Solutions Exercises Set 7

Exercise 1 (Matrix norms)

Norms on signals and systems are very important in system theory, in particular in system approximation. The purpose of this exercise is to introduce some basic facts on matrix norms. These yield a good introduction to system norms, since linear systems have very many things in common with infinite matrices.

Recall that $\mathbb{R}^{\bullet\times\bullet}$ denotes the real matrices with any, but, of course, a finite number of rows and columns. There are 3 basic operations defined on $\mathbb{R}^{\bullet\times\bullet}$:

$$\begin{array}{rcl}
\text{scalar multiplication } \circ: & \alpha \in \mathbb{R}, M \in \mathbb{R}^{\bullet\times\bullet} & \mapsto \alpha\circ M \in \mathbb{R}^{\bullet\times\bullet}, \\
\text{addition } +: & M_1, M_2 \in \mathbb{R}^{\bullet\times\bullet}, \text{ with } \dim(M_1) = \dim(M_2) & \mapsto M_1 + M_2 \in \mathbb{R}^{\bullet\times\bullet}, \\
\text{multiplication } *: & M_1, M_2 \in \mathbb{R}^{\bullet\times\bullet}, \text{ with } \mathrm{coldim}(M_1) = \mathrm{rowdim}(M_2) & \mapsto M_1 * M_2 \in \mathbb{R}^{\bullet\times\bullet}.
\end{array}$$

In practice, neither the $\circ$ nor the $*$ are written.

The main role, and an exceedingly important one at that, on the mathematical stage for matrices is as linear transformations on $\mathbb{R}^{\bullet}$:

$$M \in \mathbb{R}^{\bullet\times\bullet}, v \in \mathbb{R}^{\bullet}, \text{ with } \dim(v) = \mathrm{rowdim}(M) \mapsto Mv \in \mathbb{R}^{\bullet}.$$

In the sequel we will not mention explicitly the dimension restrictions that must hold for addition, multiplication, and matrix action on vectors.

There are 3 types of norms on $\mathbb{R}^{\bullet\times\bullet}$:

Vector norms, $||\cdot|| : \mathbb{R}^{\bullet\times\bullet} \to \mathbb{R}_+$, in which only the vector space structure is considered. Vector norms satisfy the axioms

1. $||M|| > 0$ for $M \neq 0$,
2. $||\alpha \circ M|| = |\alpha| \; ||M||$,
3. $||M_1 + M_2|| \leq ||M_1|| + ||M_2||$ (the *triangle inequality*, also called *sub-additivity*).

Matrix norms, (written here with 3 bars - this is not standard) $|||\cdot||| : \mathbb{R}^{\bullet\times\bullet} \to \mathbb{R}_+$, in which also the algebra (multiplication) structure is considered. Such norms satisfy, in addition to 1, 2, and 3, also the axiom

4. $|||M_1 * M_2||| \leq |||M_1||| \;\; |||M_2|||$ (called *sub-multiplicativity*).

Induced norms, in which the role of matrices as linear transformations comes to the foreground. Induced norms are defined by

$$|||M|||_{1\to 2} = \sup_{0\neq v\in\mathbb{R}^{\bullet}} \frac{||Mv||_2}{||v||_1}.$$

where $||\cdot||_1$ and $||\cdot||_2$ denote vector norms on $\mathbb{R}^{\bullet}$. We use the notation $|||\cdot|||_{1\to 1}$ if the matrix is square, and the norm on the domain and co-domain are taken to be the same. Such a norm is sometimes called *equi-induced.*

1. Prove that an induced norm is a matrix norm (whence the $|||\cdot|||_{1\to 2}$-notation, with 3 bars, is not frivolous), *i.e.*, prove that if you choose once and for all a norm on $\mathbb{R}^{\mathtt{n}}$ for all $\mathtt{n} \in \mathbb{N}$, then the resulting induced norm on matrices is a matrix norm.

$$\begin{aligned}
|||M_1 M_2|||_{1\to 3} &= \sup_{x\neq 0} \frac{\|M_1 M_2 x\|_3}{\|x\|_3} \\
&= \sup_{\substack{x\neq 0 \\ M_2 x \neq 0}} \frac{\|M_1 M_2 x\|_3}{\|M_2 x\|_2} \frac{\|M_2 x\|_2}{\|x\|_3} \\
&\leq \sup_{M_2 x\neq 0} \frac{\|M_1 M_2 x\|_3}{\|M_2 x\|_2} \sup_{x\neq 0} \frac{\|M_2 x\|_2}{\|x\|_3} \\
&\leq \sup_{0\neq v} \frac{\|M_1 v\|_3}{\|v\|_2} \sup_{x\neq 0} \frac{\|M_2 x\|_2}{\|x\|_3} \\
&= |||M_1|||_{2\to 3}\ |||M_2|||_{1\to 2}
\end{aligned}$$

