}(X) = \operatorname{vec}(W), \quad M = (I \otimes A) + (A \otimes I) \in \mathbb{R}^{n^2, n^2}.$$

Dieses Gleichungssystem zu lösen, ist numerisch betrachtet aber keine gute Idee. Eine LU Zerlegung von M benötigt $\frac{2}{3}n^6$ Flops und Speicherplatz in der Größenordnung $\mathcal{O}(n^4)$. Die Methode ist also nur für „kleine“ n geeignet. Besser ist es, die Matrix A auf eine einfachere Form zu transformieren und dann die Lyapunov-Gleichung direkt zu lösen.

Da die Lyapunov-Gleichung äquivalent zu einem linearen Gleichungssystem ist, kann man auch hier fragen wie sich kleine Störungen in den Daten auf die Lösung auswirken. Wir betrachten dazu ein lineares Gleichungssystem $Mx = b$ und

$$\tilde{M}\tilde{x} = \tilde{b}, \quad \tilde{M} = M + \Delta M, \quad \tilde{x} = x + \Delta x, \quad \tilde{b} = b + \Delta b.$$

Damit gilt

$$M\Delta x = b + \Delta b - Mx - \Delta M\tilde{x} = \Delta b - \Delta M\tilde{x}.$$

Nimmt man die Norm, so erhält man die Abschätzung

$$\|\Delta x\| \leq \|M^{-1}\| \left(\|\Delta b\| + \|\Delta M\| (\|x\| + \|\Delta x\|) \right).$$

Das impliziert schließlich

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\|M^{-1}\| \|M\| \left(\frac{\|\Delta b\|}{\|M\| \|x\|} + \frac{\|\Delta M\|}{\|M\|} \right)}{1 - \|M^{-1}\| \|\Delta M\|} \leq \frac{\|M^{-1}\| \|M\| \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta M\|}{\|M\|} \right)}{1 - \|M^{-1}\| \|\Delta M\|},$$

womit man den relativen Fehler in x beschränken kann.

Definition 3.15. Zu einer Lyapunov-Gleichung $AX + XA^T = W$ definieren wir die Separation von A durch

$$\operatorname{Sep} A := \min_{\|X\|_F=1} \|AX + XA^T\|_F.$$

Damit kann man wiederum die Konditionszahl einer Lyapunov-Gleichung definieren, nämlich

$$\kappa := \frac{2}{\operatorname{Sep} A} \|A\|_2.$$

Man beachte, dass die Konditionszahl den Operator A in der 2-Norm und die rechte Seite in der F-Norm beinhaltet. Wie für lineare Gleichungssysteme erhält man damit ein Maß wie groß der Fehler durch kleine Störungen in den Daten werden kann. Dazu betrachten wir folgendes Lemma.

Lemma 3.16. *Sei X die Lösung von $AX + XA^T = W$ und \tilde{X} die Lösung der gestörten Lyapunov-Gleichung $\tilde{A}\tilde{X} + \tilde{X}\tilde{A}^T = \tilde{W}$ mit $\|\tilde{A} - A\|_2 \leq \varepsilon\|A\|_2$ und $\|\tilde{W} - W\|_F \leq \varepsilon\|W\|_F$. Gilt $\varepsilon\kappa < 1$, wobei κ die Konditionszahl der Lyapunov-Gleichung ist, dann gilt*

$$\frac{\|\tilde{X} - X\|_F}{\|X\|_F} \leq \frac{2\varepsilon\kappa}{1 - \varepsilon\kappa}.$$

Beweis. Wir zeigen zunächst $(\text{Sep } A)^{-1} = \|(I \otimes A + A \otimes I)^{-1}\|_2$ und nutzen dann die Abschätzung für allgemeine Gleichungssysteme.

Übungsaufgabe 4.3 □

3.3.2 Bartels-Stewart

Die Idee hier ist, A auf Dreiecks-Form zu transformieren. Dann kann man X durch Rückwärtseinsetzen ([backward substitution](#)) berechnen. Welche Operationen dürfen wir anwenden? Wir multiplizieren die Lyapunov-Gleichung mit Q^{-1} von links und Q von rechts. Für die numerische Stabilität fordern wir zudem, dass Q orthogonal ist, also $Q^{-1} = Q^T$. Wir erhalten

$$\underbrace{Q^T A Q}_{\tilde{A}} \underbrace{Q^T X Q}_{\tilde{X}} + \underbrace{Q^T X Q}_{\tilde{X}} \underbrace{Q^T A^T Q}_{\tilde{A}^T} = \underbrace{Q^T W Q}_{\tilde{W}}.$$

Wir betrachten also Kongruenztransformation von A , X und W .

Satz 3.17 (Reelle Schur-Form). *Sei $A \in \mathbb{R}^{n,n}$. Dann existiert ein orthogonales $Q \in \mathbb{R}^{n,n}$, sodass*

$$Q^T A Q = T = \begin{bmatrix} A_{11} & \dots & \dots & A_{1r} \\ 0 & A_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & A_{rr} \end{bmatrix}.$$

Dabei sind $A_{jj} \in \mathbb{R}^{n_j, n_j}$, $n_j \in \{1, 2\}$ und $\sum_{i=1}^r n_j = n$. Ist $n_j = 1$, dann ist A_{jj} ein reeller Eigenwert von A . Im Fall $n_j = 2$ sind die Eigenwerte von A_{jj} ein komplex konjugiertes Eigenwert-Paar von A .

Beweis. Den Beweis findet man in Kapitel 7.4 in Golub, Van Loan: *Matrix computations*, 3. Edition, 1996. \square

Bemerkung 3.18. Wir nennen T die reelle Schur-Form von A und die Spalten von Q Schur-Vektoren. Die Berechnung erfolgt über den QR-Algorithmus. Dieser ist numerisch stabil und benötigt $25n^3$ Flops zur Berechnung von T und Q . Der benötigte Speicher liegt in der Größenordnung $\mathcal{O}(n^2)$.

Mit der reellen Schur-Form lässt sich nun per Rückwärtseinsetzen die Lösung X berechnen. Sei

$$Q^T A Q = T =: \begin{bmatrix} A_1 & A_2 \\ 0 & A_{rr} \end{bmatrix}, \quad A_1 \in \mathbb{R}^{n-n_r, n-n_r}, \quad A_{rr} \in \mathbb{R}^{n_r, n_r}.$$

Mit

$$\tilde{X} = Q^T X Q = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix}, \quad \tilde{W} = Q^T W Q = \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix}$$

ergibt sich dann aus der Lyapunov-Gleichung

$$\begin{bmatrix} A_1 & A_2 \\ 0 & A_{rr} \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} + \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ 0 & A_{rr} \end{bmatrix}^T = \begin{bmatrix} W_1 & W_2 \\ W_2^T & W_3 \end{bmatrix}.$$

Blockweise ergeben sich dadurch die drei Gleichungen

$$A_{rr} X_3 + X_3 A_{rr}^T = W_3, \quad (3.4a)$$

$$A_1 X_2 + X_2 A_{rr}^T = W_2 - A_2 X_3 =: \tilde{W}_2, \quad (3.4b)$$

$$A_1 X_1 + X_1 A_1^T = W_1 - A_2 X_2^T - X_2 A_2^T =: \tilde{W}_1. \quad (3.4c)$$

Gleichung (3.4a) können wir nach X_3 lösen. Im Fall $n_r = 1$ ist $A_{rr} \in \mathbb{R}$ und somit $X_3 = W_3 / (2A_{rr})$. Andernfalls ($n_r = 2$) ergibt sich mit

$$A_{rr} = \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \quad W_3 = \begin{bmatrix} w_1 & w_2 \\ w_2 & w_3 \end{bmatrix}, \quad X_3 = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix}$$

das Gleichungssystem

$$\begin{aligned} & \begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} + \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_1 & a_3 \\ a_2 & a_4 \end{bmatrix} = \begin{bmatrix} w_1 & w_2 \\ w_2 & w_3 \end{bmatrix} \\ \Leftrightarrow & \begin{bmatrix} 2a_1 & 2a_2 & 0 \\ a_3 & a_1 + a_4 & a_2 \\ 0 & 2a_3 & 2a_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}. \end{aligned}$$

Dies kann gelöst werden mit einer LU Zerlegung mit Pivotisierung.

Als zweites lösen wir (3.4b) nach X_2 . Offensichtlich gilt

$$\Lambda(A_1) \cap \Lambda(-A_{rr}) \subseteq \Lambda(A) \cap \Lambda(-A) = \emptyset.$$

Damit hat die Sylvester-Gleichung (3.4b), also

$$\begin{bmatrix} A_{11} & \cdots & A_{1r-1} \\ & \ddots & \vdots \\ & & A_{r-1r-1} \end{bmatrix} \begin{bmatrix} \hat{X}_1 \\ \vdots \\ \hat{X}_{r-1} \end{bmatrix} + \begin{bmatrix} \hat{X}_1 \\ \vdots \\ \hat{X}_{r-1} \end{bmatrix} A_{rr}^T = \tilde{W}_2 = \begin{bmatrix} \hat{W}_1 \\ \vdots \\ \hat{W}_{r-1} \end{bmatrix}$$

eine eindeutige Lösung. Die Teilmatrizen \hat{X}_j können durch Rückwärtseinsetzen bestimmt werden. Dazu löst man iterativ

$$A_{jj}\hat{X}_j + \hat{X}_j A_{rr}^T = \hat{W}_j - \sum_{k=j+1}^{r-1} A_{jk}\hat{X}_k.$$

für alle $j = r-1, \dots, 1$. Dies ist jeweils eine Sylvester-Gleichung in niedriger Dimension und kann somit mit dem Kronecker-Produkt gelöst werden.

Schließlich bleibt noch Gleichung (3.4c) zu lösen. Hier nutzen wir die bekannte Schur-Form von A_1 und gehen vor wie in den Schritten zuvor mit

$$A_1 = \begin{bmatrix} \tilde{A}_1 & \tilde{A}_2 \\ 0 & A_{r-1r-1} \end{bmatrix}.$$

Dies wiederholen wir so lange bis (3.4c) eine Lyapunov-Gleichung der Größe 1×1 oder 2×2 ist.

Bemerkung 3.19. *Der gegebene Algorithmus geht zurück auf Bartels und Stewart aus dem Jahr 1972. Dieser ist auch implementiert in Matlab (die Funktion `lyap`). Die Kosten liegen bei rund $32n^3$ Flops, der größte Teil davon für die reelle Schur-Form. Zudem ist der Algorithmus numerisch stabil durch die Verwendung orthogonaler Transformationen.*

3.3.3 Matrix Signum Methode

Für die Matrix Signum Methode ([sign function method](#)) definieren wir zunächst das Signum einer Matrix.

Definition 3.20. Sei $M \in \mathbb{R}^{n,n}$ mit $\Re(\lambda) \neq 0$ für alle $\lambda \in \Lambda(M)$. Sei

$$M = S \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix} S^{-1}$$

eine Jordan-Normalform von M , wobei $J_- \in \mathbb{C}^{k,k}$ und $J_+ \in \mathbb{C}^{n-k,n-k}$ so gewählt sind, dass $\Lambda(J_-) \subseteq \mathbb{C}^-$ und $\Lambda(J_+) \subseteq \mathbb{C}^+$ gilt. Dann heißt

$$\operatorname{sgn}(M) := S \begin{bmatrix} -I_k & 0 \\ 0 & I_{n-k} \end{bmatrix} S^{-1}$$

das Signum (*sign*) von M .

