ME 7247: Advanced Control Systems **Supplementary notes**

Lyapunov equations

v.2.0 (11.3.2022)

In these notes, we derive conditions under which the Lyapunov equation has a unique solution, and we explain the interplay between stability of A and positive definiteness of the solution.

1 The Sylvester equation

The (discrete) Sylvester equation is a matrix equation given by

$$A^{\mathsf{T}} X B - X + Q = 0, \tag{1}$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{m \times m}$ are square matrices, but $X, Q \in \mathbb{R}^{n \times m}$ need not be square. We are interested in the case where A, B, Q are given, and we must find X.

1.1 Existence and uniqueness of solutions

Lemma 1. *The Sylvester equation* (1) *has a unique solution* X *if and only if* $\lambda_A \lambda_B \neq 1$ *for every eigenvalue* λ_A *of* A *and eigenvalue* λ_B *of* B.

Proof. Eq. (1) is a set of mn linear equations in mn unknowns. Therefore, it has a unique solution if and only if the homogeneous equation $A^{\mathsf{T}} X B - X = 0$ admits only the trivial solution $X = 0$. This is the same as saying that our solution is unique for $Fx = g$ if $\text{null}(F) = \{0\}$. In our case, F is square (as many equations as unknowns), so a zero nullspace means range(A) is the whole space, so there is a solution for every g and this solution is unique.

Suppose $\lambda_A \lambda_B = 1$. Let $v \neq 0$ be a left eigenvector of A for λ_A and let $w \neq 0$ be a left eigenvector of B for λ_B. Now let $X = vw^* \neq 0$, and we have $A^{\mathsf{T}} X B = A^{\mathsf{T}} v w^* B = \lambda_A \lambda_B v w^* = v w^* = X$. Similarly, $A^{\mathsf{T}} \bar{X} B = \bar{\lambda}_A \bar{\lambda}_B \bar{X} = \bar{X}$. So both X and $\bar{X}$ satisfy the homogeneous equation. Consequently, so does $\text{Re}(X) = \frac{1}{2}(X + \bar{X})$ and $\text{Im}(X) = \frac{1}{2i}(X - \bar{X})$. These matrices can't both be zero (otherwise X would itself be zero), so at least one of them is a real nontrivial solution to the homogeneous equation $A^{\mathsf{T}} X B - X = 0$.

Conversely, suppose we have a solution $X \neq 0$ to the homogeneous equation $A^{\mathsf{T}} X B = X$. Let $B = PJP^{-1}$ be a Jordan decomposition of B. Rewrite the equation as $A^{\mathsf{T}} \hat{X} J = \hat{X}$ where $\hat{X} = XP \neq 0$. Pick out an eigenvalue λ_B of B. Suppose the corresponding Jordan block has size q and write $A^{\mathsf{T}} \hat{X}_\lambda J_\lambda = \hat{X}_\lambda$ with $J_\lambda \in \mathbb{C}^{q \times q}$. Since $\hat{X} \neq 0$, suppose λ_B was chosen such that $\hat{X}_\lambda \neq 0$. Let $\hat{x}_\ell$ be ℓ^{th} column of $\hat{X}_\lambda$. Writing $A^{\mathsf{T}} \hat{X}_\lambda J_\lambda = \hat{X}_\lambda$ columnwise, we obtain

$$\lambda_B A^{\mathsf{T}} \hat{x}_1 = \hat{x}_1, \qquad A^{\mathsf{T}} \hat{x}_1 + \lambda_B A^{\mathsf{T}} \hat{x}_2 = \hat{x}_2, \qquad \dots \qquad A^{\mathsf{T}} \hat{x}_{r-1} + \lambda_B A^{\mathsf{T}} \hat{x}_r = \hat{x}_r.$$

The first equation tells us that $(I - \lambda_B A^{\mathsf{T}})\hat{x}_1 = 0$. If $\hat{x}_1 \neq 0$, then λ_B^{-1} is an eigenvalue of A. Which means we have $\lambda_A \lambda_B = 1$. If this is not the case, then $\hat{x}_1 = 0$. Substituting this into the second equation, we have $(I - \lambda_B A^{\mathsf{T}})\hat{x}_2 = 0$. Repeating this argument, we conclude that $\lambda_A \lambda_B = 1$, for otherwise we would have $\hat{x}_\ell = 0$ for all ℓ, which contradicts the fact that $\hat{X}_\lambda \neq 0$. ■

2 The Lyapunov equation

The (discrete) Lyapunov equation is a special case of the Sylvester equation with $B = A$.

$$A^\mathsf{T} X A - X + Q = 0, \tag{2}$$

where A and Q are given matrices, and our goal is to solve for X. Here, all matrices are $n \times n$. Our main result describes the connections between Schur-stability of A, definiteness of solution to the Lyapunov equation, and properties of the matrices (A, Q).