2. Prove that $|||I||| \geq 1$.

$$|||I||| = |||I\ I||| \leq |||I|||\ |||I||| \qquad \Rightarrow \qquad |||I||| \geq 1$$

3. Prove that $|||I|||_{1\to 1} = 1$.

$$|||I|||_{1\to 1} = \sup_{x\neq 0} \frac{\|Ix\|_1}{\|x\|_1} = 1$$

4. Prove that if $||\cdot||$ is a vector norm on $\mathbb{R}^{\bullet\times\bullet}$, then so is the scaling $C||\cdot||$ for $C > 0$.

$C\|\cdot\|$ trivially satisfies axioms 1,2, and 3 for any $C > 0$.

Prove that if $|||\cdot|||$ is a matrix norm, then so is the scaling $C|||\cdot|||$ for $C \geq 1$.

If $C \geq 1$, then axiom 4 holds as well since

$$|||M_1 M_2||| \leq |||M_1|||\ |||M_2||| \quad \textbf{and} \quad C \geq 1 \qquad \Rightarrow \qquad C|||M_1 M_2||| \leq C|||M_1|||\ C|||M_2|||$$

For which C is the scaled induced norm $C|||\cdot|||_{1\to 2}$ an induced norm?

Obviously $C|||\cdot|||_{1\to 2}$ is also induced: change $||\cdot||_1$ to $C^{-1}||\cdot||_1$. But, if $|||\cdot|||_{1\to 1}$ is an equi-induced norm, then $|||I|||_{1\to 1} = 1$ (see 3). Hence for $C|||\cdot|||_{1\to 1}$ to be an equi-induced norm, C must be 1.

Same question for $C|||\cdot|||_{1\to 1}$.

The most 'famous' matrix norms are the *Frobenius norm*:

$$||M||_{\mathfrak{F}} := \sqrt{\mathrm{Trace}(M^\top M)}.$$

and the *spectral norm*

$$||M||_{\text{spectral}} := \text{ the square root of the maximum eigenvalue of } M^\top M.$$

5. Prove that the Frobenius norm is a matrix norm, but not, and none of its scalings, an induced norm. More precisely, prove that, unless $\mathtt{n}_1$ or $\mathtt{n}_2$ is equal to one, there do not exist norms $||\cdot||_1$ on $\mathbb{R}^{\mathtt{n}_1}$ and $||\cdot||_2||$ on $\mathbb{R}^{\mathtt{n}_2}$ such that for all $M \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$, $||M||_{\mathfrak{F}} = \sup_{0\neq v\in\mathbb{R}^{\mathtt{n}_2}} \frac{||Mv||_1}{||v||_2}$.

Let $M_{\mathtt{k}\ell}$ denotes the $(\mathtt{k},\ell)$-th element of $M \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$. Then

$$\|M\|_{\mathfrak{F}} = \sqrt{\sum_{\mathtt{k}=1}^{\mathtt{n}_1}\sum_{\ell=1}^{\mathtt{n}_2} M_{\mathtt{k}\ell}^2}.$$

This shows that $\|\cdot\|_{\mathfrak{F}}$ is merely the Euclidean norm of M viewed in the obvious way as an element of $\mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$. This yields axioms 1, 2, and 3.

To show axiom 4, let $M' \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ and $M'' \in \mathbb{R}^{\mathtt{n}_2\times\mathtt{n}_3}$. Then

$$\begin{aligned}
\|M'M''\|_{\mathfrak{F}}^2 &= \sum_{\mathtt{k}=1}^{\mathtt{n}_1}\sum_{\ell=1}^{\mathtt{n}_3}(M'M'')_{\mathtt{k}\ell}^2 \\
&= \sum_{\mathtt{k}=1}^{\mathtt{n}_1}\sum_{\ell=1}^{\mathtt{n}_3}\left(\sum_{\mathtt{j}=1}^{\mathtt{n}_2} M'_{\mathtt{kj}}M''_{\mathtt{j}\ell}\right)^2 \\
&\leq \sum_{\mathtt{k}=1}^{\mathtt{n}_1}\sum_{\ell=1}^{\mathtt{n}_3}\sum_{\mathtt{j}=1}^{\mathtt{n}_2}(M'_{\mathtt{kj}})^2\sum_{\mathtt{j}=1}^{\mathtt{n}_2}(M''_{\mathtt{j}\ell})^2 \qquad \textbf{(Cauchy–Schwartz)} \\
&= \sum_{\mathtt{k}=1}^{\mathtt{n}_1}\sum_{\ell=1}^{\mathtt{n}_2}(M'_{\mathtt{kj}})^2 \;\sum_{\ell=1}^{\mathtt{n}_3}\sum_{\mathtt{j}=1}^{\mathtt{n}_2}(M''_{\mathtt{j}\ell})^2 \\
&= \|M'\|_{\mathfrak{F}}^2\;\|M''|_{\mathfrak{F}}^2
\end{aligned}$$