Lemma 3.21. Für jedes invertierbare $T \in \mathbb{C}^{n,n}$ gilt

$$\operatorname{sgn}(TMT^{-1}) = T \operatorname{sgn}(M)T^{-1}.$$

Beweis. Übungsaufgabe (optional) □

Bemerkung 3.22. Lemma 3.21 zeigt die Wohldefiniertheit von sgn , d.h., die Wahl von S spielt bei der Jordan-Normalform keine Rolle. Sind also

$$M = S \begin{bmatrix} J_- & 0 \\ 0 & J_+ \end{bmatrix} S^{-1} = T \begin{bmatrix} K_- & 0 \\ 0 & K_+ \end{bmatrix} T^{-1}$$

zwei Jordan-Normalformen mit $J_-, K_- \in \mathbb{C}^{k,k}$ und $\Lambda(J_-), \Lambda(K_-) \subseteq \mathbb{C}^-$, $\Lambda(J_+), \Lambda(K_+) \subseteq \mathbb{C}^+$, dann gilt auch

$$S \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} S^{-1} = T \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} T^{-1}.$$

Das folgende Lemma zeigt den Zusammenhang mit Lyapunov- und Sylvester-Gleichungen.

Lemma 3.23. Seien $A \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{m,m}$, $W \in \mathbb{R}^{n,m}$ mit $\Lambda(A), \Lambda(B) \subseteq \mathbb{C}^-$. Sei X die Lösung der Sylvester-Gleichung $AX + XB = W$. Dann gilt

$$\operatorname{sgn} \begin{bmatrix} A & W \\ 0 & -B \end{bmatrix} = \begin{bmatrix} -I & -2X \\ 0 & I \end{bmatrix}.$$

Beweis. Nach Voraussetzung existiert nach Satz 3.5 eine eindeutige Lösung X der Sylvester-Gleichung. Weiterhin gilt wegen $\Lambda(A), \Lambda(B) \subseteq \mathbb{C}^-$,

$$\operatorname{sgn}(A) = -I_n, \quad \operatorname{sgn}(-B) = I_m.$$

Wir definieren $T = T^{-1}$ durch

$$T := \begin{bmatrix} I & X \\ 0 & -I \end{bmatrix}.$$

Mit $T^2 = I$ und Lemma 3.21 folgt somit

$$\begin{aligned} \operatorname{sgn} \begin{bmatrix} A & W \\ 0 & -B \end{bmatrix} &= T \operatorname{sgn} \left(T \begin{bmatrix} A & W \\ 0 & -B \end{bmatrix} T \right) T \\ &= T \operatorname{sgn} \left(\begin{bmatrix} A & AX + XB - W \\ 0 & -B \end{bmatrix} \right) T \\ &= \begin{bmatrix} I & X \\ 0 & -I \end{bmatrix} \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & X \\ 0 & -I \end{bmatrix} = \begin{bmatrix} -I & -2X \\ 0 & I \end{bmatrix}, \end{aligned}$$

was zu beweisen war. □

Dieses Lemma zeigt, dass wir aus dem Signum einer oberen Block-Dreiecksmatrix die Lösung der Sylvester-Gleichung berechnen können. Nun müssen wir noch einen Weg finden, um das Signum effektiv numerisch zu berechnen. Zu einer invertierbaren Matrix M betrachten wir dazu die Iteration

$$M_{(i+1)} = \frac{1}{2} \left(M_{(i)} + M_{(i)}^{-1} \right) \tag{3.5}$$

mit Startwert $M_{(0)} = M$. Wir wollen zeigen, dass diese Iteration tatsächlich global gegen das Signum von M konvergiert.

Satz 3.24. *Für eine invertierbare Matrix M konvergiert die Iteration (3.5) mit Startwert $M_{(0)} = M$ global gegen $\operatorname{sgn}(M)$, wobei die Konvergenz lokal quadratisch ist.*

Für den Beweis dieser Aussage brauchen wir noch zwei Lemmata.

Lemma 3.25 (skalare Signum Iteration). *Gegeben sei die rekursive Folge $\lambda_{j+1} = \frac{1}{2}(\lambda_j + 1/\lambda_j)$ mit $\lambda_0 \in \mathbb{C}$, $\lambda_0 \neq 0$. Für die Konvergenz der Folge gilt dann:*

- i) Ist $\Re(\lambda_0) = 0$, dann divergiert (λ_j) oder die Iteration bricht ab wegen Division durch Null.
- ii) Ist $\Re(\lambda_0) > 0$, dann konvergiert (λ_j) gegen 1.
- iii) Ist $\Re(\lambda_0) < 0$, dann konvergiert (λ_j) gegen -1 .

Beweis. Übungsaufgabe (optional) □

Lemma 3.26. Sei $s \in \{1, -1\}$ und J eine rechte obere Matrix mit s auf der Diagonalen und 0 auf den ersten k Superdiagonalen. Die Einträge auf den nächsten $(k+1)$ Superdiagonalen seien mit a_{ij} bezeichnet. Dann hat J^{-1} ebenfalls s auf der Diagonalen sowie k Superdiagonalen mit Einträgen 0. Zudem sind die Einträge auf den folgenden $(k+1)$ Superdiagonalen durch $-a_{ij}$ gegeben. Somit hat $\frac{1}{2}(J+J^{-1})$ oberhalb der Diagonalen $2k+1$ Nulldiagonalen stehen.

Beweis. Diese Aussage kann man durch Nachrechnen nachvollziehen. □

Nun haben wir die nötigen Mittel zusammen, um Satz 3.24 zu beweisen.

Beweis zu Satz 3.24. Übungsaufgabe 5.1a □

Bemerkung 3.27. Eine Iteration (3.5) benötigt $\mathcal{O}(n^3)$ Operationen. Der Aufwand von Bartels-Stewart entspricht dem Aufwand von ein paar Iterationen der Signum-Iteration.

3.3.4 Hammarling-Verfahren

Hierbei handelt es sich um eine Abwandlung von Bartels-Stewart, die etwas schneller ist. Diese berechnet allerdings nicht die Lösung X sondern den Cholesky Faktor G mit $X = GG^T$.

3.3.5 ADI-Iteration

ADI steht für *Alternating-Directional-Implicit*. Es handelt sich um ein Verfahren zum Lösen von sehr großen Sylvester Gleichungen mit dünn besetzten Matrizen (*sparse*). Zudem wird angenommen, dass der Rang von W klein ist. Dies ist zum Beispiel der Fall, wenn $m \ll n$.

3.4 Teilstabilisierung

Der Name des Kapitels ([partial stabilization](#)) ist etwas irreführend - es wird tatsächlich komplett stabilisiert. Wir betrachten Matrizen A , die viele Eigenwerte in \mathbb{C}^- und nur wenige in \mathbb{C}^+ hat. In dem Fall scheint es zu viel Aufwand zu sein (gerade für große Matrizen), ganz A bei der Stabilisierung zu betrachten.

Sei $Q = [Q_1, Q_2] \in \mathbb{R}^{n \times n}$ orthogonal mit $Q_2 \in \mathbb{R}^{n \times k}$, sodass

$$Q^T A Q = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}.$$

Dabei soll $\Lambda(A_{11}) \subseteq \mathbb{C}^-$ und $\Lambda(A_{22}) \subseteq \mathbb{C}^+$ für $A_{22} \in \mathbb{R}^{k \times k}$ gelten. Bilde nun

$$Q^T B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

und stabilisiere (A_{22}, B_2) . Das Paar (A_{22}, B_2) ist steuerbar, wenn (A, B) stabilisierbar ist. Zur Stabilisierung können wir den Algorithmus von Bass benutzen. Setze dafür $F_2 := -B_2^T X_2^{-1}$, wobei X_2 die Lyapunov-Gleichung

$$A_{22} X_2 + X_2 A_{22}^T = 2 B_2 B_2^T$$

löst.

Bemerkung 3.28. *Das Vorgehen hier besitzt zwei Unterschiede zum klassischen Algorithmus von Bass. Da $\Lambda(A_{22}) \subseteq \mathbb{C}^+$ gilt, können wir $\beta = 0$ wählen. Somit tritt kein βI Term auf. Zudem ist die Matrix X_2 tatsächlich invertierbar, da (A_{22}, B_2) steuerbar ist.*

Wir setzen $F := [0, F_2] Q^T$. Das ist eine stabilisierende Feedback-Matrix, da

$$Q^T (A - BF) Q = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} - \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} 0 & F_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} - B_1 F_2 \\ 0 & A_{22} - B_2 F_2 \end{bmatrix}.$$

Damit folgt

$$\Lambda(A - BF) = \Lambda(A_{11}) \cup \Lambda(A_{22} + B_2 F_2) \subseteq \mathbb{C}^-.$$

Der Vorteil davon ist, dass die zu lösende Lyapunov-Gleichung eine kleinere Dimension $k < n$ hat. In manchen Fällen kann sogar $k \ll n$ gelten. Der Nachteil ist, dass die Matrix Q benötigt wird, um A auf Block-Dreiecksform zu bringen. Um diese zu erhalten, gibt es mehrere Methoden:

- Reelle Schur-Form von A berechnen, wobei die Schur-Vektoren so geordnet sind, dass die Eigenwerte in \mathbb{C}^+ unten auftauchen. Im Algorithmus von Bartels-Stewart wird sowieso die Schur-Form berechnet, sodass die Vektoren nur umgeordnet werden müssen.
- Die Matrix $P := \frac{1}{2}(I - \operatorname{sgn} A)$ ist ein Projektor auf den A -invarianten Unterraum zu Eigenwerten in \mathbb{C}^- . (Übungsaufgabe 5.1b) Zu P berechnet man dann die dazugehörige SVD, d.h., $P = U\Sigma V^T$. Dann wählt man $Q = U$, da das Bild von P gerade der stabile invariante Unterraum ist. (Übungsaufgabe optional)
- Es gilt $A_{22} = Q_2^T A Q_2$, $B_2 = Q_2^T B$ und $F = F_2 Q_2^T$. Es wird also nur Q_2 benötigt. Mit $A^T Q_2 = Q_2 A_{22}^T$ folgt, dass Q_2 eine Orthonormalbasis des invarianten Unterraums von A^T bildet. Dies kann mit Methoden der numerischen linearen Algebra effizient gelöst werden.

3.5 Polvorgabe

Zu gegebenen Matrizen A, B und $\mathcal{P} = \{\mu_1, \dots, \mu_n\} \subseteq \mathbb{C}$ suchen wir nun eine Matrix F , sodass $\Lambda(A - BF) = \mathcal{P}$ gilt.

Definition 3.29. Die Paare (A, B) und (\tilde{A}, \tilde{B}) heißen ähnlich (*similar*), wenn eine invertierbare Matrix T existiert mit $\tilde{A} = TAT^{-1}$ und $\tilde{B} = TB$.

Definition 3.30. Sei $\rho_A(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0$ das charakteristische Polynom der Matrix $A \in \mathbb{R}^{n \times n}$. Dann heißen

$$C_1(A) := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \dots & \dots & -\alpha_{n-1} \end{bmatrix}, \quad C_2(A) := C_1(A)^T$$

die erste und zweite Begleitmatrix (*first and second companion form*) von A .

Lemma 3.31. Die erste und zweite Begleitmatrix von A haben das gleiche charakteristische Polynom wie A , d.h.,

$$\rho_{C_1(A)} = \rho_{C_2(A)} = \rho_A.$$

Beweis. Die Behauptung folgt mit dem Laplace'schen Entwicklungssatz nach der letzten Zeile (für $C_1(A) - \lambda I$) bzw. der letzten Spalte (für $C_2(A) - \lambda I$). \square

Fürs Erste sei $m = 1$ (single input), also $B \in \mathbb{R}^n$. Mit der Steuerbarkeitsmatrix $K(A, B) = [B \ AB \ \dots \ A^{n-1}B] \in \mathbb{R}^{n,n}$ gilt

$$AK(A, B) = K(A, B)A_H \quad \text{mit } A_H = C_2(A).$$

(Übungsaufgabe (optional)) Außerdem gilt

$$B = K(A, B)B_H \quad \text{mit } B_H = e_1$$

Lemma 3.32. *Das Paar $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^n$ ist steuerbar genau dann, wenn (A, B) und (A_H, B_H) mit $A_H = C_2(A)$ und $B_H = e_1$ ähnlich sind.*

Beweis. Sei (A, B) steuerbar, also die Steuerbarkeitsmatrix $K(A, B)$ invertierbar. Dann liefert $T = K(A, B)$ die behauptete Ähnlichkeit, denn

$$TA_H T^{-1} = A, \quad TB_H = B.$$

Für die Rückrichtung reicht es zu zeigen, dass (A_H, B_H) steuerbar ist, da Steuerbarkeit unter Basiswechseln invariant ist. Dies folgt aber direkt aus $K(A_H, B_H) = I$. \square

Die selbe Aussage gilt auch für die erste Begleitmatrix.

