ETH zürich IMES

Nonlinear Dynamics and Chaos II.
Assignment 3

Florian Mahlknecht

2020-04-08

1 Eigenvalues of a linearized Hamiltonian system

Given the general Hamiltonian system

$$\dot{\mathbf{x}} = J\,\mathrm{D}\mathcal{H}(\mathbf{x}) \qquad \mathbf{x} \in \mathbb{R}^{2n} \tag{1}$$

Linearization around $\mathbf{x}_0$ yields:

$$\dot{\mathbf{y}} = J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)\,\mathbf{y} \tag{2}$$

with $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$. We are interested in the eigenvalues of the matrix $J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0) \in \mathbb{R}^{2n\times 2n}$. Assuming that λ is an eigenvalue, the characteristic polynomial yields:

$$\det\left(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0) - \mathbb{I}_{2n}\,\lambda\right) = 0 \tag{3}$$

Applying the property given in the exercise statement, we know that the matrix $J^{-1}\left(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)\right)J$ must have the same eigenvalues, therefore:

$$\begin{aligned}
&\det\left(J^{-1}\left(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)\right)J - \mathbb{I}_{2n}\,\lambda\right) = \\
&= \det\left(\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)\,J - \mathbb{I}_{2n}\,\lambda\right) = \\
&= \det\left(J^{\mathsf{T}}\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0) - \mathbb{I}_{2n}\,\lambda\right) = && \mid \text{transposing argument} \\
&= \det\left(-J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0) - \mathbb{I}_{2n}\,\lambda\right) = && \mid \text{using } J \text{ skew-symmetric} \\
&= \det\left(-\mathbb{I}_{2n}\right)\,\det\left(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0) + \mathbb{I}_{2n}\,\lambda\right) = \\
&= \det\left(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0) - \mathbb{I}_{2n}\,(-\lambda)\right) = 0
\end{aligned}$$

Which is the characteristic polynomial for the eigenvalue $-\lambda$.

Note that the matrix of the linearized system of the Hamiltonian matrix $J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)$, has dimension $2n \times 2n$. Therefore in general, there will be $2n$ eigenvalues. With the above result we have shown that those eigenvalue pairs will be additive opposites, such that:

$$\lambda \in \operatorname{eig}(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)) \quad \Rightarrow \quad -\lambda \in \operatorname{eig}(J\,\mathrm{D}^2\mathcal{H}(\mathbf{x}_0)) \tag{4}$$

□ *q.e.d.*

fmahlknecht@student.ethz.ch
19-945-351

2 Gradient systems

Consider in the following the system:

$$\dot{\mathbf{x}} = -\mathrm{D}V(\mathbf{x}) \qquad \mathbf{x} \in \mathbb{R}^n \tag{5}$$

2.1 Linearized gradient system

If we linearize the system around a fixed point $\mathbf{x}_0$, using $\mathbf{y} = \mathbf{x} - \mathbf{x}_0$, we get:

$$\dot{\mathbf{y}} = -\mathrm{D}^2 V(\mathbf{x_0})\,\mathbf{y} \tag{6}$$

Note that for this to make sense $V(\mathbf{x})$ should be at least differentiable twice, i.e. $\in C^2$. The Hessian of $V(\mathbf{x})$ is then *symmetric*. For simplicity denote $A = -\mathrm{D}^2 V(\mathbf{x_0})$, with $A^\top = A$ and all *real* entries, i.e. $A \in \mathbb{R}^{n \times n}$, which yields:

$$\overline{A} = A \tag{7}$$

with $\overline{\bullet}$ denoting the complex conjugate. Suppose now that there was a complex eigenvalue $\lambda = a + i\,b$, with $\overline{\lambda} = a - i\,b$. We have in general:

$$\overline{\lambda\,\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}}$$

Applying the given facts to the definition of eigenvalue yields:

$$A\,\mathbf{x} = \lambda\,\mathbf{x} \qquad \Rightarrow \qquad A\,\overline{\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}} \qquad \Rightarrow \qquad \overline{\mathbf{x}}^\top A = \overline{\mathbf{x}}^\top \overline{\lambda} \tag{8}$$

$$\overline{\mathbf{x}}^\top A\,\mathbf{x} = \overline{\mathbf{x}}^\top \lambda\,\mathbf{x} \qquad\qquad \overline{\mathbf{x}}^\top A\,\mathbf{x} = \overline{\mathbf{x}}^\top \overline{\lambda}\,\mathbf{x} \tag{9}$$

where we left multiplied with $\overline{\mathbf{x}}^\top$ and right multiplied with $\mathbf{x}$ respectively. Since the left hand side is equal in both equations in (9), we have:

$$\overline{\mathbf{x}}^\top \lambda\,\mathbf{x} = \overline{\mathbf{x}}^\top \overline{\lambda}\,\mathbf{x} \tag{10}$$