We now show that no matter which norms $\|\cdot\|_1$ on $\mathbb{R}^{\mathtt{n}_1}$ and $\|\cdot\|_2$ on $\mathbb{R}^{\mathtt{n}_2}$, we choose, there is no constant $C > 0$, such that

$$\frac{\|M|_{2\to1}}{\|M\|_{\mathfrak{F}}} = C, \quad \textbf{for all } M \in \mathbb{R}^{\mathtt{n}_1\times n_2}.$$

This result is quite easy to prove for equi-induced norms, but we want the general case. The proof which we give, is courtesy of Vijaya–Sekhar Chellaboina and Wassim M. Haddad, "Is the Frobenius norm induced?", *IEEE Transactions on Automatic Control*, volume 40, pages 2137–2139, 1995. It is quite surprising that such a basic fact about matrices was apparently first proven in a control journal.

In fact, we first prove that the only unitarily invariant (see 7) induced norms on $\mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ are those which are proportional to the spectral norm.

<u>Proposition</u>: Let $J \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ be the matrix with a 1 in the $(1,1)$ element and 0 everywhere else. Then $\|\cdot\|$ is a unitarily invariant induced norm on $\mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ iff

$$\|M\| = \|J\|\;\|M\|_{\textbf{spectral}}.$$

<u>Proof</u>: Let M be a dyad, *i.e.*, a matrix of rank one: $M = ab^\top$ with $a \in \mathbb{R}^{\mathtt{n}_1}$ and $b \in \mathbb{R}^{\mathtt{n}_2}$. Then by the SVD there exist unitary matrices U and V, such that

$$Uab^\top V^\top = \|a\|_{\mathbb{E}_{\mathtt{n}_1}}\|b\|_{\mathbb{E}_{\mathtt{n}_2}}J,$$

where $\|\cdot\|_{\mathbb{E}_\mathtt{n}}$ denotes the Euclidean norm on $\mathbb{R}^\mathtt{n}$. By unitary invariance, therefore,

$$\|ab^\top\| = \|a\|_{\mathbb{E}_{\mathtt{n}_1}}\|b\|_{\mathbb{E}_{\mathtt{n}_2}}\|J\|.$$

This proves the proposition for dyads.

Let $\|\cdot\|$ be a norm on $\mathbb{R}^{\mathtt{n}}$. The *dual norm* $\|\cdot\|^*$ is defined by

$$\|x\|^* := \max_{0\neq y\in\mathbb{R}^{\mathtt{n}}} \frac{|y^\top x|}{\|y\|}.$$

It is easy to prove that $\|\cdot\|^*$ also defines a norm on $\mathbb{R}^{\mathtt{n}}$. Further the dual of the dual is the original (the chic way of saying this is that *dualization is an involution on the space of norms on* $\mathbb{R}^{\mathtt{n}}$). In other words, $(\|\cdot\|^*)^* = \|\cdot\|$ (for a proof of this, see the book by Roger A. Horn and Charles R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985, theorem 5.5.14). Also, it is easy to see that the Euclidean norm is self-dual. Now compute the norm of $ab^\top$ as an induced norm from $\mathbb{R}^{\mathtt{n}_2}$, $\|\cdot\|_2$ to $\mathbb{R}^{\mathtt{n}_1}$, $\|\cdot\|_1$. Obviously

$$\|ab^\top\|_{2\to1} = \sup_{x\in\mathbb{R}^{\mathtt{n}_2}} \frac{|b^\top x|}{\|x\|_2}\|a\|_1 = \|a\|_1\|b\|_2^*.$$

It follows that if $\|\cdot\|$ is both induced and unitarily invariant, we obtain the following relation between the various norms:

$$\|a\|_1\|b\|_2^* = \|a\|_{\mathbb{E}_{\mathtt{n}_1}}\|b\|_{\mathbb{E}_{\mathtt{n}_2}}\|J\|.$$

Therefore (take a fixed b), there exists a $\rho > 0$, such that

$$\|a\|_1 = \rho\|a\|_{\mathbb{E}_{\mathtt{n}_1}}, \quad \text{for all } a\in\mathbb{R}^{\mathtt{n}_1}, \qquad \text{and} \qquad \|b\|_2^* = \frac{\|J\|}{\rho}\|b\|_{\mathbb{E}_{\mathtt{n}_2}}, \quad \text{for all } b\in\mathbb{R}^{\mathtt{n}_2}.$$