Lemma 3.33. *Seien $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^n$ und $A_{RNF} = C_1(A)$, $B_{RNF} = e_n$. Dann ist (A, B) steuerbar genau dann, wenn (A, B) und (A_{RNF}, B_{RNF}) ähnlich sind.*

Beweis. Sei (A, B) steuerbar, also (A, B) ähnlich zu (A_H, B_H) nach Lemma 3.32. Wir zeigen, dass (A_H, B_H) und (A_{RNF}, B_{RNF}) ähnlich sind. Dazu bemerken wir, dass

$$A_{RNF}K(A_{RNF}, B_{RNF}) = K(A_{RNF}, B_{RNF})A_H, \quad B_{RNF} = K(A_{RNF}, B_{RNF})B_H.$$

Die Paare (A, B) und (A_H, B_H) sind also ähnlich, wenn $K(A_{RNF}, B_{RNF})$ invertierbar ist. Dies ist der Fall, da

$$K(A_{RNF}, B_{RNF}) = \begin{bmatrix} 0 & \dots & \dots & 0 & 1 \\ \vdots & & \ddots & 1 & * \\ \vdots & \ddots & \ddots & \ddots & * \\ 0 & 1 & \ddots & & \vdots \\ 1 & * & \dots & \dots & * \end{bmatrix}. \quad \square$$

Definition 3.34. Das Paar $(A_{RNF}, B_{RNF}) = (C_1(A), e_n)$ heißt Regelungsnormalform (*controllable canonical form*) von (A, B) .

Definition 3.35. Eine Menge $\mathcal{P} \subseteq \mathbb{C}$ heißt abgeschlossen unter komplexer Konjugation, falls $\mathcal{P} = \overline{\mathcal{P}}$, d.h., für $\lambda \in \mathcal{P}$ gilt auch $\bar{\lambda} \in \mathcal{P}$.

Satz 3.36 (Polvorgabe). Sei $(A, B) \in \mathbb{R}^{n,n} \times \mathbb{R}^{n,m}$, $B \neq 0$ und $\lambda_{k+1}, \dots, \lambda_n$ die nicht-steuerbaren Eigenwerte von A . Sei zudem $\mathcal{P} \subseteq \mathbb{C}$. Dann sind die folgenden Aussagen äquivalent.

- i) Es gibt eine Matrix $F \in \mathbb{R}^{m,n}$, sodass $\Lambda(A - BF) = \mathcal{P}$ gilt.
- ii) $\mathcal{P} = \{\mu_1, \dots, \mu_k, \lambda_{k+1}, \dots, \lambda_n\}$, wobei die Menge $\{\mu_1, \dots, \mu_k\}$ abgeschlossen ist unter komplexer Konjugation.

Gilt ii), so ist F für $m = 1$ eindeutig bestimmt.

Beweis. Sei zunächst (A, B) nicht steuerbar. Ohne Beschränkung der Allgemeinheit sei (A, B) in Kalman-Form gegeben, d.h.,

$$A = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}.$$

Dabei ist (A_1, B_1) steuerbar und $\Lambda(A_3) = \{\lambda_{k+1}, \dots, \lambda_n\}$. Dann gilt für beliebiges $F = [F_1, F_2]$,

$$A - BF = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} - \begin{bmatrix} B_1 \\ 0 \end{bmatrix} [F_1 \quad F_2] = \begin{bmatrix} A_1 - B_1 F_1 & A_2 - B_1 F_2 \\ 0 & A_3 \end{bmatrix}.$$

Damit folgt

$$\Lambda(A - BF) = \Lambda(A_1 - B_1 F_1) \cup \Lambda(A_3).$$

Für die Aussage des Satzes reicht es also, nur den steuerbaren Teil (A_1, B_1) zu betrachten. Im Folgenden nehmen wir daher an, dass (A, B) steuerbar ist.

i) \Rightarrow ii) Da $A_1 - B_1 F_1$ eine reelle Matrix ist, muss $\Lambda(A_1 - B_1 F_1)$ abgeschlossen bezüglich komplexer Konjugation sein. Somit hat \mathcal{P} die behauptete Form.

ii) \Rightarrow i) Sei (A, B) steuerbar und $\mathcal{P} = \{\mu_1, \dots, \mu_n\} = \overline{\mathcal{P}}$. Wir betrachten zunächst den Fall $m = 1$. Da (A, B) steuerbar ist, ist das Paar nach Lemma 3.33 ähnlich zu seiner Regelungsnormalform. Ohne Beschränkung der Allgemeinheit sei (A, B) also gegeben durch

$$A = C_1(A), \quad B = e_n.$$

Mit $F = [f_1 \ \dots \ f_n] \in \mathbb{R}^{1,n}$ gilt dann

$$A - BF = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 1 \\ -(\alpha_0 + f_1) & -(\alpha_1 + f_2) & \dots & \dots & -(\alpha_{n-1} + f_n) \end{bmatrix}.$$

Das charakteristische Polynom von $A - BF$ ist gegeben durch

$$\rho(\lambda) = \lambda^n + (\alpha_{n-1} + f_n)\lambda^{n-1} + \dots + (\alpha_0 + f_1).$$

Wegen $\mathcal{P} = \bar{\mathcal{P}}$ existieren Koeffizienten $\beta_0, \dots, \beta_{n-1} \in \mathbb{R}$, sodass

$$(\lambda - \mu_1)(\lambda - \mu_2) \dots (\lambda - \mu_n) = \lambda^n + \beta_{n-1}\lambda^{n-1} + \dots + \beta_0.$$

Die Matrix F lässt sich also eindeutig bestimmen durch $f_j = \beta_{j-1} - \alpha_{j-1}$, $j = 1, \dots, n$.

Ist $m > 1$, dann können wir das Problem auf den Fall $m = 1$ zurückführen. Da $B \neq 0$ ist, gibt es einen Vektor $v \in \mathbb{R}^m$, sodass $b := Bv \neq 0$. Ziel ist es nun, eine Matrix $G \in \mathbb{R}^{m,n}$ zu konstruieren, sodass $(A - BG, b)$ steuerbar ist. Dann folgt nämlich aus dem ersten Teil des Beweises die eindeutige Existenz von $f \in \mathbb{R}^{1,n}$, sodass

$$\Lambda(A - BG - bf) = \mathcal{P}.$$

Somit erfüllt die Matrix $F := G + vf$ das Gewünschte.

Sei $x_1 := b = Bv$ und $\{x_1, \dots, x_\ell\}$ die maximale Menge linear unabhängiger Vektoren, sodass $x_j - Ax_{j-1} \in \text{Bild } B$ für $j = 2, \dots, \ell$. Wir definieren $\mathcal{H} := \text{span}\{x_1, \dots, x_\ell\}$ und zeigen $\ell = n$. Zunächst gilt für alle $u \in \mathbb{R}^m$, dass

$$Ax_\ell + Bu \in \mathcal{H}.$$

Wäre dies nämlich nicht der Fall, dann widerspricht $x_{\ell+1} := Ax_\ell + Bu$ der Maximalität von ℓ . Für $u = 0$ impliziert dies $Ax_\ell \in \mathcal{H}$ und somit

$$\text{Bild } B \subseteq \mathcal{H}.$$

Wegen $x_j - Ax_{j-1} \in \text{Bild } B$ folgt dann auch $Ax_j \in \mathcal{H}$ für $j = 1, \dots, \ell - 1$. Die Menge \mathcal{H} ist also A -invariant. Mit der angenommenen Steuerbarkeit von (A, B) folgt dann

$$\mathbb{R}^n = \text{Bild } K(A, B) \subseteq \mathcal{H},$$

also $\ell = n$.

Nach Konstruktion der x_j existieren $u_1, \dots, u_{n-1} \in \mathbb{R}^m$, sodass

$$x_j - Ax_{j-1} = Bu_{j-1}, \quad j = 2, \dots, n.$$

Wir wählen u_n beliebig und definieren $X := [x_1, \dots, x_n] \in \mathbb{R}^{n,n}$, $U := [u_1, \dots, u_n] \in \mathbb{R}^{m,n}$. Dann erfüllt die Matrix $G := -UX^{-1}$ die Gleichung $Gx_j = -u_j$ und damit für $j = 1, \dots, n-1$,

$$(A - BG)x_j = Ax_j + Bu_j = x_{j+1}, \quad x_1 = b.$$

Es gilt also $K(A - BG, b) = X$, sodass $(A - BG, b)$ steuerbar ist. \square

Bemerkung 3.37. *Um die Eindeutigkeit der Feedback-Matrix F im Fall $m > 1$ zu erhalten, muss man weitere Bedingungen angeben. Man kann diese Freiheit ausnutzen, um die Pole des rückgekoppelten Systems so robust wie möglich gegenüber Störungen zu machen. Numerisch sinnvoll ist daher, F so zu wählen, dass $A - BF$ diagonalisierbar ist und die Eigenvektormatrix zu F minimale Kondition hat. Mehr zum Thema der robusten Polvorgabe findet man in einer Arbeit von Kautsky, Nichols und Van Dooren, 1985.*

Korollar 3.38. *Das Paar (A, B) ist genau dann steuerbar, wenn es zu jeder Menge $\mathcal{P} \subseteq \mathbb{C}$, $\mathcal{P} = \overline{\mathcal{P}}$, eine Feedback-Matrix F gibt, sodass $\Lambda(A - BF) = \mathcal{P}$.*

Beweis. Die Hinrichtung ist ein Spezialfall von Satz 3.36. Für die Rückrichtung nehmen wir an, dass (A, B) nicht steuerbar ist. Dann existiert nach Satz 2.29 ein $\lambda \in \mathbb{C}$ und ein $z \neq 0$, sodass $z^T B = 0$ und $z^T(\lambda I - A) = 0$. Sei \mathcal{P} nun eine Menge mit $\lambda \notin \mathcal{P}$. Dann gibt es kein F , sodass $\Lambda(A - BF) = \mathcal{P}$, da

$$z^T(\lambda I - (A - BF)) = z^T(\lambda I - A) + z^T BF = 0. \quad \square$$

Lemma 3.39. *Sei (A, B) steuerbar und in Regelungsnormalform gegeben, $m = 1$. Zudem sei $F \in \mathbb{R}^{1,n}$ eine Feedback-Matrix, die $\Lambda(A - BF) = \mathcal{P} = \overline{\mathcal{P}}$ erfüllt. Dann gilt:*

- i) Ist $\lambda \in \mathcal{P} \cap \Lambda(A)$, dann ist jeder Eigenvektor von A zu λ auch ein Eigenvektor von $A - BF$ zu λ .*
- ii) Ist $\lambda \in \mathcal{P} \setminus \Lambda(A)$, dann gilt $Fx = 1$ mit $x := (A - \lambda I)^{-1}B$ und x ist ein Eigenvektor von $A - BF$ zu λ .*

Beweis. Für Punkt i) kann man die Ackermann Formel aus Kapitel 3.5.2 nutzen.

