ME 6402 – Lecture 10 [1]

Lyapunov's Linearization Method

February 6 2025

[1] Based on notes created by Murat Arcak and licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Overview:

- Further tools for studying systems based on their linearization
- Define region of attraction
- Obtain Lyapunov estimates of the region of attraction

Additional Reading:

- Khalil, Chapter 4.3-4.7, 8.2

Motivation

In the last class, we introduced the Lyapunov equation[2] and showed that it can be used to test whether or not a matrix A is Hurtwitz, by defining some positide definite matrix Q and solving the Lyapunov equation for P. Then, if the Lyapunov equation has a positive definite solution, we conclude that A is Hurwitz.

[2] $PA + A^T P = -Q$

However, this method has no computational advantage over calculating the eigenvalues of A. In fact, the eigenvalues of A actually tell us more about the behavior of the system compared to solving the Lyapunov equation.

Importantly, the benefit of the Lyapunov equation is rather as a procedure for finding a Lyapunov function for any linear system $\dot{x} = Ax$ when A is Hurwitz, and then allowing us to draw conclusions about the system when Ax is perturbed (including a nonlinear perturbation).

TLDR; We will establish that $V = x^T P x$ (derived using the linearized system) is locally a Lyapunov function for the nonlinear system.

Lyapunov's Linearization Method

Let's go back to our nonlinear system:

$$\dot{x} = f(x), \quad f(0) = 0,$$

where $f : D \to \mathbb{R}^n$ is C^1 and $D \subset \mathbb{R}^n$ is a neighborhood of the equilibrium point $x = 0$.

We can rewrite the nonlinear system to be in the form $\dot{x} = Ax + g(x)$ by leveraging the mean value theorem[3]:

[3] If $f(x)$ is C^1 on $[a, b]$ then there exists c such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

$$\frac{\partial f_i}{\partial x}(z_i) = \frac{f_i(x) - f_i(0)}{x - 0}$$

with z_i being a point on the line segment connecting 0 and x. This can be rearranged to get our desired form:

$$\begin{aligned} f_i(x) &= \cancel{f_i(0)} + \frac{\partial f_i}{\partial x}(z_i)x \\ &= \frac{\partial f_i}{\partial x}(0)x + \underbrace{\left[\frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0)\right]x}_{g_i(x)} \\ f(x) &= Ax + g(x) \end{aligned}$$

The function $g_i(x)$ satisfies:

$$|g_i(x)| \leq \left\| \frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0) \right\| \|x\|$$

By continuity of $[\partial f / \partial x]$, we see that

$$\frac{\|g(x)\|}{\|x\|} \to 0 \quad \text{as} \quad \|x\| \to 0$$

This suggests that in a small neighborhood of the origin we can approximate the nonlinear system by its linearization about the origin.

The following theorem provides the conditions under which we can draw conclusions about the stability of the origin as an equilibrium point for the nonlinear system by investigating the stability as an equilibrium point for the linear system. This is known as Lyapunov's indirect method.

Theorem: (4.7 in Khalil). *Let $x = 0$ be an equilibrium point for the nonlinear system $\dot{x} = f(x)$, where $f : D \to \mathbb{R}^n$ is C^1 and $D \subset \mathbb{R}^n$ is a neighborhood of the origin. Let*

$$A = \left. \frac{\partial f(x)}{\partial x} \right|_{x=0}$$

Then,

1. *If $\Re\{\lambda_i(A)\} < 0$ for all eigenvalues of A, then the origin is asymptotically stable for the nonlinear system.*
2. *If $\Re\{\lambda_i(A)\} > 0$ for some eigenvalue of A, then the origin is unstable for the nonlinear system.*

Note: We can conclude only *local* asymptotic stability from this linearization. Inconclusive if A has eigenvalues on the imaginary axis.