Dualizing the latter and using $(\|\cdot\|_2^*)^* = \|\cdot\|_2$, together with $\|\cdot\|_{\mathbb{E}_{\mathtt{n}}} = \|\cdot\|_{\mathbb{E}_{\mathtt{n}}}^*$, yields

$$\|\cdot\|_1 = \rho\|\cdot\|_{\mathbb{E}_{\mathtt{n}_1}}, \qquad \|\cdot\|_2 = \frac{\rho}{\|J\|}\|\cdot\|_{\mathbb{E}_{\mathtt{n}_2}}.$$

Conclude that

$$\begin{aligned}\|M\|_{2\to1} &= \sup_{0\neq x\in\mathbb{R}^{\mathtt{n}_2}} \frac{\|Mx\|_2}{\|x\|_1}\\ &= \|J\|\ \sup_{0\neq x\in\mathbb{R}^{\mathtt{n}_2}} \frac{\|Mx\|_{\mathbb{E}_{\mathtt{n}_1}}}{\|x\|_{\mathbb{E}_{\mathtt{n}_2}}}\\ &= \|J\|\ \|M\|_{\textbf{spectral}}. && \text{(see 6)}\end{aligned}$$

Since (see 8)

$$\|M\|_{\mathfrak{F}} = \sqrt{\sigma_1^2(M) + \sigma_2^2(M) + \cdots + \sigma_{\mathrm{rank}(M)}^2(M)}$$

and $\|M\|_{\textbf{spectral}} = \sigma_1(M)$, we see that $c\|M\|_{\mathfrak{F}}$ is not an induced norm on $\mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ (unless $\mathtt{n}_1$ and/or $\mathtt{n}_2$ are 1).

6. Prove that the spectral norm is an induced norm.

The induced norm of $M\in\mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ viewed as a map from $\mathbb{E}_{\mathtt{n}_1}$ to $\mathbb{E}_{\mathtt{n}_2}$ is defined by

$$\|M\|_{\mathbb{E}_{\mathtt{n}_1}\to\mathbb{E}_{\mathtt{n}_2}} = \sqrt{\sup_{0\neq x\in\mathbb{R}^{\mathtt{n}_2}} \frac{x^\top M^\top Mx}{x^\top x}}.$$

Since $M^\top M \geq 0$, this supremum equals the square root of the maximum eigenvalue of $M^\top M$, *i.e.*, $\|M\|_{\mathbb{E}_{\mathtt{n}_1}\to\mathbb{E}_{\mathtt{n}_2}} = \|M\|_{\textbf{spectral}}$.

7. Call a vector norm *unitarily invariant* if $||UMV|| = ||M||$ for all unitary matrices U, V (U, V square and $U^\top U = I, V^\top V = I$). Prove that the Frobenius and the spectral norm are both unitarily invariant.

(i)

$$\begin{aligned}\|M\|_{\mathfrak{F}}^2 &= \operatorname{trace}\big(UMV(UMV)^\top\big) \\ &= \operatorname{trace}(UMVV^\top M^\top U^\top) \\ &= \operatorname{trace}(UMM^\top U^\top) \\ &= \operatorname{trace}(U^\top UMM^\top) && (\operatorname{trace}(AB) = \operatorname{trace}(BA)) \\ &= \operatorname{trace}(MM^\top).\end{aligned}$$

This shows that the Frobenius norm is unitarily invariant.

(ii) $UMV(UMV)^\top = UMM^\top U^\top$ **and** $U^\top = U^{-1}$, **shows that** $UMV(UMV)^\top$ **and** $MM^\top$ **are similar, and hence have the same eigenvalues. Therefore** M **and** UMV **have the same spectral norm.**

8. Denote by $(\sigma_1(M), \sigma_2(M), \ldots, \sigma_{\min\{\mathtt{n}_1,\mathtt{n}_2\}}(M))$ the SV's of $M \in \mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$. Prove that a unitarily invariant norm on $\mathbb{R}^{\mathtt{n}_1\times\mathtt{n}_2}$ is a convex function of the SV's. Express the Frobenius norm and the spectral norm in terms of the SV's.