Übungsaufgabe 5.2

□

Allgemein ist das Problem der Polvorgabe (so wie die meisten inversen Probleme) sehr schlecht konditioniert. So können kleinste Störungen in F die Eigenwerte von $A - BF$ in die rechte Halbebene wechseln lassen. Eine Stabilisierung des geschlossenen Systems kann also nicht garantiert werden. Bevor wir nun also zwei Algorithmen zur Polvorgabe betrachten, wollen wir uns kurz mit den Auswirkungen von Störungen in den Daten auf die Eigenwerte von $A - BF$ beschäftigen. Dazu geben wir folgende Definition.

Definition 3.40. Wir definieren den Abstand zur Nichtsteuerbarkeit als

$$d_{uc}(A, B) := \min_{\lambda \in \mathbb{C}} \sigma_n[A - \lambda I, B],$$

also den kleinsten Singulärwert der Matrix $[A - \lambda I, B]$. Weiter definieren wir für eine diagonalisierbare Matrix $A \in \mathbb{R}^{n,n}$ die (skalierte) spektrale Konditionszahl als $\kappa := \|G_0\| \|G_0^{-1}\|$, wobei die Spalten von G_0 normierte Eigenvektoren von A sind.

Satz 3.41 (Mehrmann/Xu, 1996). Sei (A, B) steuerbar mit $m = 1$ und $\mathcal{P} = \bar{\mathcal{P}} = \{\lambda_1, \dots, \lambda_n\}$ die Menge der gewünschten Pole. Dazu betrachten wir das gestörte Problem mit

$$\hat{A} := A + \delta_A, \quad \hat{B} := B + \delta_B, \quad \hat{\mathcal{P}} := \{\lambda_1 + \delta_1, \dots, \lambda_n + \delta_n\}.$$

Dabei nehmen wir an, dass sowohl die λ_j also auch die $\lambda_j + \delta_j$ paarweise verschieden sind. Weiter seien die Störungen klein im Sinne von

$$\|\delta_A\|, \|\delta_B\|, \|\delta_\lambda\| \leq \varepsilon < \frac{1}{4} d_{uc}(A, B).$$

Dabei bezeichnet δ_λ den Vektor $[\lambda_1, \dots, \lambda_n]^T$. Seien $F, \hat{F} \in \mathbb{R}^{1,n}$ die entsprechenden Feedback-Matrizen und $\hat{\kappa}$ die spektrale Konditionszahl der Matrix $\hat{A} - \hat{B}\hat{F}$. Dann gibt es zu jedem Eigenwert μ_j von $A - B\hat{F}$ einen Pol $\lambda_j \in \mathcal{P}$, sodass

$$|\lambda_j - \mu_j| \leq \varepsilon (1 + \hat{\kappa}(1 + \|\hat{F}\|)).$$

Der Satz zeigt, dass wenn $\hat{\kappa}$ oder $\|\hat{F}\|$ groß sind, dann müssen die Eigenwerte von $A - B\hat{F}$ nicht unbedingt nah an den gewünschten Polen liegen. Gleiches gilt, wenn (A, B) nah an der Nichtsteuerbarkeit ist. Dies kann sogar zur Folge haben, dass wir Eigenwerte von $A - B\hat{F}$ in der rechten Halbebene erhalten, obwohl wir den Algorithmus zur Stabilisierung einsetzen. In diesem Sinne ist das Problem der Polvorgabe nicht gut konditioniert.

3.5.1 Der Polvorgabe-Algorithmus von Miminis und Paige

Gegeben sei (A, B) mit $m = 1$. Ferner sei (A, B) steuerbar und in Hessenberg-Form, d.h., A ist in oberer Dreiecks-Form mit zusätzlicher Subdiagonalen und $B = \beta e_1$. Gesucht ist ein $f \in \mathbb{R}^{1,n}$, sodass $\Lambda(A - Bf) = \mathcal{P} = \bar{\mathcal{P}} = \{\mu_1, \dots, \mu_n\}$.

Unterscheiden sich zwei Matrizen A und \tilde{A} nur in der ersten Zeile und ist $A = RQ$ eine RQ-Zerlegung von A ist, dann gibt es eine RQ-Zerlegung von \tilde{A} der Form $\tilde{A} = \tilde{R}Q$, wobei R und \tilde{R} sich ebenfalls nur in der ersten Zeile unterscheiden.

Sei μ_1 kein Eigenwert von A . Wir betrachten die RQ-Zerlegung $A - \mu_1 I = RQ$, also $Q = R^{-1}(A - \mu_1 I)$. Die Matrix Q ist damit als Produkt einer Dreiecks-Matrix und einer Hessenberg-Matrix ebenfalls eine Hessenberg-Matrix. Da $B = \beta e_1$, unterscheidet sich die Matrix $A - Bf - \mu_1 I$ nur in der ersten Zeile von $A - \mu_1 I$. Somit gilt

$$A - Bf - \mu_1 I = \tilde{R}Q,$$

wobei sich R und \tilde{R} nur in der ersten Zeile unterscheiden. Es ist μ_1 ein Eigenwert von $A - Bf$ genau dann, wenn \tilde{R} singularär ist, d.h.,

$$\begin{aligned} 0 = \tilde{r}_{11} &= e_1^T \tilde{R} e_1 = e_1^T (A - Bf - \mu_1 I) Q^T e_1 \\ &= e_1^T (A - \mu_1 I) Q^T e_1 - e_1^T B f Q^T e_1 \\ &= e_1^T R e_1 - \beta f Q^T e_1 \\ &= r_{11} - \beta e_1^T Q f^T. \end{aligned}$$

Das gesuchte f erfüllt also die Gleichung $e_1^T Q f^T = r_{11}/\beta$. Damit hat f die Form

$$f^T = Q^T \begin{bmatrix} r_{11}/\beta \\ f_2^T \end{bmatrix}$$

für ein unbekanntes $f_2 \in \mathbb{R}^{1,n-1}$. Wir müssen also noch f_2 berechnen. Die Matrix $A - Bf$ ist ähnlich zu

$$Q(A - Bf)Q^T = Q(A - Bf - \mu_1 I)Q^T + \mu_1 I = Q\tilde{R} + \mu_1 I.$$

Dabei ist Q eine Hessenberg-Matrix, \tilde{R} eine obere Dreiecks-Matrix mit $\tilde{r}_{11} = 0$ und $\mu_1 I$ eine Diagonal-Matrix. Daraus folgt

$$Q(A - Bf)Q^T = \begin{bmatrix} \mu_1 & * \\ 0 & A_2 - B_2 f_2 \end{bmatrix}.$$

Dabei ist $A_2 = (QAQ^T)_{2:n,2:n}$ wieder in Hessenberg-Form. Weiterhin gilt wegen der Form von Q (Hessenberg) und B , dass

$$QB = \begin{bmatrix} * & \beta q_{21} & 0 & \dots & 0 \end{bmatrix}^T = \begin{bmatrix} * \\ B_2 \end{bmatrix}.$$

Somit hat B_2 wieder die Form $B_2 = \beta_2 e_1 \in \mathbb{R}^{n-1}$. Nun muss rekursiv f_2 bestimmt werden, sodass $\Lambda(A_2 - B_2 f_2) = \{\mu_2, \dots, \mu_n\}$ gilt. Im Fall $n = 2$, also $A_2 \in \mathbb{R}^{1 \times 1}$ vereinfacht sich dies, denn es muss gelten

$$\mu_2 \in \Lambda(A_2 - B_2 f_2) = A_2 - B_2 f_2.$$

Also kann $f_2 = (A_2 - \mu_2)/B_2$ gewählt werden.

Bemerkung 3.42. *Nimmt man an, dass A in unreduzierter Hessenberg-Form ist, so gilt für die RQ Zerlegung $r_{11} = 0$ genau dann, wenn A singulär ist. Dies ist von Vorteil, wenn man den Algorithmus für ein $\mu \in \mathcal{P} \cap \Lambda(A)$ verwendet.*

Beispiel 3.43. *Wir wollen F berechnen für*

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathcal{P} = \{-1, -2\}.$$

Es gilt $\Lambda(A) = \{1\}$ und $\beta = 1$. Zunächst machen wir eine RQ-Zerlegung von $A - \mu_1 I$ mit $\mu_1 = -1$,

$$R = \frac{1}{\sqrt{2}} \begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}, \quad Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

Algorithmus 3.2 Miminis und Paige (1982)

Input: (A, B) in Hessenberg-Form, $\mathcal{P} = \{\mu_1, \dots, \mu_n\}$

Output: f

```
1:  $A_1 := A, B = \beta_1 e_1$ 
2: for  $i = 1, \dots, n - 1$  do
3:   berechne eine RQ-Zerlegung von  $A_i - \mu_i I = R_i Q_i$ 
4:    $\tau_i = (R_i)_{11} / \beta_i$ 
5:    $\beta_{i+1} = \beta_i (Q_i)_{21}$ 
6:    $A_{i+1} = (Q_i A_i Q_i^T)_{2:n, 2:n}$ 
7: end for
8:  $f_n = (A_n - \mu_n) / \beta_n$ 
9: for  $i = n - 1, \dots, 1$  do
10:   $f_i = Q_i^T \begin{bmatrix} \tau_i \\ f_{i+1} \end{bmatrix}$ 
11: end for
12:  $f = f_1$ 
```

Somit hat f die Form

$$f^T = Q^T \begin{bmatrix} r_{11}/\beta_1 \\ f_2 \end{bmatrix}, \quad \frac{r_{11}}{\beta_1} = \frac{4}{\sqrt{2}} = 2\sqrt{2}.$$

Für die Berechnung von f_2 brauchen wir

$$A_2 = (Q A Q^T)_{22} = 1, \quad B_2 = \beta_2 = \beta_1 Q_{21} = \frac{1}{\sqrt{2}}.$$

Es gilt also $f_2 = (A_2 - \mu_2) / \beta_2 = 3\sqrt{2}$ und somit

$$f^T = Q^T \begin{bmatrix} r_{11}/\beta_1 \\ f_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} \\ 3\sqrt{2} \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

Die Matrix $A - Bf$ hat dann die Eigenwerte -1 und -2 .

Bemerkung 3.44. Die Kosten von Algorithmus 3.2 liegen bei $\mathcal{O}(n^4)$ Flops. Da die Matrizen, für die die RQ-Zerlegung berechnet wird, Hessenberg sind, kann dies jedoch verbessert werden. Somit erreicht man $\mathcal{O}(n^3)$ Flops. Zudem müssen in dem gegebenen Algorithmus μ_1, \dots, μ_n reell sein, da sonst Q_i komplex werden kann. Es existieren Abwandlungen, wo man auch für komplexe Werte von μ_i reell bleibt.

Bemerkung 3.45. Da nur Transformationen mit orthogonalen Matrizen durchgeführt werden, ist der Algorithmus numerisch stabil. Man beachte aber die schlechte Kondition des Problems der Polvorgabe (vgl. Satz 3.41).

3.5.2 Die Ackermann Formel

Die Ackermann Formel gibt eine Alternative zum Algorithmus von Miminis und Paige, um die Feedback-Matrix F zu berechnen. Sei dazu wieder (A, B) steuerbar, $m = 1$ und $\mathcal{P} = \bar{\mathcal{P}} = \{\mu_1, \dots, \mu_n\}$ die Menge der gewünschten Pole von $A - BF$. Sei T die Transformationsmatrix, die (A, B) auf Regelungsnormalform bringt, d.h.,

$$T^{-1}AT = C_1(A), \quad T^{-1}B = e_n.$$

Zunächst zeigen wir, dass $e_1^T T^{-1}K(A, B) = e_n^T$ gilt. (Übungsaufgabe 5.3a) Zudem gilt offensichtlich, dass

$$e_1^T (T^{-1}AT)^k = e_1^T C_1(A)^k = \begin{cases} e_{k+1}^T & \text{für } k = 0, \dots, n-1 \\ [-\alpha_0, \dots, -\alpha_{n-1}] & \text{für } k = n. \end{cases}$$

Mit diesen Vorbereitungen können wir nun folgenden Satz beweisen.