Given that $\overline{\mathbf{x}}^\top \mathbf{x}$ is the length squared > 0, we must have:

$$\lambda = \overline{\lambda} \quad \Rightarrow \quad \mathrm{Im}(\lambda) = 0 \tag{11}$$

$\square$ *q.e.d.*

2.2 Asymptotic stability

First, let us realize that while moving on solutions of eq. (5), we will have a strict decrease on $V(x)$:

$$\frac{dV(\mathbf{x}(t))}{dt} = \frac{dV}{d\mathbf{x}}\dot{\mathbf{x}} = -\frac{dV}{d\mathbf{x}} \cdot \frac{dV}{d\mathbf{x}} < 0 \qquad \Longleftrightarrow \qquad \frac{dV}{d\mathbf{x}} \neq \mathbf{0} \tag{12}$$

as long as we are not on a critical point of V, which is a fixed point of the gradient system eq. (5). Excluding fixed points means excluding *constant* trajectories.

Florian Mahlknecht
fmahlknecht@student.ethz.ch
19-945-351

If along this strict decrease in a neighborhood of $\mathbf{x}_0$ we approach the fixed point $\mathbf{x}_0$, i.e. towards the value $V(\mathbf{x}_0)$, we will have an asymptotically stable fixed point. This is nothing but Lyapunovs direct method with the Lyapunov function $\tilde{V}(\mathbf{x}) = V(\mathbf{x}) - V(\mathbf{x}_0)$. For applying Lyapunov we need to impose that $V(\mathbf{x})$ shows an *isolated minimum* in x_0, i.e.

$$\tilde{V}(\mathbf{x}) > 0 \quad \forall \quad \mathbf{x} \in U(\mathbf{x}_0) \setminus \{\mathbf{x}_0\} \quad \Longleftrightarrow \quad \mathrm{D}^2 V(\mathbf{x}_0) > 0 \quad \text{for } V \in C^2 \tag{13}$$

2.3 Absence of periodic orbits

Ad absurdum suppose that there was a non trivial T-periodic solution of eq. (5), such that:

$$\tilde{\mathbf{x}}(t) = \tilde{\mathbf{x}}(t+T) \neq const \qquad \forall\ t \in \mathbb{R}, \quad T > 0 \tag{14}$$

Given that it must not be constant, the derivative in time will be different from zero. From an alternative point of view, by uniqueness of solutions, along this trajectory we are not allowed to encounter any fixed points, i.e.

$$\dot{\tilde{\mathbf{x}}}(t) \neq \mathbf{0} \qquad \forall\ t \in \mathbb{R} \tag{15}$$

by periodicity from eq. (14) we must have:

$$V(\tilde{\mathbf{x}}(t)) = V(\tilde{\mathbf{x}}(t+T)) \tag{16}$$

By looking at $V(\tilde{\mathbf{x}}(t))$ as a function of time, defined at least on the closed time interval $[t, t+T]$, and being differentiable, we can apply Rolle's theorem from Calculus, which given eq. (16) concludes that $\exists\ t' \in [t, t+T]$ such that:

$$\frac{dV}{dt}(t') = 0 \tag{17}$$

However, as we have already shown in eq. (12),

$$\frac{dV(\tilde{\mathbf{x}}(t))}{dt} = -\left\|\dot{\tilde{\mathbf{x}}}(t)\right\|^2 < 0 \tag{18}$$

This is strictly negative due to eq. (15), which then constitutes a contradiction to eq. (17).

$\square$ *q.e.d.*

Note that equivalently, eq. (18) would imply that $V(\tilde{\mathbf{x}}(t)$ is a strictly monotonic decreasing function of time, which collides with eq. (16), and is therefore a contradiction again.

2.4 Numerical method for finding local extrema

We may use the results we have obtained so far to find local extrema for a general smooth function $f(\mathbf{x})$. The goal is to translate extrema of the function to asymptotically stable fixed points of the system eq. (5).