By the SVD, there exist unitary matrices U **and** V, **such that**

$$UMV = \begin{cases} \big[\operatorname{diag}(\sigma_1(M), \sigma_2(M), \ldots, \sigma_{\mathtt{n}_1}(M)) \;\; 0_{\mathtt{n}_2-\mathtt{n}_1}\big] & \textbf{if } \mathtt{n}_1 < \mathtt{n}_2 \\ \begin{bmatrix} \operatorname{diag}\big(\sigma_1(M), \sigma_2(M), \ldots, \sigma_{\mathtt{n}_2}(M)\big) \\ 0_{\mathtt{n}_1-\mathtt{n}_2} \end{bmatrix} & \textbf{if } \mathtt{n}_1 \geq \mathtt{n}_2 \end{cases} \tag{SVD}$$

Hence, if $\|\cdot\|$ **is a unitarity invariant norm,** $\|M\| = \|UMV\|$ **is obviously a function of the entries of** UMV, **whence of the singular values of** M.

Call this map $f : \mathbb{R}_+^{\min\{\mathtt{n}_1,\mathtt{n}_2\}} \to \mathbb{R}_+$.

$$f \textbf{ convex} :\Leftrightarrow f\big(\alpha x_1 + (1-\alpha)x_2\big) \leq \alpha f(x_1) + (1-\alpha) f(x_2), \textbf{ for all } x_1, x_2 \in \mathbb{R}_+^{\min\{\mathtt{n}_1,\mathtt{n}_2\}}, 0 \leq \alpha \leq 1.$$

This, however, is an immediate consequence of the convexity of norms

$$\|\alpha M_1 + (1-\alpha)M_2\| \leq \alpha\|M_1\| + (1-\alpha)\|M_2\|$$

applied to diagonal matrices.

To compute $\|M\|_{\textbf{spectral}}$, **observe that the non-zero eigenvalues of** $M^\top M$ **equal those of**

$$\operatorname{diag}\big(\sigma_1^2(M), \sigma_2^2(M), \ldots, \sigma_{\min\{\mathtt{n}_1,\mathtt{n}_2\}}^2(M)\big).$$

It follows that $\|M\|_{\textbf{spectral}} = \sigma_1(M)$. **To compute** $\|M\|_{\mathfrak{F}}$, **observe that**

$$\operatorname{trace}(M^\top M) = \operatorname{trace}\big((UMV)^\top(UMV)\big) = \sigma_1^2(M) + \sigma_2^2(M) + \cdots + \sigma_{\min\{\mathtt{n}_1,\mathtt{n}_2\}}^2(M).$$

Hence

$$\|M\|_{\mathfrak{F}} = \sqrt{\sigma_1^2(M) + \sigma_2^2(M) + \cdots + \sigma_{\min\{\mathtt{n}_1,\mathtt{n}_2\}}^2(M)}.$$

There appear to be more inequalities involving norms, singular values, and eigenvalues than there are fishes in the sea. Almost weird, for example, are the *Ky-Fan* norms. Let

$$\sigma(M) = (\sigma_1(M), \sigma_2(M), \ldots, \sigma_{\mathrm{rank}(M)}(M))$$

denote the non-zero SV's of M. Define for $\mathtt{k} \in \mathbb{N}$ and $1 \leq p < \infty$, the $(\mathtt{k}, p)$ Ky-Fan norm by

$$||M||_{(\mathtt{k},p)} := \left(\sum_{\mathtt{k}'=1,2,\ldots,\mathtt{k}} \sigma_{\mathtt{k}'}(M)^p \right)^{\frac{1}{p}}$$

(define $\sigma_{\mathtt{k}}(M) = 0$ for $\mathtt{k} > \mathrm{rank}(M)$). Well, believe it or not, $|| \cdot ||_{(\mathtt{k},p)}$ is a matrix norm for all $\mathtt{k} \in \mathbb{N}$ and $1 \leq p < \infty$. For $\mathtt{k} = 1$ the Ky-Fan norm becomes the spectral norm. The $(\mathrm{rank}(M), p)$ Ky-Fan norm is called the *Schatten* p-norm. For $p = 2$ we obtain the Frobenius norm, and for $p = 1$ the *Trace norm*.

Exercise 2 (Balancing, reduction, and stability)

The purpose of this exercise is (i) to make you reflect on the relations and the differences between asymptotically stable systems and systems in which the norm of the state decreases, and (ii) to show that balancing leads to contractivity, but that strict contractivity and asymptotic stability of the reduced system is not automatic.

Consider the autonomous system

$$\Sigma : \frac{d}{dt}x = Ax, \tag{1}$$

with $A \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$. Recall from the lectures the notions of a Hurwitz matrix, stability, asymptotic stability, and the related conditions on the spectrum of A.

In this exercise, let, for $v \in \mathbb{R}^\bullet$, $||v||$ denote the *Euclidean norm,* defined by $||v|| = \sqrt{v^\top v}$, and for $M \in \mathbb{R}^{\bullet\times\bullet}$, let $||M||$ denote the *Euclidean induced norm,* $||M|| := \sup_{0\neq v\in\mathbb{R}^\bullet} \frac{||Mv||}{||v||}$. Of course, this norm is what in exercise 1 was called the *spectral norm.*

Consider also the associated equation

$$\boxed{A^\top P + PA = Q} \tag{2}$$

relating $P, Q \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$, called the *Lyapunov equation.* It is one of the most important equations of all of systems theory.