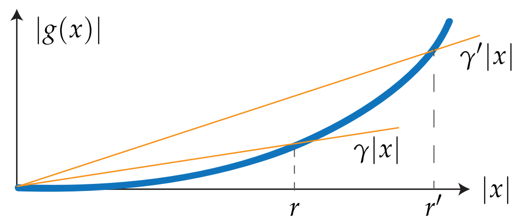
Proof: Find $P = P^T > 0$ such that $A^T P + PA = -Q < 0$. Use $V(x) = x^T P x$ as a Lyapunov function for the nonlinear system $\dot{x} = Ax + g(x)$.

$$\begin{aligned}\dot{V}(x) &= x^T P \dot{x} + \dot{x}^T P x \\ &= x^T P(Ax + g(x)) + (Ax + g(x))^T P x \\ &= x^T (PA + A^T P)x + 2x^T P g(x) \\ &= -x^T Q x + 2x^T P g(x)\end{aligned}$$

The second term is (in general) indefinite, but since we know that $\|g(x)\|/\|x\| \to 0$ as $x \to 0$, we can find a ball around the origin where the second term is negative definite. This is mathematically stated as: for any $\gamma > 0$, there exists $r > 0$ such that

$$\|g(x)\| < \gamma \|x\|, \quad \forall \|x\| < r$$

see the illustration below for the case $x \in \mathbb{R}$.



We can use this to bound our previous expression for $\dot{V}(x)$:

$$\begin{aligned}\dot{V}(x) &= -x^T Q x + 2x^T P g(x) \\ &< -x^T Q x + 2\gamma \|P\| \|x\|^2, \quad \forall \|x\| < r\end{aligned}$$

This can be further bounded by observing the bounds on $x^T Q x$:

$$\lambda_{\min}(Q)|x|^2 \leq x^T Q x \leq \lambda_{\max}(Q)|x|^2$$

Thus, plugging this bound into our expression:

$$\begin{aligned}\dot{V}(x) &< -\lambda_{\min}(Q)|x|^2 + 2\gamma \|P\| |x|^2 \\ &< -\left[\lambda_{\min}(Q) - 2\gamma \|P\|\right] |x|^2\end{aligned}$$

Finally, we can choose $\gamma < \frac{\lambda_{\min}(Q)}{2\|P\|}$ so that $\dot{V}(x)$ is negative definite in a ball of radius $r(\gamma)$ around the origin. We can then appeal to Lyapunov's Stability Theorem (previous lecture) to conclude (local) asymptotic stability.

Region of Attraction

$$R_A = \{x : \phi(t, x) \to 0\} \tag{1}$$

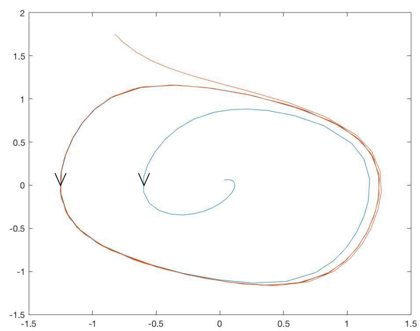
"Quantifies" local asymptotic stability. Global asymptotic stability: $R_A = \mathbb{R}^n$.

Proposition: If $x = 0$ is asymptotically stable, then its region of attraction is an open, connected, invariant set. Moreover, the boundary is formed by trajectories.

Example: van der Pol system in reverse time:

$$\begin{aligned} \dot{x}_1 &= -x_2 \\ \dot{x}_2 &= x_1 - x_2 + x_2^3 \end{aligned} \tag{2}$$

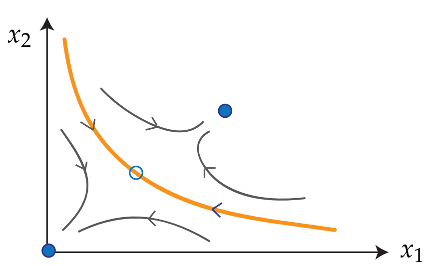
The boundary is the (unstable) limit cycle. Trajectories starting within the limit cycle converge to the origin.



Note: A limit cycle is an isolated periodic orbit.