Satz 3.46. Sei (A, B) steuerbar mit $m = 1$. Ferner sei $\mathcal{P} = \bar{\mathcal{P}} = \{\mu_1, \dots, \mu_n\}$ mit dazugehörigen Polynom

$$\Psi(x) := \prod_{k=1}^n (x - \mu_k) = (x - \mu_1) \cdots (x - \mu_n).$$

Dann erfüllt die Feedback-Matrix $F := e_n^T K(A, B)^{-1} \Psi(A)$ die Eigenschaft $\Lambda(A - BF) = \mathcal{P}$.

Beweis. Übungsaufgabe 5.3b □

4 LQ-Optimalsteuerung

Betrachten wir das LTV-System

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(0) = x_0 \tag{4.1}$$

für $A(t) \in \mathbb{R}^{n,n}$ und $B(t) \in \mathbb{R}^{n,m}$. Ziel ist es nun, x nahe zur Null zu steuern. Dabei suchen wir die Steuerung $u \in \mathcal{U}_{\text{ad}}$, welche zusammen mit dem resultierenden x das folgende *Kostenfunktional* (**cost functional**) minimiert

$$J(x, u) := \frac{1}{2}x(t_*)^T Mx(t_*) + \frac{1}{2} \int_0^{t_*} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} Q(t) & S(t) \\ S^T(t) & R(t) \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt.$$

Hier sind $t_* > 0$ und $M, Q(t) \in \mathbb{R}^{n,n}$ symmetrisch, $R(t) \in \mathbb{R}^{m,m}$ symmetrisch, $S(t) \in \mathbb{R}^{n,m}$. Zudem setzen wir voraus, dass

$$M \geq 0, \quad R(t) > 0, \quad \begin{bmatrix} Q(t) & S(t) \\ S^T(t) & R(t) \end{bmatrix} \geq 0$$

für alle $t \in [0, t_*]$. Es werden also die Endabweichung $x(t_*)$ von 0 sowie die Abweichungen von $x(t)$ und $u(t)$ von 0 mit Kosten belegt. Erlaubt man $t_* = \infty$, dann ist die Stabilisierung natürlich asymptotisch gemeint.

Da das LTV-System linear in $x(t)$ und $u(t)$ ist und das Kostenfunktional quadratisch ist, spricht man vom Linear-Quadratischen (LQ) Optimalsteuerungsproblem. Der gemischte Term $x(t)^T S(t)u(t)$ wird in der Praxis selten benutzt. Dieser entsteht zum Beispiel, wenn man $y(t) = C(t)x(t) + D(t)u(t)$ gewichten will.

Beispiel 4.1. *Wir betrachten ein mechanisches System zweiter Ordnung, welches optimal gesteuert werden soll. Dieses lässt sich durch Umformung in die obige Form bringen. Die Kontrollsystem ist gegeben durch*

$$\begin{bmatrix} 2m & \\ & m \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} d \\ -d \end{bmatrix} + \begin{bmatrix} 0 \\ u \end{bmatrix}.$$

Ziel ist es, mit einem möglichst kleinem Eingang u (also möglichst kleinen Kosten) zu erreichen, dass x_1 einer bestimmten Trajektorie γ folgt. Die Ausgangsgleichung ist also gegeben durch $y = Cx = x_1$. Wir minimieren daher das Kostenfunktional

$$J(x, u) = \frac{1}{2}\|x_1(t_*) - \gamma(t_*)\|^2 + \frac{1}{2} \int_0^{t_*} \|x_1(t) - \gamma(t)\|^2 + \beta_0 \|u(t)\|^2 dt.$$

Bemerkung 4.2. *Sei u_* stückweise stetig und eine optimale Steuerung des Optimalsteuerungsproblems. Sei x_* die dazugehörige Lösung des LTI Systems*

$$\dot{x}(t) = Ax(t) + Bu_*(t), \quad x(0) = x_0.$$

Dann existiert ein Lagrange Multiplikator $\lambda(t) \in \mathbb{R}^n$, sodass

$$\begin{bmatrix} I_n & & \\ & -I_n & \\ & & 0 \end{bmatrix} \begin{bmatrix} \dot{x}(t) \\ \dot{\lambda}(t) \\ \dot{u}(t) \end{bmatrix} = \begin{bmatrix} A & 0 & B \\ Q & A^T & S \\ S^T & B^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ \lambda(t) \\ u(t) \end{bmatrix}, \quad \begin{aligned} x(0) &= x_0, \\ \lambda(t_*) &= Mx(t_*). \end{aligned}$$

Dabei handelt es sich um ein DAE-Randwertproblem, da wir einen Anfangswert für x aber einen (unbekannten) Endwert für λ vorgegeben haben. Numerisch einfacher und günstiger ist allerdings der Weg über die Riccati-Gleichungen, den wir hier betrachten wollen.

Eine alternative Lösung des Problems liefert der folgende Satz.

Satz 4.3 (Darstellungssatz). Sei (x, u) eine Lösung von (4.1). Zudem sei das Kostenfunktional J mit A, B, M, Q, R, S und t_* wie oben definiert und

$$\hat{P} := BR^{-1}B^T, \quad \hat{A} = A - BR^{-1}S^T, \quad \hat{Q} = Q - SR^{-1}S^T.$$

Weiter sei $W(t) = W(t)^T \in \mathbb{R}^{n,n}$ eine symmetrische Lösung der Riccati-Differentialgleichung

$$\dot{W}(t) = W(t)\hat{P}(t)W(t) - \hat{A}(t)^T W(t) - W(t)\hat{A}(t) - \hat{Q}(t). \quad (4.2)$$

Mit $F(t) := R(t)^{-1}(S(t)^T + B(t)^T W(t))$ gilt dann

$$\begin{aligned} J(x, u) &= \frac{1}{2}x_0^T W(0)x_0 + \frac{1}{2}x(t_*)^T (M - W(t_*))x(t_*) \\ &\quad + \frac{1}{2} \int_0^{t_*} (u(t) + F(t)x(t))^T R(t)(u(t) + F(t)x(t)) dt. \end{aligned}$$

Beweis. Wir zeigen, dass für die angegebene Wahl von $W(t)$ und $F(t)$ folgende Gleichung gilt,

$$\text{LHS} := \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \frac{d}{dt}(x^T W x) = (u + Fx)^T R(u + Fx) =: \text{RHS}.$$

Daraus folgt dann mit Integration von 0 bis t_* die Behauptung. Es gilt

$$\begin{aligned}
\frac{d}{dt}(x^T W x) &= \dot{x}^T W x + x^T \dot{W} x + x^T W \dot{x} \\
&= \begin{bmatrix} x \\ \dot{x} \end{bmatrix}^T \begin{bmatrix} \dot{W} & W \\ W & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix} \\
&= \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} I & 0 \\ A & B \end{bmatrix}^T \begin{bmatrix} \dot{W} & W \\ W & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ A & B \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\
&= \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T W + W A + \dot{W} & W B \\ B^T W & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}.
\end{aligned}$$

Damit folgt, dass die linke Seite gegeben ist durch

$$LHS = \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} Q + A^T W + W A + \dot{W} & S + W B \\ S^T + B^T W & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}.$$

Für die rechte Seite ergibt sich

$$RHS = \begin{bmatrix} x \\ u \end{bmatrix}^T [F \ I]^T R [F \ I] \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} F^T R F & F^T R \\ R F & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}.$$

Nach Definition von F gilt $RF = S^T + B^T W$. Benutzt man noch die Eigenschaft, dass W Lösung der Riccati-Differentialgleichung (4.2) ist, so ergibt sich

$$\begin{aligned}
Q + A^T W + W A + \dot{W} &= Q - \hat{Q} + W \hat{P} W + (A - \hat{A})^T W + W(A - \hat{A}) \\
&= F^T R F.
\end{aligned}$$

und damit $LHS = RHS$. □

Bemerkung 4.4. Die Aussage aus Satz 4.3 gilt für jede symmetrische Lösung der Riccati-DGL (4.2). Insbesondere gilt sie für jene, die $W(t_*) = M$ erfüllen. In dem Fall fällt der Term $\frac{1}{2}x(t_*)^T(M - W(t_*))x(t_*)$ weg. Der Integrand ist nach Voraussetzung nicht negativ und verschwindet für $u(t) = -F(t)x(t)$. Damit ist der optimale Wert von J gegeben durch

$$J_{opt} = x_0^T W(0)x_0.$$

Die optimale Steuerung $u(t) = -F(t)x(t)$ ist also eine Feedback-Steuerung.

Wir suchen nun also Lösungen der Riccati-DGL mit $W(t_*) = M$. Die Berechnung der optimalen Zustands-Trajektorie $x(t)$ bedarf also zunächst der Lösung eines Riccati-Endwertproblems rückwärts in der Zeit. Dies ist nötig zur Bestimmung von F . Danach löst man dann das Anfangswertproblem des LTV Systems

$$\dot{x}(t) = (A(t) - B(t)F(t))x(t), \quad x(0) = x_0$$

vorwärts in der Zeit.

Wir betrachten noch den Spezialfall konstanter Matrizen A , B , Q , S und R , also den zeitinvarianten Fall. Da die Lösung W der Riccati-DGL (4.2) zeitvariant ist, wäre nach der bisher bekannten Lösung F trotzdem zeitabhängig. Wir wollen nun eine Lösung finden, sodass F konstant ist. Wir gehen vor wie folgt:

- Anstatt $W(t_*) = M$ zu setzen, suchen wir eine Lösung der *algebraischen Riccati-Gleichung*

$$0 = W\hat{P}W - \hat{A}^T W - W\hat{A} - \hat{Q}. \quad (4.3)$$

Dann gilt $\dot{W} = 0$ und damit sind W sowie F konstant.

- Um den störenden Term $x(t_*)^T(M - W(t_*))x(t_*)$ verschwinden zu lassen, sorgen wir dafür, dass $x(t_*) = 0$ gilt. Dazu nehmen wir an, dass F stabilisierend für (A, B) ist und betrachten $t_* \rightarrow \infty$.

Satz 4.5. Seien $A, Q \in \mathbb{R}^{n,n}$, $B, S \in \mathbb{R}^{n,m}$ und $R \in \mathbb{R}^{m,m}$. Weiter seien Q und R symmetrisch sowie

$$R > 0, \quad \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \geq 0.$$

Wir definieren wieder

$$\hat{P} := BR^{-1}B^T, \quad \hat{A} := A - BR^{-1}S^T, \quad \hat{Q} := Q - SR^{-1}S^T$$

und W als eine symmetrische Lösung der algebraischen Riccati-Gleichung (4.3), sodass $F := R^{-1}(S^T + B^T W)$ stabilisierend für (A, B) ist. Dann wird das Minimum von

$$J(x, u) = \frac{1}{2} \int_0^\infty \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt$$

unter $\dot{x} = Ax + Bu$ mit $x(0) = x_0$ für die Feedback-Steuerung $u(t) := -Fx(t)$ angenommen. Das Minimum ist gegeben durch $\frac{1}{2}x_0^T W x_0$.