Theorem 1. *Consider the Lyapunov equation* (2)*.*

1. *Suppose A is Schur-stable.*
   (a) *There exists a unique solution to the Lyapunov equation, and $X = \sum_{k=0}^{\infty} (A^\mathsf{T})^k Q A^k$.*
   (b) *If $Q \succeq 0$, then $X \succeq 0$.*
   (c) *If $Q \succeq 0$, then $X \succ 0$ if and only if (A, Q) is observable.*
2. *If X is a solution to the Lyapunov equation, then*
   (a) *If $Q \succeq 0$ and $X \succ 0$, then all eigenvalues of A satisfy $|\lambda| \le 1$.*
   (b) *If $Q \succeq 0$ and $X \succeq 0$ and (A, Q) is detectable, then A is Schur-stable.*

Proof. We prove each item separately. Also, we will make use of some technical lemmas regarding observability and detectability, which may be found in Appendix B.

1. Suppose A is Schur-stable. The Lyapunov equation is a Sylvester equation with $B = A$. Since A is Schur-stable, we have $|\lambda_A \lambda_B| = |\lambda_A| \cdot |\lambda_B| < 1$, so by Lemma 1, the Lyapunov equation has a unique solution. The proposed infinite sum converges.[1] We can also see by direct substitution that this X satisfies the Lyapunov equation, proving Item (a). Due to the special form of the infinite sum, X will inherit symmetry and definiteness properties from Q. So if $Q \succeq 0$, then $X \succeq 0$ and we have proven Item (b). To prove Item (c), multiply the Lyapunov equation by $v^*(\dots)v$, where (λ, v) is an eigenpair of A, and obtain

$$(|\lambda|^2 - 1) v^* X v + v^* Q v = 0. \tag{3}$$

   Since A is Schur-stable, $|\lambda| < 1$. If $X \succ 0$, the first term is negative, which means the second term must be positive. Since $Q \succeq 0$, we deduce that $Qv \neq 0$. By Lemma 7, (A, Q) is observable. Suppose instead that $X \not\succ 0$. Since $X \succeq 0$ from Item (a), there must exist some $z \neq 0$ such that $Xz = 0$. Using the formula from Item (a), we have

$$0 = z^\mathsf{T} X z = \sum_{k=0}^{\infty} (A^k z)^\mathsf{T} Q (A^k z) = \sum_{k=0}^{\infty} \left\| Q^{1/2} A^k z \right\|^2.$$

   Therefore $Q^{1/2} A^k z = 0$ for all k, so $Q A^k z = 0$ for all k. This implies that z is in the nullspace of the observability matrix (A, Q), so (A, Q) is not observable.

[1] See Lemma 6 in Appendix A for a proof of this fact.

2. Suppose X is a solution to the Lyapunov equation. Let (λ, v) be an eigenpair of A, and obtain (3) again. If $Q \succeq 0$, the second term is ≥ 0 so the first term must be ≤ 0. If $X \succ 0$, we deduce $(|\lambda|^2 - 1) \leq 0$, so $|\lambda| \leq 1$ and we have proven Item (a). If $X \succeq 0$ and (A, Q) is detectable, then by Lemma 8, whenever $|\lambda| \geq 1$, we have $Qv \neq 0$, so $v^* Q v > 0$. But the first term is ≥ 0, a contradiction since the two terms must sum to zero. So we conclude that there can be no eigenvalues of A satisfying $|\lambda| \geq 1$, so A is Schur-stable and we have proven Item (b). ∎

2.1 Connection to Gramians

The Lyapunov equations for the observability and controllability Gramians are

$$A^\mathsf{T} Q A - Q + C^\mathsf{T} C = 0 \quad \text{and} \quad A P A^\mathsf{T} - P + B B^\mathsf{T} = 0.$$

If A is Schur-stable, we can apply Theorem 1 and Lemma 7 to conclude that:

(i) $Q \succ 0 \iff (A, C^\mathsf{T} C)$ observable $\iff (A, C)$ observable.