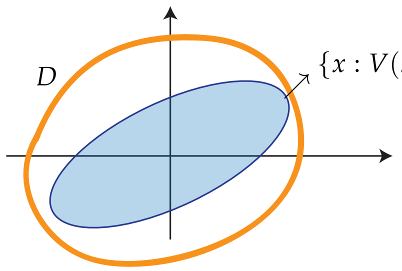
Example: bistable switch:

$$\begin{aligned} \dot{x}_1 &= -ax_1 + x_2 \\ \dot{x}_2 &= \frac{x_1^2}{1 + x_1^2} - bx_2 \end{aligned} \tag{3}$$



Estimating the Region of Attraction with a Lyapunov Function

Suppose $\dot{V}(x) < 0$ in $D - \{0\}$. The level sets of V inside D are invariant and trajectories starting in them converge to the origin. Therefore we can use the largest levet set of V that fits into D as an (under)approximation of the region of attraction.



This estimate depends on the choice of Lyapunov function. A simple (but often conservative) choice is: $V(x) = x^T P x$ where P is selected for the linearization (see p.1).

Example

Code demonstrating this example is provided online.

Consider again the van der Pol system (a slightly different version for now):

$$\begin{aligned}\dot{x}_1 &= -x_2 \\ \dot{x}_2 &= x_1 + (x_1^2 - 1)x_2\end{aligned}$$

We will leverage our Lyapunov equation to find the region of attraction.

First, we linearize the system about the origin:

$$A = \left.\frac{\partial f(x)}{\partial x}\right|_{x=0} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

We can observe that A is Hurwitz, with eigenvalues $\lambda = -0.5 \pm 0.866i$. We can then find a P such that $A^T P + PA = -Q$ where Q is positive definite.

Selecting $Q = I$, the unique solution for P (the solution to the Lyapunov equation $PA + A^T P = -I$) is:

$$P = \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 1 \end{bmatrix}$$

Thus, the quadratic function $V(x) = x^T P x$ is a Lypaunov function for the system in a certain neighborhood of the origin.

Since we want to estimate the region of attraction, we need to find the largest level set of $V(x)$ that fits into the domain D such that

$$\Omega_c = \{V(x) \leq c\}$$

First, checking our derivative condition:

$$
\begin{aligned}
\dot{V}(x) &= \dot{x}^T P x + x^T P \dot{x} \\
&= \begin{bmatrix} -x_2 & x_1 + (x_1^2 - 1)x_2 \end{bmatrix} \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&\quad + \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} -x_2 \\ x_1 + (x_1^2 - 1)x_2 \end{bmatrix} \\
&= 2(-1.5x_1x_2 + 0.5x_2^2 - 0.5x_1(x_1 + (x_1^2 - 1)x_2) + x_2(x_1 + (x_1^2 - 1)x_2)) \\
&= 2(0.5x_2^2 - 0.5x_1^2 - 0.5x_1^3x_2 + x_2^2x_1^2) \\
&= -(x_1^2 + x_2^2) - (x_1^3x_2 - 2x_1^2x_2^2) \\
&\leq -\|x\|_2^2 + |x_1||x_1x_2||x_1 - 2x_2| \\
&\leq -\|x\|_2^2 + \frac{\sqrt{5}}{2}\|x\|_2^4
\end{aligned}
$$

which uses $|x_1| \leq \|x\|_2$, $|x_1x_2| \leq \|x\|_2^2/2$, and $|x_1 - 2x_2| \leq \sqrt{5}\|x\|_2$.

Thus, we can conclude that $\dot{V}(x)$ is negative definite within a ball of radius $r^2 = \frac{2}{\sqrt{5}} = 0.8944$ around the origin.

Finally, we can find a level set within this open ball ($\Omega_c \subset B_r(0)$) by choosing:

$$
c < \min_{\|x\|_2 = r} V(x) = \lambda_{\min}(P) r^2
$$

which gives us:

$$
c = 0.617 < 0.69(0.8944) = 0.6171
$$

The set Ω_c with $c = 0.6171$ is then an underapproximation of the region of attraction for the origin.

A less conservative estimate can be obtained by plotting contours of $\dot{V}(x) = 0$ and $V(x) = c$ for increasing values of c until we find the largest level set that fits into the domain $\{\dot{V}(x) < 0\}$.