Local maxima Suppose $\mathbf{x}_0$ is an isolated local maximum of $f(\mathbf{x})$, given that it is smooth we have:

$$\mathrm{D}f(\mathbf{x}_0) = \mathbf{0}$$
$$\mathrm{D}^2 f(\mathbf{x}_0) < 0$$

By choosing $V(\mathbf{x}) = -f(\mathbf{x})$, $\mathbf{x}_0$ becomes a (nonlinear) asymptotically stable fixed point of the gradient system given in eq. (5), since we satisfy the previously derived condition in eq. (13):

$$\mathrm{D}^2 V(\mathbf{x}_0) = -\mathrm{D}^2 f(\mathbf{x}_0) > 0 \tag{19}$$

Local minima the same reasoning holds for a local minimum in $\mathbf{x}_0$, which yields:

$$\mathrm{D}f(\mathbf{x}_0) = \mathbf{0}$$
$$\mathrm{D}^2 f(\mathbf{x}_0) > 0$$

We therefore choose $V(\mathbf{x}) = f(\mathbf{x})$, which again guarantees asymptotic stability for $\mathbf{x}_0$:

$$\mathrm{D}^2 V(\mathbf{x}_0) = \mathrm{D}^2 f(\mathbf{x}_0) > 0 \tag{20}$$

Algorithm By applying a simple forward Euler discretization scheme on the gradient system in eq. (5), we can derive an algorithm which seeks the minimum in an iterative fashion, as illustrated in algorithm 1. As shown before, for the maximum we can simply change the sign. The algorithm is indeed known as gradient descent, or gradient ascent in case of a maximum.

Algorithm 1 Gradient descent

function GetMinimum($f(\mathbf{x}), \mathbf{x}_0$)
  $\Delta t \leftarrow 0.001$ ▷ Denotes the timesteps to take
  $maxIterations = 10000$

  $i \leftarrow 0$
  $\mathbf{x}^\star \leftarrow \mathbf{x}_0$
  $\Delta\mathbf{x} \leftarrow -\mathrm{D}f(\mathbf{x}^\star)\,\Delta t$

  while $\|\Delta\mathbf{x}\| > \varepsilon$ & $i < maxIterations$ **do**
    $\mathbf{x}^\star \leftarrow \mathbf{x}^\star + \Delta\mathbf{x}$
    $\Delta\mathbf{x} \leftarrow -\mathrm{D}f(\mathbf{x}^\star)\,\Delta t$
    $i \leftarrow i + 1$
  end while
  return $\mathbf{x}^\star$
end function

Note that this algorithm is sensitive on the choice of the timestep as well as of the initial condition. To guarantee convergence, we need to choose the initial condition to be in the neighborhood of the true extrema, where the Lyapunov theorem can be applied. Furthermore, Δt needs to be small enough to guarantee stability in the forward integartion scheme. Of course the step size is a trade of between stability and convergence rate, i.e. the number of necessary iterations.

Florian Mahlknecht
fmahlknecht@student.ethz.ch
19-945-351

3 Graph as manifold

Given a smooth function

$$\mathbf{f} : X \to Y \qquad X \subset \mathbb{R}^n, \text{ manifold} \tag{21}$$

the graph

$$\text{graph}(\mathbf{f}) = \{(\mathbf{x}, \mathbf{y}) \in X \times Y : \mathbf{y} = \mathbf{f}(\mathbf{x})\} \tag{22}$$

is a n-dimensional manifold, for which we can write out a global parametrization

$$\begin{aligned} \Phi : X &\to \text{graph}(\mathbf{f}) \\ \mathbf{x} &\mapsto (\mathbf{x}, \mathbf{f}(\mathbf{x})) \\ \Phi^{-1} : \text{graph}(\mathbf{f}) &\to X \\ (\mathbf{x}, \mathbf{f}(\mathbf{x})) &\mapsto \mathbf{x} \end{aligned}$$

Note that the global parametrization can always be applied to any relatively open neighborhood around $\mathbf{x} \in X$, because X itself is assumed to be a manifold, which guarantees such a neighborhood's existence.

Φ and Φ^{-1} are both smooth and differentiable, with the derivatives:

$$\begin{aligned} \mathrm{D}\Phi &= \begin{bmatrix} \mathbb{I}_n \\ \mathrm{D}\mathbf{f} \end{bmatrix} \\ \mathrm{D}\Phi^{-1} &= \begin{bmatrix} \mathbb{I}_n & \mathbf{0} \end{bmatrix} \end{aligned}$$

where $\mathrm{D}\mathbf{f}$ denotes the Jacobian of $\mathbf{f}$. As we can see it is a one to one mapping, and all we need for reconstructing the manifold is the $\mathbf{x}$ position: The function value $\mathbf{f}(\mathbf{x})$ has no influence in the encoding, as $\mathrm{D}\Phi^{-1}$ shows.