(ii) $P \succ 0 \iff (A^\mathsf{T}, B B^\mathsf{T})$ observable $\iff (A^\mathsf{T}, B^\mathsf{T})$ observable $\iff (A, B)$ controllable.

2.2 Monotonicity results

In certain instances, it can be useful to replace the Lyapunov equation by a corresponding inequality. Let's investigate when this is possible and what other properties follow.

Lemma 2. *The following statements are equivalent.*

(i) The matrix A is Schur-stable.

(ii) There exists a matrix $X \succ 0$ such that $A^\mathsf{T} X A - X \prec 0$.

Proof. Suppose A is Schur-stable. Let $Q = I$, so (A, Q) is observable. By Theorem 1, Eq. (2) has a unique solution and $X \succ 0$. Moreover, we have $A^\mathsf{T} X A - X = -I \prec 0$. Conversely, suppose $X \succ 0$ and $A^\mathsf{T} X A - X \prec 0$. Then define $Q := -(A^\mathsf{T} X A - X) \succ 0$. Now (2) is satisfied and (A, Q) is detectable since Q is invertible. By Theorem 1, we conclude that A is Schur-stable. ∎

Lemma 3. *Suppose A is Schur-stable. If X_i and Q_i satisfy*

$$A^\mathsf{T} X_1 A - X_1 + Q_1 = 0 \qquad \textit{and} \qquad A^\mathsf{T} X_2 A - X_2 + Q_2 = 0,$$

then if $Q_1 \succeq Q_2$, we have $X_1 \succeq X_2$.

Proof. Subtracting one equation from the other, obtain $A^\mathsf{T}(X_1 - X_2)A - (X_1 - X_2) + (Q_1 - Q_2) = 0$. From Theorem 1, if $Q_1 - Q_2 \succeq 0$, then $X_1 - X_2 \succeq 0$. ∎

Note. The converse of Lemma 3 is *not* true in general, so $X_1 \succeq X_2 \not\Longrightarrow Q_1 \succeq Q_2$. In fact, if we arbitrarily pick some $X \succ 0$ and Schur-stable A, then $A^\mathsf{T} X A - X$ may be indefinite.

Lemma 4. *Suppose A is Schur-stable. Let X_0 be the unique solution to the Lyapunov equation $A^\mathsf{T} X_0 A - X_0 + Q = 0$. Then we have:*

- *If X satisfies $A^\mathsf{T} X A - X + Q \prec 0$, then $X_0 \prec X$. In other words, X_0 is the* minimal solution *among all solutions of this Lyapunov inequality.*
- *If X satisfies $A^\mathsf{T} X A - X + Q \succ 0$, then $X_0 \succ X$. In other words, X_0 is the* maximal solution *among all solutions of this Lyapunov inequality.*

Proof. Subtracting the Lyapunov equation from the inequality, obtain $A^\mathsf{T}(X - X_0)A - (X - X_0) \prec 0$. Multiplying the above by $A^\mathsf{T}(\dots)A$ and iterating, we conclude that

$$(X - X_0) \succ A^\mathsf{T}(X - X_0)A \succ (A^\mathsf{T})^2(X - X_0)A^2 \succ \dots \succ (A^\mathsf{T})^k(X - X_0)A^k.$$

Since A is Schur-stable, $A^k \to 0$ as $k \to \infty$, so we conclude that $X - X_0 \succ 0$. The second claim of Lemma 4 can be proved in an analogous manner. ∎

Note. If we apply Lemma 4 to a case where $Q \succeq 0$, then $X_0 \succeq 0$, therefore all solutions to the inequality $A^\mathsf{T} X A - X + Q \prec 0$ also satisfy $X \succ X_0 \succeq 0$ automatically. The same is not true if we reverse the inequality. If we have a solution to $A^\mathsf{T} X A - X + Q \succ 0$, then all we can say is that $X_0 \succ X$ and $X_0 \succeq 0$, so X need not be positive definite.

Finally, we have the following result that relates detectability to a Lyapunov-like inequality.

Lemma 5. *The following are equivalent.*

(i) (A, C) is detectable.

(ii) There exists $Y \succ 0$ such that $A^\mathsf{T} Y A - Y - C^\mathsf{T} C \prec 0$.

This matrix inequality in Lemma 5 is similar to the observability Gramian, but notice that A need not be stable, and there is a negative sign in front of the $C^\mathsf{T} C$ term.