Beweis. Übungsaufgabe 6.1 □

Bemerkung 4.6. *Der vorherige Satz liefert insbesondere eine weitere Methode der Stabilisierung. Setze dazu zum Beispiel $S = 0$, $Q = I$, $R = I$. Dies ergibt $u(t) = -Fx(t) = -B^T W x(t)$, wobei W Lösung der algebraischen Riccati-Gleichung (4.3) ist, die in diesem Fall die Form*

$$0 = WBB^T W - A^T W - W A - Q.$$

hat. Im Fall $S = 0$, $Q = 0$, $R = \frac{1}{2}I$ ist die algebraische Riccati-Gleichung gegeben durch

$$0 = 2WBB^T W - A^T W - W A \quad \Leftrightarrow \quad AW^{-1} + W^{-1}A^T = 2BB^T.$$

Die Feedback-Matrix ist gleich $F = 2B^T W$. Dies ist also – bis auf den Vorfaktor 2 – nichts anderes als der Algorithmus von Bass.

Dass (A, B) in Satz 4.5 stabilisierbar ist, ist offensichtlich notwendig, da man sonst kein stabilisierendes F finden kann. Weiter gilt

$$(A, B) \text{ stabilisierbar} \Leftrightarrow (\hat{A}, B) \text{ stabilisierbar} \Leftrightarrow (\hat{A}, \hat{P}) \text{ stabilisierbar.}$$

(Übungsaufgabe (6.2a)) Häufig nimmt man an, dass die Matrix mit Q , S , R positiv definit ist. Dies ist zwar nicht notwendig, aber hilfreich. Daraus folgt dann unter anderem, dass $\hat{Q} = Q - SR^{-1}S^T > 0$ gilt (Schurkomplement). Weiter folgt daraus aber auch, dass für einen endlichen Wert $J(x, u)$ notwendigerweise u stabilisierend wirkt, da sonst der Integrand nicht gegen 0 konvergiert. Ist die Matrix nicht positiv definit, dann kann der folgende Fall auftreten: Gilt $Q = 0$ und $S = 0$, dann steht im Integral nur $u^T R u$, wobei R positiv definit ist. Somit ist $u \equiv 0$ ein Minimierer von J ohne, dass $x \rightarrow 0$ gelten muss.

4.1 Die algebraische Riccati-Gleichung (ARE)

In diesem Abschnitt beschäftigen wir uns mit der Lösung der algebraischen Riccati-Gleichung (ARE). Man beachte, dass diese im Gegensatz zur Riccati-DGL (mit Anfangsbedingung) unendlich viele Lösungen besitzen kann. Auch unsymmetrische Lösungen sind möglich.

Bemerkung 4.7. *Riccati war ein italienischer Mathematiker. Das Wort Riccati ist das italienische Wort für Erpressung.*

Gegeben seien die Matrix $\hat{A} \in \mathbb{R}^{n,n}$ und die symmetrischen Matrizen $\hat{P}, \hat{Q} \in \mathbb{R}^{n,n}$ aus Satz 4.5. Wir suchen eine symmetrische Lösung W der algebraischen Riccati-Gleichung

$$0 = W\hat{P}W - \hat{A}^T W - W\hat{A} - \hat{Q}, \quad (4.4)$$

die stabilisierend ist, d.h., $\Lambda(\hat{A} - \hat{P}W) \subseteq \mathbb{C}^-$. Das impliziert dann $\Lambda(A - BF) \subseteq \mathbb{C}^-$ mit $F = R^{-1}(S^T + B^T W)$, da

$$A - BF = A - BR^{-1}(S^T + B^T W) = \hat{A} - BR^{-1}B^T W = \hat{A} - \hat{P}W.$$

Eine Frage ist, ob die Stabilisierbarkeit von (\hat{A}, \hat{P}) zusammen mit $\hat{Q} > 0$ hinreichend für die Existenz einer stabilisierenden Lösung der ARE ist.

Offensichtlich gilt: W ist genau dann eine Lösung von (4.4), wenn

$$\mathcal{H} \begin{bmatrix} I \\ W \end{bmatrix} := \begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \begin{bmatrix} I \\ W \end{bmatrix} = \begin{bmatrix} I \\ W \end{bmatrix} (\hat{A} - \hat{P}W).$$

Aus dieser Darstellung folgt mit Lemma 1.22, dass das Bild von $[I, W]^T$ ein n -dimensionaler \mathcal{H} -invarianter Unterraum ist. Ist W eine stabilisierende Lösung, dann gilt $\Lambda(\hat{A} - \hat{P}W) \subseteq \mathbb{C}^-$ und somit gehört der n -dimensionale Unterraum zu Eigenwerten in \mathbb{C}^- . Hier nutzen wir, dass $[I, W]^T$ vollen Spaltenrang hat. Wir wollen die Frage untersuchen, ob auch die Rückrichtung gilt. Dafür betrachten wir zunächst Eigenschaften von \mathcal{H} .

Definition 4.8. Eine Matrix $\mathcal{H} \in \mathbb{R}^{2n,2n}$ heißt Hamiltonisch, falls für

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

die Gleichheit $\mathcal{H}^T J + J\mathcal{H} = 0$ bzw. $\mathcal{H}^T = J\mathcal{H}J$ gilt.

Lemma 4.9. Sei $\mathcal{H} \in \mathbb{R}^{2n,2n}$ Hamiltonisch. Dann ist \mathcal{H} ähnlich zu $-\mathcal{H}$. Zudem ist das Spektrum von \mathcal{H} symmetrisch bezüglich der reellen und imaginären Achse.

Beweis. Übungsaufgabe 6.2b □

Bemerkung 4.10. Das Spektrum einer Hamiltonischen Matrix besteht also aus den komplexen Quadrupeln $\lambda, -\lambda, -\bar{\lambda}, \bar{\lambda}$. Für rein reelle oder imaginäre Eigenwerte reduziert sich dies zu $\lambda, -\lambda$. Falls Null-Eigenwerte auftreten, so haben diese stets eine gerade algebraische Vielfachheit. Insbesondere gibt es genauso viele Eigenwerte in \mathbb{C}^- wie in \mathbb{C}^+ . Es gibt also genau dann n Eigenwerte in \mathbb{C}^- , wenn es keinen Eigenwert auf der imaginären Achse gibt.

Satz 4.11. *Sei*

$$\mathcal{H} = \begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \in \mathbb{R}^{2n,2n} \quad (4.5)$$

Hamiltonisch und das Bild von $[I, W^T]^T$ ein n -dimensionaler \mathcal{H} -invarianter Unterraum zu den Eigenwerten in \mathbb{C}^- . Dann ist W eine symmetrische, stabilisierende Lösung der algebraischen Riccati-Gleichung (4.4). Gilt zusätzlich $\hat{Q}, \hat{P} \geq 0$, dann ist W positiv semidefinit. Bei $\hat{Q} > 0, \hat{P} \geq 0$ gilt sogar $W > 0$.

Beweis. Nach Voraussetzung existiert eine Matrix $B \in \mathbb{R}^{n,n}$ mit $\Lambda(B) \subseteq \mathbb{C}^-$, sodass

$$\mathcal{H} \begin{bmatrix} I \\ W \end{bmatrix} = \begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \begin{bmatrix} I \\ W \end{bmatrix} = \begin{bmatrix} I \\ W \end{bmatrix} B. \quad (4.6)$$

Die erste Block-Gleichung besagt $\hat{A} - \hat{P}W = B$, sodass W also stabilisierend ist. Zusammen mit der zweiten Gleichung folgt

$$-\hat{Q} - \hat{A}^T W = WB = W\hat{A} - W\hat{P}W,$$

was bedeutet, dass W (4.4) löst. Wir zeigen nun, dass W symmetrisch ist. Dazu multiplizieren wir die erste Zeile von (4.6) mit W^T von links und erhalten $W^T \hat{A} - W^T \hat{P}W = W^T B$. Nun betrachten wir die Transponierte der zweiten Zeile von (4.6) und nutzen aus, dass \hat{Q} symmetrisch ist. Das liefert $-\hat{Q} - W^T \hat{A} = B^T W^T$. Addieren wir beide Gleichungen, so ergibt sich

$$-M := -\hat{Q} - W^T \hat{P}W = W^T B + B^T W^T.$$

Dabei ist $-M$ symmetrisch. Dies ist eine Lyapunov-Gleichung für W^T und die bekannte Lösungsformel aus Satz 3.7 ergibt

$$W^T = \int_0^\infty e^{B^T s} M e^{Bs} ds.$$

Daraus folgt, dass W symmetrisch ist.

Schließlich untersuchen wir noch die Definitheit von W . Gilt $\hat{Q}, \hat{P} \geq 0$, dann folgt $M \geq 0$ und somit auch $W \geq 0$. Analog impliziert $\hat{Q} > 0, \hat{P} \geq 0$, dass $M > 0$ und somit $W > 0$. \square

Nun stellt sich die Frage, wann \mathcal{H} genau n Eigenwerte in \mathbb{C}^- hat.

Satz 4.12. Sei \mathcal{H} die Hamiltonische Matrix (4.5) aus Satz 4.11. Gelten die beiden Bedingungen

- $\hat{P} > 0$ oder ($\hat{P} \geq 0$ und (\hat{A}, \hat{P}) stabilisierbar),
- $\hat{Q} > 0$ oder ($\hat{Q} \geq 0$ und (\hat{A}, \hat{Q}) entdeckbar)

so ist $\Re(\lambda) \neq 0$ für alle Eigenwerte $\lambda \in \Lambda(\mathcal{H})$.

Beweis. Sei $i\omega \in \Lambda(\mathcal{H})$ für ein $\omega \in \mathbb{R}$ und $[x_1^T, x_2^T]^T$ ein zugehöriger Eigenvektor, wobei $x_1, x_2 \in \mathbb{C}^n$. Dann gilt

$$\begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = i\omega \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Multiplikation von links mit $[x_2^*, x_1^*]$ liefert für die rechte Seite

$$[x_2^* \ x_1^*] i\omega \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = i\omega(x_2^*x_1 + x_1^*x_2) = 2i\omega \Re(x_2^*x_1) \in i\mathbb{R}.$$

Auf der linken Seite ergibt sich

$$[x_2^* \ x_1^*] \begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2^*\hat{A}x_1 - x_1^*\hat{A}^T x_2 - x_2^*\hat{P}x_2 - x_1^*\hat{Q}x_1.$$

Da $x_2^*\hat{A}x_1 - x_1^*\hat{A}^T x_2 \in i\mathbb{R}$ und $x_2^*\hat{P}x_2 - x_1^*\hat{Q}x_1 \in \mathbb{R}$, muss $x_2^*\hat{P}x_2 + x_1^*\hat{Q}x_1 = 0$ gelten. Da sowohl $\hat{P} \geq 0$ als auch $\hat{Q} \geq 0$, gilt

$$x_2^*\hat{P}x_2 + x_1^*\hat{Q}x_1 = 0 \quad \Rightarrow \quad x_2^*\hat{P}x_2 = x_1^*\hat{Q}x_1 = 0 \quad \Rightarrow \quad \hat{Q}x_1 = \hat{P}x_2 = 0.$$

Wir zeigen nun, dass $x_1 = 0 = x_2$ gilt. Gilt $\hat{P} > 0$, dann folgt direkt $x_2 = 0$. Nehmen wir also an, dass (\hat{A}, \hat{P}) stabilisierbar ist. Daraus folgt, dass $\text{Rang} [\hat{A} - \lambda I \ \hat{P}] = n$ für alle λ mit $\Re(\lambda) \geq 0$. Aus der zweiten Zeile der Eigenwert-Gleichung folgt

$$i\omega x_2 = -\hat{Q}x_1 - \hat{A}^T x_2 = -\hat{A}^T x_2.$$

Mit konjugierter Transposition führt dies auf

$$x_2^*(\hat{A} - i\omega I) = 0 \quad \Rightarrow \quad x_2^* [\hat{A} - i\omega I \ \hat{P}] = 0.$$

Dies steht im Widerspruch zur Stabilisierbarkeit, außer wenn $x_2 = 0$ gilt.

