Discussion 9A

The following notes are useful for this discussion: Note 13, Note 15.

1. Towards Upper-Triangularization By An Orthonormal Basis

Previously in this course, we have seen the value of changing our coordinates to be eigenbasis-aligned, because we can then view the system as a set of parallel scalar systems. Diagonalization causes these scalar equations to be fully uncoupled such that they can be solved separately. But even when we cannot diagonalize, we can *upper-triangularize* such that we can still solve the equations one at a time, from the "bottom up".

To better understand the steps involved, we will use the following concrete example:

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

and solve the general case by abstracting variables. Note that there is a datahub link to a jupyter notebook on the website, which will allow you to perform the numerical calculations quickly.[1]

(a) 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 $\vec{u}_1, \cdots, \vec{u}_{n-1}$, such that $\text{span}(\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-1}) = \mathbb{R}^n$. **To make things concrete, consider $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$. Can you get an orthonormal basis where the first vector is a multiple of this vector?**

(HINT: What was the last discussion all about? Also, the given vector isn't normalized yet!)

[1]This particular matrix has an additional special property of symmetry, but we won't be invoking that here.

(b) Now consider a real eigenvalue λ_1, and the corresponding (normalized) eigenvector $\vec{v}_1 \in \mathbb{R}^n$ of $M \in \mathbb{R}^{n \times n}$ ($M\vec{v}_1 = \lambda_1 \vec{v}_1$). We know we can extend $\vec{v}_1$ to an orthonormal basis of $\mathbb{R}^n$. We will denote the basis by

$$U = \begin{bmatrix} | & | & \cdots & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ | & | & \cdots & | \end{bmatrix} \tag{2}$$

where $\vec{u}_1 = \vec{v}_1$ (note that this eigenvector is already normalized).

Our goal is to look at what the matrix M looks like in the coordinate system defined by the basis U. **Compute $U^\top M U$ by writing $U = \begin{bmatrix} \vec{v}_1 & R \end{bmatrix}$, where $R := \begin{bmatrix} | & | & & | \\ \vec{r}_1 & \vec{r}_2 & \cdots & \vec{r}_{n-1} \\ | & | & & | \end{bmatrix}$.** *(Note : $\vec{r}_i = \vec{u}_{i+1}$)*

(c) **Verify that $U^{-1} = U^\top$, where U is the matrix we get from Gram-Schmidt process.**

(d) **Look at the first column and the first row of $U^\top MU$ and show that:**

$$M = U \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} U^\top \tag{3}$$

where $Q = R^\top MR$. Here, $\vec{a}$ is a vector related to M, R, and $\vec{v}_1$ (we will show the relation!).

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

$$Q = \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix} \begin{bmatrix} \lambda_2 & \vec{b}^\top \\ \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}^\top \tag{4}$$

where we have taken $\vec{v}_2 \in \mathbb{R}^{n-1}$, a normalized eigenvector of Q, associated with eigenvalue λ_2. Again $\vec{v}_2$ is extended into an orthonormal basis to form $\begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}$.
Plug this form of Q into M above, to show that:

$$M = \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix} \begin{bmatrix} \lambda_1 & \breve{a}_1 & \breve{\vec{a}}_{\text{rest}}^\top \\ 0 & \lambda_2 & \vec{b}^\top \\ \vec{0} & \vec{0} & P \end{bmatrix} \begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}^\top \tag{5}$$

where we define $\breve{\vec{a}}$ to be the "adjusted" $\vec{a}$ to account for the subsitution of Q; $\breve{\vec{a}}^\top = \vec{a}^\top \begin{bmatrix} \vec{v}_2 & Y \end{bmatrix}$.

(f) **(PRACTICE) Show that the matrix $\begin{bmatrix} \vec{v}_1 & R\vec{v}_2 & RY \end{bmatrix}$ is still orthonormal.**

(g) **(PRACTICE)** We have shown how to upper triangularize a 3×3 and a 2×2 matrix. **How can we generalize this process to any $n \times n$ matrix M?**

(h) **(PRACTICE) Show that the characteristic polynomial of square matrix M is the same as that of the square matrix UMU^{-1} for any invertible U.** You should use the key property $\det(AB) = \det(A)\det(B)$ for square matrices.

Contributors:

- Neelesh Ramachandran.
- Yuxun Zhou.
- Edward Wang.
- Anant Sahai.
- Sanjit Batra.
- Pavan Bhargava.
- Kumar Krishna Agrawal.