Proof. Suppose $Y \succ 0$ satisfies $A^\mathsf{T} Y A - Y - C^\mathsf{T} C \prec 0$. Suppose (A, C) is *not* detectable. By Lemma 8, there exists (λ, v) such that $v \neq 0$, $Av = \lambda v$, $|\lambda| \geq 1$, and $Cv = 0$. Multiply the inequality by $v^*(\dots)v$ and obtain $(|\lambda|^2 - 1)v^* Y v < 0$. But $Y \succ 0$ and $|\lambda| \geq 1$, a contradiction. So we conclude (A, C) is detectable.

Conversely, suppose (A, C) is detectable. Then there exists a matrix L such that $A + LC$ is Schur-stable. By Theorem 1, the Lyapunov equation $(A + LC)X(A + LC)^\mathsf{T} - X + (I + LL^\mathsf{T}) = 0$ has a solution $X \succ 0$. Therefore, $(A + LC)X(A + LC)^\mathsf{T} - X + LL^\mathsf{T} \prec 0$. Using properties of Schur complements, this is equivalent to

$$0 \prec \begin{bmatrix} X - LL^\mathsf{T} & A + LC \\ (A + LC)^\mathsf{T} & X^{-1} \end{bmatrix} = \begin{bmatrix} X & A \\ A^\mathsf{T} & X^{-1} + C^\mathsf{T} C \end{bmatrix} - \begin{bmatrix} -L \\ C^\mathsf{T} \end{bmatrix} \begin{bmatrix} -L \\ C^\mathsf{T} \end{bmatrix}^\mathsf{T} \preceq \begin{bmatrix} X & A \\ A^\mathsf{T} & X^{-1} + C^\mathsf{T} C \end{bmatrix}$$

Applying Schur complements again, this is equivalent to $X^{-1} + C^\mathsf{T} C - A^\mathsf{T} X^{-1} A \succ 0$ and $X \succ 0$. Letting $Y = X^{-1}$, and rearranging, we obtain $Y \succ 0$ and $A^\mathsf{T} Y A - Y - C^\mathsf{T} C \prec 0$, as required. ∎

Note. An analogous result to Lemma 5 holds for stabilizability. Namely, (A, B) is stabilizable if and only if there exists $X \succ 0$ such that $AXA^\mathsf{T} - X - BB^\mathsf{T} \prec 0$.

A Convergence of an infinite matrix sum

Lemma 6. *Suppose A is Schur-stable. The following infinite sum converges.*

$$\sum_{k=0}^{\infty} A^k Q (A^\mathsf{T})^k$$

Proof. We divide the proof into several steps.

Step 1. First, we show that if A is Schur-stable, then $\lim_{k\to\infty} A^k = 0$. To see why this is so, write A is Jordan normal form: $A = PJP^{-1}$, and use the fact that $A^k = PJ^kP^{-1}$. We will prove that $J^k \to 0$, which implies that $A^k \to 0$. The matrix J is block diagonal and made up of the Jordan blocks J_λ corresponding to the eigenvalues of A. Each Jordan block looks like

$$J_\lambda = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \lambda & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} = \lambda I + S,$$

where S is the shift matrix (1's on the super-diagonal and zeros everywhere else). Since λI and S commute, we can apply the binomial theorem to expand powers of J_λ. Powers of S correspond to additional shifts, so if $S \in \mathbb{R}^{m\times m}$, we have $S^m = 0$. So when $k \geq m-1$, we have

$$(\lambda I + S)^k = \sum_{\ell=0}^{k} \binom{k}{\ell} S^\ell \lambda^{k-\ell} = \sum_{\ell=0}^{m-1} \binom{k}{\ell} S^\ell \lambda^{k-\ell} = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \cdots & \binom{k}{m-1}\lambda^{k-m+1} \\ 0 & \lambda^k & k\lambda^{k-1} & \ddots & \vdots \\ 0 & 0 & \lambda^k & \ddots & \binom{k}{2}\lambda^{k-2} \\ \vdots & \vdots & \ddots & \ddots & k\lambda^{k-1} \\ 0 & 0 & \cdots & 0 & \lambda^k \end{bmatrix}$$

As $k \to \infty$, the exponential terms involving powers of λ dominate, since the binomial coefficients are polynomials in degree at most $m-1$. Since A is Schur-stable, $|\lambda| < 1$, so we have $J_\lambda^k \to 0$, and therefore $J^k \to 0$ and $A^k \to 0$.