Analog zeigt man $x_1 = 0$. Dazu nutzt man, dass Stabilisierbarkeit und Entdeckbarkeit duale Konzepte sind. \square

Bemerkung 4.13. Durch Betrachten von $-\mathcal{H}$ können in Satz 4.12 die Relationszeichen $\geq, >$ durch $\leq, <$ ersetzt werden.

Bemerkung 4.14. Die Voraussetzungen in Satz 4.12 sind nicht notwendig. Dazu betrachten wir den Fall $\hat{A} = I$ und $\hat{Q} = 0$. Dieses System ist nicht entdeckbar, aber

$$\mathcal{H} = \begin{bmatrix} I & * \\ 0 & -I \end{bmatrix}$$

hat die Eigenwerte ± 1 .

Nun bleibt noch eine Frage. Wenn \mathcal{H} einen n -dimensionalen invarianten Unterraum zu Eigenwerten in \mathbb{C}^- hat, ist dieser dann immer von der Form $[I, W]^T$?

Satz 4.15. Seien $\hat{A}, \hat{P}, \hat{Q} \in \mathbb{R}^{n,n}$, wobei \hat{P}, \hat{Q} symmetrisch sind, $\hat{P} \geq 0$, $\hat{Q} > 0$ gilt und (\hat{A}, \hat{P}) stabilisierbar ist. Des Weiteren sei \mathcal{H} wieder die Hamiltonische Matrix (4.5) und das Bild von $[U^T, V^T]^T$ für $U, V \in \mathbb{R}^{n,n}$ ein n -dimensionaler invarianter Unterraum von \mathcal{H} zu den Eigenwerten in \mathbb{C}^- . Dann ist U invertierbar und

$$\text{Bild} \begin{bmatrix} U \\ V \end{bmatrix} = \text{Bild} \begin{bmatrix} I \\ W \end{bmatrix}$$

für $W = VU^{-1}$.

Beweis. Nach den Voraussetzungen gibt es eine Matrix $B \in \mathbb{R}^{n,n}$ mit $\Lambda(B) \subseteq \mathbb{C}^-$, sodass

$$\begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} B. \quad (4.7)$$

Wir multiplizieren die erste Gleichung mit V^T von links und das Transponierte der zweiten Gleichung mit U von rechts. Da \hat{P} und \hat{Q} symmetrisch sind, ergibt die Summe dieser Gleichungen

$$-V^T \hat{P} V - U^T \hat{Q} U = V^T U B + B^T V^T U. \quad (4.8)$$

Somit ist $U^T V = V^T U$ die symmetrische Lösung einer Lyapunov-Gleichung.

Wir zeigen nun, dass $\text{Ker } U$ ein B -invarianter Unterraum ist. Sei $w \in \text{Ker } U$. Dann gilt

$$UBw = (\hat{A}U - \hat{P}V)w = -\hat{P}Vw.$$

Mit Gleichung (4.8) ergibt sich

$$\begin{aligned} w^T V^T \hat{P} V w &= w^T V^T \hat{P} V w + w^T U^T \hat{Q} U w \\ &= -w^T (V^T U B + B^T V^T U) w \\ &= -w^T (U^T V B + B^T V^T U) w = 0. \end{aligned}$$

Wegen $\hat{P} \geq 0$ folgt dann, dass $0 = \hat{P} V w = -U B w$ gilt. Damit ist also $B w \in \text{Ker } U$.

Als nächstes zeigen wir, dass $\text{Ker } U = \{0\}$ gilt. Wir führen einen Widerspruchsbeweis und nehmen an, dass der Kern nicht nur den Nullvektor enthält. Da $\text{Ker } U$ ein B -invarianter Unterraum ist, gibt es ein $0 \neq z \in \text{Ker } U$, sodass

$$B z = \lambda z.$$

(Übungsaufgabe (optional)) Dabei gilt $\lambda \in \mathbb{C}^-$. Multiplizieren wir (4.7) mit z , dann ergibt sich für $v := V z$,

$$\lambda \begin{bmatrix} 0 \\ v \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} B z = \begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} z = \begin{bmatrix} -\hat{P} v \\ -\hat{A}^T v \end{bmatrix}.$$

Es gilt also

$$v^T [\hat{A} + \lambda I \quad \hat{P}] = 0.$$

Dies steht im Widerspruch zur Stabilisierbarkeit von (\hat{A}, \hat{P}) . Man beachte, dass $v \neq 0$ gelten muss, da sonst $[U^T, V^T]^T z = 0$ gelten würde, was ein Widerspruch dazu ist, dass Bild $[U^T, V^T]^T$ die Dimension n hat. \square

Wir fassen die bisherigen Ergebnisse in einem Satz zusammen.

Satz 4.16. *Seien $\hat{A}, \hat{P}, \hat{Q} \in \mathbb{R}^{n,n}$. Seien zudem $\hat{P} \geq 0$, $\hat{Q} > 0$ und (\hat{A}, \hat{P}) stabilisierbar. Dann hat die algebraische Riccati-Gleichung (4.4) mindestens eine Lösung. Zudem ist genau eine Lösung W stabilisierend, symmetrisch und positiv definit.*

Beweis. Folgt aus den vorherigen Resultaten zusammen mit der Eigenschaft, dass es maximal eine stabilisierende Lösung gibt. \square

Algorithmus 4.1 Laub (1979)

Input: Die zur ARE (4.4) gehörige Hamiltonische Matrix \mathcal{H} (ohne Eigenwerte in $i\mathbb{R}$).

Output: Stabilisierende Lösung $W \geq 0$ der ARE

1: Berechne die reelle Schur-Form

$$Q^T \mathcal{H} Q = R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$$

mit einer orthogonalen Matrix Q . Dabei sollen die Diagonaleinträge bzw. Diagonalblöcke von R so geordnet werden, dass $\Lambda(R_{11}) \subseteq C^-$ und $\Lambda(R_{22}) \subseteq C^+$ gilt.

2: Partitioniere $Q = \begin{bmatrix} U & Q_{12} \\ V & Q_{22} \end{bmatrix}$

3: **if** U singular **then** STOP

4: **end if**

5: $W := VU^{-1}$

Bemerkung 4.17. Die Kosten für den Algorithmus 4.1 liegen bei $\mathcal{O}(n^3)$. Die numerischen Eigenschaften bezüglich der Stabilität sind nur mittelmäßig. Es gibt Probleme, wenn Eigenwerte nah an der imaginären Achse liegen. Durch Rundungsfehler kann es dann passieren, dass ein stabiler Eigenwert instabil wird. Auch wenn es theoretisch bekannt ist, dass einer der Eigenwerte aufgrund der Symmetrie stabil sein muss, so ist dies praktisch nicht zu entscheiden.

Bemerkung 4.18. Alternativ zum Algorithmus 4.1 kann man auch eine spezielle orthogonale Transformationsmatrix Q benutzen, eine sogenannte orthogonal-symplektische Matrix. Leider gibt es keinen vollständig zufriedenstellenden Algorithmus zur Berechnung solch einer Matrix Q . Der aktuell beste ist von Chu, Liu und Mehrmann (2007) und wurde von Watkins, Mehrmann und Schröder (2009) effizienter gemacht.

Definition 4.19. Eine orthogonale Matrix $Q \in \mathbb{R}^{2n,2n}$ heißt orthogonal-symplektisch, wenn $U, V \in \mathbb{R}^{n,n}$ existieren, sodass

$$Q = \begin{bmatrix} U & -V \\ V & U \end{bmatrix}.$$

Lemma 4.20. *Die Menge orthogonal-symplektischer Matrizen formen eine Gruppe bezüglich der Matrix Multiplikation. Zudem erhalten Transformationen mit einer orthogonal-symplektischer Matrix Q die Eigenschaft Hamiltonisch zu sein, d.h., für eine Hamiltonische Matrix \mathcal{H} ist auch $Q^T \mathcal{H} Q$ Hamiltonisch.*

Beweis. Übungsaufgabe 6.2a □

Lemma 4.21. *Sei \mathcal{H} Hamiltonisch ohne Eigenwerte auf der imaginären Achse. Dann gibt es eine orthogonal-symplektische Matrix*

$$Q = \begin{bmatrix} U & -V \\ V & U \end{bmatrix}, \text{ sodass } Q^T \mathcal{H} Q = \begin{bmatrix} R_{11} & R_{12} \\ 0 & -R_{11}^T \end{bmatrix}$$

Hamiltonisch ist, d.h., $R_{12} = R_{12}^T$. Zudem ist R_{11} in reeller Schur-Form mit $\Lambda(R_{11}) \subseteq \mathbb{C}^-$.

Beweis. Da es keine rein imaginären Eigenwerte gibt, gibt es einen genau n -dimensionalen \mathcal{H} -invarianten Unterraum zu den Eigenwerten in \mathbb{C}^- . Es gibt also $U, V \in \mathbb{C}^{n,n}$ und eine Matrix $B \in \mathbb{C}^{n,n}$, sodass

$$\mathcal{H} \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} U \\ V \end{bmatrix} B. \tag{4.9}$$

Wir können annehmen, dass $[U^T, V^T]^T$ eine Orthonormalbasis des Unterraumes ist. Dies impliziert $U^T U + V^T V = I$. Wir betrachten nun die reelle Schur-Form von B , d.h., $B = T S_B T^{-1}$ mit orthogonaler Matrix T . Damit erhalten wir

$$\mathcal{H} \begin{bmatrix} UT \\ VT \end{bmatrix} = \begin{bmatrix} UT \\ VT \end{bmatrix} S_B.$$

Man beachte, dass dies nicht die Orthonormalität zerstört, da

$$\begin{bmatrix} UT \\ VT \end{bmatrix}^T \begin{bmatrix} UT \\ VT \end{bmatrix} = [T^T U^T, T^T V^T] \begin{bmatrix} UT \\ VT \end{bmatrix} = T^T (U^T U + V^T V) T = T^T T = I.$$

Wir können in (4.9) also annehmen, dass $[U^T, V^T]^T$ eine Orthonormalbasis des Unterraumes und B in reeller Schur-Form gegeben ist. Wie in Satz 4.11 zeigt man $U^T V = V^T U$. Dies impliziert dann, dass die Matrix

$$Q := \begin{bmatrix} U & -V \\ V & U \end{bmatrix}$$

orthogonal ist. Somit ist Q eine orthogonal-symplektische Matrix. Wir erhalten

$$\mathcal{H} \begin{bmatrix} U & -V \\ V & U \end{bmatrix} = \begin{bmatrix} U & -V \\ V & U \end{bmatrix} \begin{bmatrix} B & R_{12} \\ 0 & R_{22} \end{bmatrix} =: QR,$$

wobei $R = Q^{-1}\mathcal{H}Q$ nach Lemma 4.20 Hamiltonisch ist. Das impliziert $R_{22} = -B^T$ und $R_{12} = R_{12}^T$. [Übungsaufgabe 6.2b](#) \square

Im Folgenden Abschnitt wollen wir die Newton-Methode zur Lösung der algebraischen Riccati-Gleichung benutzen.

4.2 Newton-Methode zur Lösung der ARE

Wir beginnen mit der Definition der Fréchet-Ableitung.

Definition 4.22. *Seien \mathcal{X}, \mathcal{Z} normierte Räume und $f: \mathcal{X} \rightarrow \mathcal{Z}$ ein Operator. Falls für $x, y \in \mathcal{X}$ der Grenzwert*

$$f'_x(y) := \lim_{t \rightarrow 0} \frac{f(x + ty) - f(x)}{t}$$

existiert, so heißt dieser Fréchet-Ableitung von f an der Stelle x in Richtung y .

