

Questions

1. Towards upper-triangulation by an orthonormal basis

In lecture, we have been motivated by the goal of getting to a coordinate system in which the eigenvalues of a matrix representing a linear operation are on the diagonal. When this is done to the A matrix representing a dynamical system (whether in continuous-time as a system of differential equations or in discrete-time as a relationship between the next state and the previous one), we can view the system as a cascade of scalar systems — with each one potentially being an input to the ones that come “after” it. We saw this in lecture, but it is good to spend more time to really understand this argument.

Note that in the next homework, you will be asked to derive this in a more formal way using induction. Here we will just provide some key steps along the way to a recursive understanding. Here, as in lecture, we will restrict attention to matrices that have all real eigenvalues.

In order for you to better understand the steps, you can consider a concrete case

$$S_{[3 \times 3]} = \begin{bmatrix} \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

and figure out the general case by abstracting variables. This particular matrix has an additional special property of symmetry, but we won't be invoking that here.

(a) Characteristic polynomial warm-ups: **Show that the characteristic polynomial of square matrix A is the same as that of the square matrix $T^{-1}AT$ for any invertible T .**

(b) Consider a non-zero vector $\vec{u}_0 \in \mathbb{R}^n$. **Can you think of a way to extend it to a set of basis vectors for \mathbb{R}^n ?** In other words, find $\vec{u}_1, \dots, \vec{u}_{n-1}$, such that $\text{span}(\vec{u}_0, \vec{u}_1, \dots, \vec{u}_{n-1}) = \mathbb{R}^n$. To begin with, consider

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}.$$

Can you get an orthonormal basis from what you just constructed?

(c) Now consider a real eigenvalue λ_0 , and the corresponding eigenvector $\vec{g}_0 \in \mathbb{R}^n$ of a square matrix $M \in \mathbb{R}^{n \times n}$. From the previous part, we can extend \vec{g}_0 to an orthonormal basis of \mathbb{R}^n , denoted by

$$V = [\vec{v}_0, \vec{v}_1, \dots, \vec{v}_{n-1}]$$

where $\vec{v}_0 = \frac{\vec{g}_0}{\|\vec{g}_0\|}$.

Our goal is to look at what the matrix M looks like in the coordinate system defined by the basis V .

Compute $V^T M V$ by writing $V = [\vec{v}_0, R]$, where $R \triangleq [\vec{v}_1, \dots, \vec{v}_{n-1}]$. If you prefer, you can do this and the next question with the concrete $S_{[3 \times 3]}$ first.

(d) Define $Q = R^T MR$. Look at the first column and the first row of $V^T MV$ and **show that**

$$M = V \begin{bmatrix} \lambda_0 & \vec{a}^T \\ \vec{0} & Q \end{bmatrix} V^T$$

Here the \vec{a} is just something arbitrary.

(e) **What can you say about the characteristic polynomial $\det(\lambda I - Q)$ of Q in relationship to the characteristic polynomial of the original M ?** Recall that Q is an $(n-1) \times (n-1)$ matrix.

(f) Now, we can recurse on Q to get:

$$Q = [\vec{u}_0, Y] \begin{bmatrix} \lambda_1 & \vec{b}^T \\ \vec{0} & P \end{bmatrix} [\vec{u}_0, Y]^T$$

where we have taken $\vec{u}_0 \in \mathbb{R}^{n-1}$, a eigenvector of Q , associated with eigenvalue λ_1 . Again \vec{u}_0 is extended into an orthonormal basis $[\vec{u}_0, \vec{u}_1, \dots, \vec{u}_{n-2}]$ of \mathbb{R}^{n-1} . We denote $Y \triangleq [\vec{u}_1, \dots, \vec{u}_{n-2}]$.

Plug this into M to show that:

$$M = [\vec{v}_0, R\vec{u}_0, RY] \begin{bmatrix} \lambda_0 & a_1 & \vec{a}^T \\ 0 & \lambda_1 & \vec{b}^T \\ \vec{0} & \vec{0} & P \end{bmatrix} [\vec{v}_0, R\vec{u}_0, RY]^T$$

Again, using the concrete case may help you first.

(g) **Show that the matrix $[\vec{v}_0, R\vec{u}_0, RY]$ is still orthonormal.**

(h) Perform the above process recursively - what will you get in the end?

Contributors:

- Yuxun Zhou.
- Edward Wang.
- Anant Sahai.
- Sanjit Batra.