□ *q.e.d.*

4 Tangent space parametrization independence

Assume for some $\mathbf{x} \in M$ to have two distinct parametrization of the manifold from two open sets $V, W \subset \mathbb{R}^k$, to the relatively open set $U \subset M \subset \mathbb{R}^n$ such that

$$\begin{aligned} \Phi_1 : V &\to U \\ \mathbf{v} &\mapsto \Phi_1(\mathbf{v}) \in M \\ \Phi_2 : W &\to U \\ \mathbf{w} &\mapsto \Phi_1(\mathbf{w}) \in M \end{aligned}$$

We can define a trivial linear identity mapping between the two open sets on the manifold, such that:

fmahlknecht@student.ethz.ch
19-945-351

$$
\begin{aligned}
I : U &\to U \\
\mathbf{x} &\mapsto \mathbb{I}_n\, \mathbf{x}
\end{aligned}
$$

As we have seen in class, any linear mapping between manifolds *induces* a linear mapping between the open sets V and W., i.e.

$$F : V \to W \qquad F = \Phi_2^{-1} \circ I \circ \Phi_1 \tag{23}$$

By definition of manifold, the parametrization must be such that $\exists\, \mathbf{y} \in V$ for which

$$\Phi_1(\mathbf{y}) = \Phi_2(F(\mathbf{y})) = \mathbf{x}$$

Remind that in our simplified case, Φ_1 and Φ_2 are just different parameterizations of the *same* k dimensional manifold, such that both $V, W \subset \mathbb{R}^k$, therefore F maps just vectors from $\mathbb{R}^k$ to $\mathbb{R}^k$. Since V, W are hence two open sets in the euclidean space, $\mathrm{D}F$ is well defined, as discussed in class. Moreover, we have all other mappings well defined, as illustrated in fig. 1.

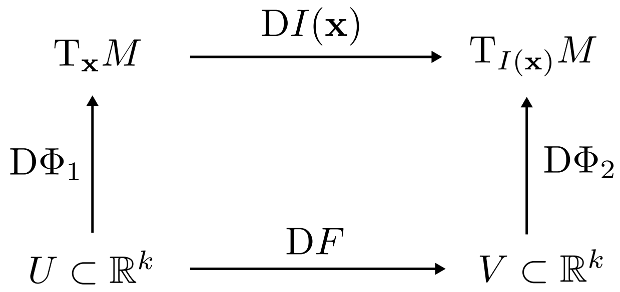

Figure 1: Illustration of differential mapping involved

In this way, the chain rule yields:

$$\mathrm{D}I(\mathbf{x}) = \mathrm{D}\Phi_2 \circ \mathrm{D}F \circ \mathrm{D}\Phi_1^{-1} \tag{24}$$

In our special case, Φ_1 and Φ_2 are just different parameterization of the same manifold, which is why there must be the identity mapping I between the same manifold. Its differential, in fact in any point, is simply the identity matrix $\mathbb{I}_n$. Equation (24) yields therefore:

$$\mathbb{I}_n = \mathrm{D}\Phi_2 \circ \mathrm{D}F \circ \mathrm{D}\Phi_1^{-1}$$

In other words, we map back always in the same identical tangent space, no matter which arbitrary parametrization we follow. As expected we have a one to one mapping between $\mathrm{T}_{\mathbf{x}}M$ and $\mathrm{T}_{\mathbf{I(x))}}M$, and can therefore conclude:

$$\mathrm{T}_{\mathbf{x}}M \equiv \mathrm{T}_{\mathbf{\Phi_2(F(\Phi_1^{-1}(x)))}}M \tag{25}$$

i.e. the tangent spaces are equal independently of the parametrization.

□ *q.e.d.*

Florian Mahlknecht
fmahlknecht@student.ethz.ch
19-945-351

5 Tangent bundle

Similar to the graph example, we can proof the tangent bundle to be a manifold, thanks to the ordered tuple which keeps track of the position:

$$TM \coloneqq \bigcup_{\mathbf{x}\in M} (\mathbf{x}, \mathrm{T}_{\mathbf{x}}M) \tag{26}$$

For any $\mathbf{x} \in M$, we know that there exists a neighborhood $U_{\mathbf{x}} \subset M$, since M is a manifold, for which we can apply the parametrization:

$$\begin{aligned} \Phi(\mathbf{x}) &= (\mathbf{x}, \mathrm{T}_{\mathbf{x}}M) & \Phi &: M \to TM \\ \Phi^{-1}(\mathbf{x}, \mathrm{T}_{\mathbf{x}}M) &= \mathbf{x} & \Phi^{-1} &: TM \to M \end{aligned}$$

The mapping is diffeomorphic, since the underlying operations, identity and differentiation are both differentiable and continuous. TM is therefore a manifold, of the same dimension of the underlying manifold M.

□ *q.e.d.*

fmahlknecht@student.ethz.ch
19-945-351