Im Falle der algebraischen Riccati-Gleichung definieren wir

$$f(W) := W\hat{P}W - \hat{A}^T W - W\hat{A} - \hat{Q} = 0.$$

Dabei sind nach Definition \hat{P} und \hat{Q} symmetrisch. Nehmen wir an, dass W symmetrisch ist, dann gilt

$$\begin{aligned} f'_W(Y) &= \lim_{t \rightarrow 0} \frac{1}{t} \left[(W + tY)\hat{P}(W + tY) - \hat{A}^T(W + tY) - (W + tY)\hat{A} - \hat{Q} \right. \\ &\quad \left. - (W\hat{P}W - \hat{A}^T W - W\hat{A} - \hat{Q}) \right] \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left[t^2 Y\hat{P}Y + t(Y\hat{P}W + W\hat{P}Y - \hat{A}^T Y - Y\hat{A}) \right] \\ &= -(\hat{A} - \hat{P}W)^T Y - Y(\hat{A} - \hat{P}W). \end{aligned}$$

Damit lässt sich nun die Newton Iteration zum Finden von Nullstellen anwenden. Wir betrachten also die Iteration $W_{k+1} = W_k + Y_k$, wobei das Update Y_k

durch $-f'_{W_k}(Y_k) = f(W_k)$ berechnet wird. Dies führt dann auf den folgenden Algorithmus.

Algorithmus 4.2 Newton-Methode zur Lösung der ARE

Input: $\hat{A}, \hat{P}, \hat{Q}$

Output: W_∞

- 1: Wähle W_0 symmetrisch
- 2: **for** $k = 0, 1, \dots$ **do**
- 3: Löse

$$(\hat{A} - \hat{P}W_k)^T Y_k + Y_k(\hat{A} - \hat{P}W_k) = f(W_k) \quad (4.10)$$

- 4: nach Y_k und wähle $W_{k+1} = W_k + Y_k$.
 - 5: **end for**
-

Bemerkung 4.23. Die Kosten einer Iteration von Algorithmus 4.2 sind abhängig von der Methode, die zur Lösung der Lyapunov-Gleichung verwendet wird. Bei Bartels-Stewart oder der Signums-Methode war dies $\mathcal{O}(n^3)$. Allerdings werden oft nur wenige Iterationen gebraucht. Benutzt wird dieser Algorithmus oft, um eine gegebene Näherungslösung zu verbessern. In endlicher Arithmetik erreicht diese Methode den bestmöglichen relativen Fehler

$$\frac{\|\tilde{W} - W_*\|}{\|W_*\|} \approx \kappa_{ARE} \cdot \varepsilon.$$

Dabei ist κ_{ARE} die Konditionszahl der Lyapunov-Gleichung aus Definition 3.15 und ε die Maschinengenauigkeit.

Es bleibt die Frage nach der Konvergenz von Algorithmus 4.2.

Satz 4.24. Seien $\hat{A}, \hat{P}, \hat{Q} \in \mathbb{R}^{n,n}$, wobei $\hat{Q} > 0$ und $\hat{P} \geq 0$ symmetrisch sind und (\hat{A}, \hat{P}) stabilisierbar ist. Sei zudem $W_0 \in \mathbb{R}^{n,n}$ symmetrisch und stabilisierend für (\hat{A}, \hat{P}) , d.h. $\Lambda(\hat{A} - \hat{P}W_0) \subseteq \mathbb{C}^-$. Bezeichnet (W_k) die Folge der Newton-Iterierten aus Algorithmus 4.2 und W_* die stabilisierende Lösung der ARE $f(W) = 0$, dann gilt

- i) $\Lambda(\hat{A} - \hat{P}W_k) \subseteq \mathbb{C}^-$ und $W_k = W_k^T$ für alle $k \geq 0$,
- ii) $0 < W_* \leq \dots \leq W_2 \leq W_1$,

$$iii) \lim_{k \rightarrow \infty} W_k = W_*,$$

iv) das Verfahren konvergiert quadratisch, d.h., es gibt einen Wert $\gamma > 0$, sodass für $k \geq 1$,

$$\|W_* - W_{k+1}\| \leq \gamma \|W_* - W_k\|^2.$$

Beweis. Existenz, Eindeutigkeit und positive Definitheit von W_* wurde bereits im vorigen Kapitel gezeigt.

i) Induktion über k . Für $k = 0$ gilt dies nach Voraussetzung. Nun gelte $W_{k-1} = W_{k-1}^T$ und $\Lambda(\hat{A} - \hat{P}W_{k-1}) \subseteq \mathbb{C}^-$. Mit (4.10) und $Y_{k-1} = W_k - W_{k-1}$ gilt

$$\begin{aligned} & (\hat{A} - \hat{P}W_{k-1})^T(W_k - W_{k-1}) + (W_k - W_{k-1})(\hat{A} - \hat{P}W_{k-1}) \\ & \quad = W_{k-1}\hat{P}W_{k-1} - \hat{A}^T W_{k-1} - W_{k-1}\hat{A} - \hat{Q} \\ \Leftrightarrow & \quad (\hat{A} - \hat{P}W_{k-1})^T W_k + W_k(\hat{A} - \hat{P}W_{k-1}) = \underbrace{-W_{k-1}\hat{P}W_{k-1} - \hat{Q}}_{<0}. \end{aligned} \tag{4.11}$$

Nach Satz 3.9 über Lyapunov-Gleichungen folgt, dass $W_k = W_k^T > 0$ gilt. Formen wir nun (4.11) weiter um, so gilt

$$\hat{A}^T W_k + W_k \hat{A} = -W_{k-1} \hat{P} W_{k-1} - \hat{Q} + W_{k-1} \hat{P} W_k + W_k \hat{P} W_{k-1}.$$

Wir ziehen $2W_k \hat{P} W_k$ auf beiden Seiten ab und erhalten

$$\begin{aligned} & (\hat{A} - \hat{P}W_k)^T W_k + W_k(\hat{A} - \hat{P}W_k) \\ & \quad = -\hat{Q} - (W_k - W_{k-1})\hat{P}(W_k - W_{k-1}) - W_k \hat{P} W_k. \end{aligned} \tag{4.12}$$

Die rechte Seite ist symmetrisch und negativ definit. Daraus folgt erneut mit Satz 3.9, dass $\hat{A} - \hat{P}W_k$ asymptotisch stabil ist.

ii) Wir subtrahieren nun (4.11) für $k + 1$ von (4.12),

$$\begin{aligned} & (\hat{A} - \hat{P}W_k)^T(W_k - W_{k+1}) + (W_k - W_{k+1})(\hat{A} - \hat{P}W_k) \\ & \quad = -(W_k - W_{k-1})\hat{P}(W_k - W_{k-1}) \end{aligned}$$

Die rechte Seite ist symmetrisch und negativ semidefinit und somit folgt mit Satz 3.10, dass $W_k - W_{k+1} \geq 0$ und damit $W_k \geq W_{k+1}$ für $k \geq 1$ gilt.

iii) Die Folge (W_k) ist monoton fallend und nach unten beschränkt. Also existiert der Grenzwert

$$W_\infty := \lim_{k \rightarrow \infty} W_k.$$

Für den Grenzwert gilt:

- $\Lambda(\hat{A} - \hat{P}W_\infty) \subseteq \mathbb{C}^- \cup i\mathbb{R}$,
- Bilden des Grenzprozesses zeigt, dass W_∞ (4.11) erfüllt, wenn man W_k und W_{k-1} durch W_∞ ersetzt. Somit erfüllt W_∞ die ARE.
- Das Bild von $[I, W_\infty]^T$ spannt einen invarianten Unterraum von $\mathcal{H} = \begin{bmatrix} \hat{A} & -\hat{P} \\ -\hat{Q} & -\hat{A}^T \end{bmatrix}$ zu den Eigenwerten in $\mathbb{C}^- \cup i\mathbb{R}$ auf. Da \mathcal{H} jedoch keine rein-imaginären Eigenwerte hat, muss $\Lambda(\hat{A} - \hat{P}W_\infty) \subseteq \mathbb{C}^-$ gelten. Also ist W_∞ eine stabilisierende Lösung der ARE. Aufgrund der Eindeutigkeit folgt damit $W_\infty = W_*$.

iv) Für die Fehlerabschätzung betrachten wir (4.12) im Grenzprozess $k \rightarrow \infty$. Es folgt

$$(\hat{A} - \hat{P}W_*)^T W_* + W_*(\hat{A} - \hat{P}W_*) = -\hat{Q} - W_*\hat{P}W_*,$$

also

$$\begin{aligned} & (\hat{A} - \hat{P}W_k)^T W_* + W_*(\hat{A} - \hat{P}W_k) \\ &= -\hat{Q} - W_*\hat{P}W_* + (W_* - W_k)^T \hat{P}W_* + W_*\hat{P}(W_* - W_k). \end{aligned}$$

Subtrahieren wir von dieser Gleichung (4.12), dann ergibt sich

$$\begin{aligned} & (\hat{A} - \hat{P}W_k)^T (W_* - W_k) + (W_* - W_k)(\hat{A} - \hat{P}W_k) \\ &= -W_*\hat{P}W_* + (W_* - W_k)\hat{P}W_* + W_*\hat{P}(W_* - W_k) \\ & \quad + W_k\hat{P}W_k + (W_k - W_{k-1})\hat{P}(W_k - W_{k-1}) \\ &= (W_* - W_k)\hat{P}(W_* - W_k) + (W_k - W_{k-1})\hat{P}(W_k - W_{k-1}). \end{aligned}$$

Wir subtrahieren $2(W_* - W_k)\hat{P}(W_* - W_k)$ und erhalten

$$\begin{aligned} & (\hat{A} - \hat{P}W_*)^T (W_* - W_k) + (W_* - W_k)(\hat{A} - \hat{P}W_*) \\ &= -(W_* - W_k)\hat{P}(W_* - W_k) + (W_k - W_{k-1})\hat{P}(W_k - W_{k-1}) =: G. \end{aligned}$$

Da $\hat{A} - \hat{P}W_*$ asymptotisch stabil ist, können wir die Lösungsformel aus Satz 3.7 anwenden. Dies ergibt

$$W_k - W_* = -(W_* - W_k) = \int_0^\infty e^{t(\hat{A} - \hat{P}W_*)^T} G e^{t(\hat{A} - \hat{P}W_*)} dt.$$

Dabei gilt $G \leq (W_k - W_{k-1})\hat{P}(W_k - W_{k-1})$, da $\hat{P} \geq 0$ gilt. Also folgt

$$\begin{aligned} \|W_k - W_*\|_2 &\leq \int_0^\infty \|e^{t(\hat{A} - \hat{P}W_*)^T} G e^{t(\hat{A} - \hat{P}W_*)}\|_2 dt \\ &\leq \|W_k - W_{k-1}\|_2^2 \underbrace{\|\hat{P}\|_2 \int_0^\infty \|e^{t(\hat{A} - \hat{P}W_*)^T} e^{t(\hat{A} - \hat{P}W_*)}\|_2 dt}_{=:\gamma} \\ &= \gamma \|W_k - W_{k-1}\|_2^2 \\ &\leq \gamma \|W_* - W_{k-1}\|_2^2. \end{aligned}$$

Im letzten Schritt haben wir genutzt, dass $0 \leq A \leq B$ die Ungleichung $\|A\|_2 \leq \|B\|_2$ impliziert (Übungsaufgabe (optional)) mit $A = W_{k-1} - W_k$ und $B = W_{k-1} - W_*$.

□

Bemerkung 4.25. Wie üblich liegt das Problem beim Newtonverfahren im Finden eines guten Startwertes. Hier müssen wir ein W_0 finden, sodass $\hat{A} - \hat{P}W_0$ asymptotisch stabil ist. Dies erfordert wieder das Lösen einer Riccati-Gleichung. Daher wird das Verfahren bevorzugt als Nachiteration verwendet.

Bemerkung 4.26. Wegen Punkt 1) im Satz 4.24 sind stets alle Lyapunov-Gleichungen eindeutig lösbar.

4.3 Signum Methode zur Lösung der ARE