The significance of this equation in terms of Lyapunov theory is as follows. Let $V(x) = x^\top Px$. Then the derivative of V along solutions of Σ, $\overset{\bullet}{V}_\Sigma$, is given by $\overset{\bullet}{V}_\Sigma(x) = x^\top Qx$.

As a warm-up, remind yourself of the equivalence of

1. Σ is asymptotically stable: all solutions of Σ go to zero as $t \to \infty$.
2. A is Hurwitz.
3. $\exists K \in \mathbb{R}_+, \varepsilon > 0$ such that $||e^{At}|| \leq Ke^{-\varepsilon t}$ for $t \geq 0$.
4. for a $Q = Q^\top \leq 0$ with (A, Q) observable, there is a (unique) solution P to (2) that is symmetric and > 0.

Note that the smallest possible K in the third statement can be quite large, and certainly there is no reason for K always to be 1: asymptotically stable systems can exhibit large overshoots.

Moreover, if A is Hurwitz, and for some $P = P^\top > 0$ (2) gives a $Q = Q^\top \leq 0$, then (A, Q) is an observable pair. If $P = P^\top > 0$ and $Q = Q^\top \leq 0$, then $x^\top Px$ is called a *quadratic Lyapunov function* for Σ. If moreover (A, Q) is observable, the we call it a *strict* quadratic Lyapunov function. Thus Σ is (asymptotically) stable iff it admits a (strict) quadratic Lyapunov function.

<u>Note</u>: this nomenclature is not standard. That Lyapunov functions should be non-increasing along solutions ($Q \leq 0$) is standard. Sometimes, however, Lyapunov functions are not required to be positive for $x \neq 0$ ($P > 0$ is not required).

For the warm-up and the pre-amble, we refer to any of the few good introductory books on Linear System Theory, for example: [1] J.W. Polderman and J.C. Willems, Introduction to Mathematical Systems Theory: A Behavioral Approach, Springer Verlag, 1998. See chapter 7.

Call $\Sigma_1 : \frac{d}{dt}x = A_1 x$ *equivalent* to $\Sigma :\Leftrightarrow \exists$ a non-singular $S \in \mathbb{R}^{\mathtt{n}\times\mathtt{n}}$ such that $A_1 = SAS^{-1}$.

Call Σ *contractive* if every solution $x : \mathbb{R} \to \mathbb{R}^{\mathtt{n}}$ of Σ satisfies $t_1 \geq t_0 \Rightarrow ||x(t_1)|| \leq ||x(t_0)||$, and *strictly contractive* if $t_1 > t_0 \Rightarrow ||x(t_1)|| < ||x(t_0)||$.

1. Prove that strict contractivity implies asymptotic stability.

$$\begin{aligned}
\textbf{strict contractivity} &\Rightarrow \|e^{A}x(0)\| < \|x(0)\|, \quad \textbf{for all } x(0) \in \mathbb{R}^{\mathtt{n}} \\
&\Rightarrow \|e^{A}\| < 1 \\
&\Rightarrow \|e^{AN}\| < \|e^{A}\|^{N}, \quad \textbf{for all } N \in \mathbb{N} \\
&\Rightarrow e^{AN} \xrightarrow{N\to\infty} 0 \\
&\Leftrightarrow e^{At} \xrightarrow{t\to\infty} 0 \\
&\Leftrightarrow \textbf{ asymptotic stability}
\end{aligned}$$

2. Prove that contractivity + asymptotic stability imply strict contractivity.

$$\begin{aligned}
\textbf{contractivity} &\Rightarrow \|e^{At}x(0)\| \textbf{ is non-increasing for } t \geq 0 \\
&\Rightarrow \frac{d}{dt}\|e^{At}x(0)\|^2 \leq 0, \qquad \textbf{for } t \geq 0 \\
&\Rightarrow x^{\top}(0)(A + A^{\top})x(0) \leq 0 \qquad \textbf{(evaluate derivative at } t = 0\textbf{)} \\
&\Rightarrow A + A^{\top} \leq 0 \qquad \textbf{(holds for all } x(0) \in \mathbb{R}^{\mathtt{n}}\textbf{)}
\end{aligned}$$