Step 2. Next, we show that when k is sufficiently large, $\|A^k\|$ is bounded by a decaying exponential in k. We already know from Step 1 that $\lim_{k\to\infty} = 0$, so $\lim_{k\to\infty} \|A^k\| = 0$. Note that the limit being zero does not mean that $\|A^k\|$ decays monotonically to zero. It may increase at first, and it may oscillate as it decays.

Let $\rho(A)$ be the spectral radius of A (largest eigenvalue magnitude). Schur-stability of A implies that $\rho(A) < 1$. Pick $\varepsilon \in \big(0, 1-\rho(A)\big)$. Then,

$$\rho\left(\frac{1}{1-\varepsilon}A\right) = \frac{\rho(A)}{1-\varepsilon} < 1.$$

Therefore, $\frac{1}{1-\varepsilon}A$ is Schur-stable, and $\lim_{k\to\infty}\left(\frac{1}{1-\varepsilon}A\right)^k = 0$. By the definition of the limit, there exists k_0 such that for all $k \geq k_0$, we have $\left\|\left(\frac{1}{1-\varepsilon}A\right)^k\right\| < 1$, which rearranges to $\|A^k\| < (1-\varepsilon)^k$.

Step 3. To show that our infinite sum is convergent, it suffices to show that it is *absolutely convergent*. In other words, we will prove that the series

$$\sum_{\ell=0}^{\infty} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\|$$

is convergent. Define k_0 as in Step 2, pick $k \geq k_0$, and apply the triangle inequality and submultiplicativity of the matrix norm to obtain

$$\begin{aligned}
\sum_{\ell=0}^{k} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| &= \sum_{\ell=0}^{k_0-1} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| + \sum_{\ell=k_0}^{k} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| \\
&\leq \sum_{\ell=0}^{k_0-1} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| + \sum_{\ell=k_0}^{k} \left\| (A^{\mathsf{T}})^{\ell} \right\| \|Q\| \left\| A^{\ell} \right\| \\
&\leq \sum_{\ell=0}^{k_0-1} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| + \sum_{\ell=0}^{k} (1-\varepsilon)^{2\ell} \|Q\| \\
&\leq \sum_{\ell=0}^{k_0-1} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| + \sum_{\ell=0}^{\infty} (1-\varepsilon)^{2\ell} \|Q\| \\
&\leq \sum_{\ell=0}^{k_0-1} \left\| A^{\ell} Q (A^{\mathsf{T}})^{\ell} \right\| + \frac{\|Q\|}{1-(1-\varepsilon)^2}
\end{aligned}$$

The right-hand side is independent of k, which shows that the left-hand side is uniformly bounded for all k. Since the left-hand side is an increasing function of k, it must converge as $k \to \infty$. This shows that our original series is absolutely convergent, and hence convergent. ■

B Observability and detectability

These are some technical lemmas we used in the proofs for Theorem 1.

Lemma 7 (observability)**.** *Let $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{m\times n}$ be given matrices. The following statements are equivalent.*

(i) The pair (A, C) is observable.

(ii) The pair $(A, C^\mathsf{T} C)$ is observable.

(iii) The eigenvalues of $A + LC$ may be freely assigned by suitable choice of L.

(iv) The observability matrix $\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$ has full column rank.

(v) For all $\lambda \in \mathbb{C}$, the matrix $\begin{bmatrix} C \\ A - \lambda I \end{bmatrix}$ has full column rank.

(vi) If $\lambda \in \mathbb{C}$ and $0 \neq v \in \mathbb{C}^n$ satisfy $Av = \lambda v$, then $Cv \neq 0$.

Lemma 8 (detectability)**.** *Let $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{m\times n}$ be given matrices. The following statements are equivalent.*

(i) The pair (A, C) is detectable.

(ii) The pair $(A, C^\mathsf{T} C)$ is detectable.

(iii) There exists a matrix L such that $A + LC$ is Schur-stable.

(iv) For all $\lambda \in \mathbb{C}$ with $|\lambda| \geq 1$, the matrix $\begin{bmatrix} C \\ A - \lambda I \end{bmatrix}$ has full column rank.

(v) If $\lambda \in \mathbb{C}$ and $0 \neq v \in \mathbb{C}^n$ satisfy $|\lambda| \geq 1$ and $Av = \lambda v$, then $Cv \neq 0$.

Items (v) and (vi) of Lemma 7 and Items (iv) and (v) of Lemma 8 are commonly known as the *Popov–Belevitch–Hautus* (PBH) test. We omit the proofs of Lemmas 7 and 8 as they are standard results and can be found in any linear systems textbook.