Assume contractivity but not strict, then

there exists $0 \neq x(0) \in \mathbb{R}^{\mathtt{n}}$ **and** $t > 0$ **such that** $\|e^{At}x(0)\| = \|x(0)\|$

$$\begin{aligned}
&\Rightarrow \|e^{At'}x(0)\| = \|x(0)\|, \quad \textbf{for } t' \in [0,t] \qquad \textbf{(monotonicity)} \\
&\Rightarrow \frac{d}{dt}\|e^{At'}x(0)\|^2 = 0, \quad \textbf{for } t' \in [0,t] \\
&\Rightarrow \big(e^{At'}x(0)\big)^{\top}(A + A^{\top})\big(e^{At'}x(0)\big) = 0, \quad \textbf{for } t' \in [0,t] \qquad \textbf{(evaluate derivatives at 0)} \\
&\Rightarrow (A + A^{\top})\big(e^{At'}x(0)\big) = 0, \quad \textbf{for } t' \in [0,t] \qquad (A + A^{\top} \leq 0) \\
&\Rightarrow (A + A^{\top})A^{\mathtt{k}}x(0) = 0, \quad \textbf{for all } \mathtt{k} \in \mathbb{Z}_+ \qquad \textbf{(differentiate and evaluate at } t' = 0\textbf{)} \\
&\Rightarrow (A + A^{\top})e^{At}x(0) = 0, \quad \textbf{for all } t \in \mathbb{R}_+ \\
&\Rightarrow \|e^{At}x(0)\|^2 = \|x(0)\|^2, \quad \textbf{for all } t \in \mathbb{R}_+
\end{aligned}$$

which contradicts asymptotic stability.

3. Prove that Σ is asymptotically stable iff it is equivalent to a strictly contractive Σ_1.

(if)

$$\begin{aligned}
\Sigma_1 \textbf{ strictly contractive} &\Rightarrow \Sigma_1 \textbf{ asymptotically stable} && \textbf{(see 1)}\\
&\Leftrightarrow SAS^{-1} \textbf{ Hurwitz} \\
&\Leftrightarrow A \textbf{ Hurwitz} && (SAS^{-1} \textbf{ and } A \textbf{ have the same eigenvalues})\\
&\Leftrightarrow \Sigma \textbf{ asymptotically stable}
\end{aligned}$$

(only if)

$$\begin{aligned}
&\Sigma_1 \textbf{ asymptotically stable} \\
&\quad\Leftrightarrow \exists P = P^\top > 0, \textbf{ such that } A^\top P + PA < 0 && \textbf{([1, thm 7.4.7, 3])}\\
&\quad\Leftrightarrow \exists S \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}} \textbf{ non-singular, such that } A^\top S^\top S + S^\top SA < 0 && \textbf{(take } S \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}} : P = S^\top S\textbf{)}\\
&\quad\Leftrightarrow SAS^{-1} + (SAS^{-1})^\top < 0 && \textbf{(premultiply with } (S^\top)^{-1}\\
& && \textbf{and postmultiply with } S^{-1}\textbf{)}\\
&\quad\Rightarrow \Sigma_1 \textbf{ strictly contractive}
\end{aligned}$$

4. Prove that for a symmetric system Σ $(:\Leftrightarrow A = A^\top)$ the following are equivalent:

   (a) Σ is strictly contractive.

   (b) Σ is asymptotically stable.

   (c) $A < 0$.

   What is the analogous statement for contractivity?

(a) $\Rightarrow$ (b) **See 1.**

(b) $\Rightarrow$ (c) **(b) $\Leftrightarrow$ A Hurwitz $\Rightarrow$ the eigenvalues of A are $< 0 \Rightarrow (A = A^\top) A \leq 0 \Leftrightarrow$ (c)**

(c) $\Rightarrow$ (a) **Let $x(0) \neq 0$. Then**

$$\frac{d}{dt}\|e^{At}x(0)\|^2 = \big(e^{At}x(0)\big)^\top (A + A^\top)\big(e^{At}x(0)\big) < 0,$$

so that $\|e^{At}x(0)\|^2$ is monotonically decreasing for $t \geq 0$, which implies strict contractivity.
The analogous statement for contractivity is the equivalence for $A = A^\top$ of

(a) **Σ is contractive,**

(b) **Σ is stable (*i.e.*, all solutions are bounded on $[0, \infty)$), and**

(c) **$A \leq 0$.**

5. Prove that the following conditions are equivalent

   (a) Σ is contractive

   (b) $||e^{At}|| \leq 1$ for all $t > 0$

   (c) $A + A^\top \leq 0$

   (d) Σ and $\frac{d}{dt}x = A^\top x$ have a common quadratic Lyapunov function

   What is the analogous equivalences for strict contractivity?

$$(a) \Leftrightarrow \|e^{At}x(0)\| \le \|x(0)\|, \quad \textbf{for all } x(0) \in \mathbb{R}^n \textbf{ and } t \ge 0$$
$$\Leftrightarrow \sup_{0 \ne x(0) \in \mathbb{R}^n} \frac{\|e^{At}x(0)\|}{\|x(0)\|} \le 1, \quad \textbf{for all } t \ge 0$$
$$\Leftrightarrow \|e^{At}\| \le 1, \quad \textbf{for all } t \ge 0$$
$$\Leftrightarrow \|e^{At}\| \le 1, \quad \textbf{for all } t > 0$$
$$\Leftrightarrow (b)$$

$$(a) \Leftrightarrow \frac{d}{dt}\|e^{At}x(0)\|^2 \le 0, \quad \textbf{for all } x(0) \in \mathbb{R}^n \textbf{ and } t \ge 0$$
$$\Leftrightarrow \big(e^{At}x(0)\big)^\top (A + A^\top)\big(e^{At}x(0)\big), \quad \textbf{for all } x(0) \in \mathbb{R}^n \textbf{ and } t \ge 0$$
$$\Leftrightarrow A + A^\top \le 0$$
$$\Leftrightarrow (c)$$

$$(c) \Leftrightarrow A + A^\top \le 0$$
$$\Leftrightarrow \|x\|^2 \textbf{ is a Lyapunov function for both } \frac{d}{dt}x = Ax \textbf{ and } \frac{d}{dt}x = A^\top x$$
$$\Rightarrow \frac{d}{dt}x = Ax \textbf{ and } \frac{d}{dt}x = A^\top x \textbf{ have a common Lyapunov function}$$
$$\Leftrightarrow (d)$$

$$(d) \Leftrightarrow \exists Q = Q^\top > 0 \textbf{ such that } A^\top Q + QA \le 0 \textbf{ and } AQ + QA^\top \le 0$$
$$\Leftrightarrow \exists Q = Q^\top > 0 \textbf{ such that } (A + A^\top)Q + Q(A + A^\top) \le 0 \qquad \textbf{(add)}$$
$$\Rightarrow \frac{d}{dt}x = (A + A^\top)x \textbf{ is stable}$$
$$\Leftrightarrow A + A^\top \le 0 \qquad \textbf{(see 4)}$$
$$\Leftrightarrow (c)$$

The analogous statement for strict contractivity is. The following statements are equivalent:

(a) **Σ is strictly contractive,**

(b) **$\|e^{At}\| < 1$ for all $t > 0$,**

(c) **$A + A^\top < 0$, and**

(d) **Σ and $\frac{d}{dt}x = A^\top x$ have a common proper quadratic Lyapunov function.**

6. Assume that $\frac{d}{dt}x = Ax + Bu, y = Cx + Du$ is controllable, observable, stable ($:\Leftrightarrow A$ Hurwitz, as is often done we identify stability of the input/output system Σ with asymptotic stability of the autonomous system $\frac{d}{dt}x = Ax$), and balanced (see lecture 7). Prove that $\frac{d}{dt}x = Ax$ is strictly contractive. This shows that balanced state representations bring stability strongly into evidence.

The system $\frac{d}{dt}x = Ax$ is balanced iff

$\exists\, \Delta = \operatorname{diag}(\sigma_1, \sigma_2, \cdots, \sigma_n)$ **with** $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n,$ **s.t.** $A\Delta = \Delta A^\top = -BB^\top,\ A^\top \Delta + \Delta A = -C^\top C.$

Thus $\frac{d}{dt}x = Ax$ and $\frac{d}{dt}x = A^\top x$ have a common proper Lyapunov function, and, by 5, $\frac{d}{dt}x = Ax$ is strictly contractive.

7. Assume again that $\frac{d}{dt}x = Ax + Bu, y = Cx + Du$ is controllable, observable, stable (we call an input/output system stable if A is Hurwitz), and balanced. Partition A as $\left[\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right]$. Prove that $\frac{d}{dt}x_1 = A_{11}x_1$ is contractive, but not always strictly contractive. Whence stability of the reduced systems $\frac{d}{dt}x_1 = A_{11}x_1 + B_1 u, y = C_1 x + D_1 u$ is not an automatism. Do you recall from the lectures which simple condition on the Hankel SV's yields strict contractivity, and hence stability?

For example,

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -1/2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad C = [0 \;\; 1], \qquad D = 0$$

yields a balanced system with $\Delta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$**, but the reduced system** $\tilde{A} = 0$**,** $\tilde{B} = 0$**,** $\tilde{C} = 0$**,** $\tilde{D} = 0$ **is not asymptotically stable. The simple condition on the Hankel singular values that guarantees stability of the reduced system is:**

the neglected singular values are distinct from the non-neglected singular values.

Watch out: the discrete time case is significantly different!