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. If we have a diagonalization, then these scalar equations are fully uncoupled, and can therefore be treated completely separately. But even when we cannot diagonalize, we can upper-triangularize in a way that allows us to solve the equations one at a time, from the "bottom up".

In this problem, to better understand the steps involved, we will use the following concrete example:

$$M = S_{[3\times3]} = \begin{bmatrix} \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{bmatrix} \tag{1}$$

and figure out 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 to connect the symbolic analysis to an example, but without being time-consuming in the process.[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!

$\{\vec{u}_0, \vec{e}_1 \ldots \vec{e}_n\} \rightarrow$ G.S.

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

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

R

$n \times n-1$

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** $R \triangleq \begin{bmatrix} | & | & \cdots & | \\ \vec{r}_1 & \vec{r}_2 & \cdots & \vec{r}_{n-1} \\ | & | & \cdots & | \end{bmatrix}$. *(Note : $\vec{r}_i = \vec{u}_{i+1}$)*

$$\begin{bmatrix} \vec{v}_1^T \\ R^T \end{bmatrix} M [\vec{v}_1, R] = \begin{bmatrix} \vec{v}_1^T \\ R^T \end{bmatrix} [M\vec{v}_1 \quad MR] = \begin{bmatrix} \vec{v}_1^T \\ R \end{bmatrix} [\lambda_1 \vec{v}_1 \quad MR]$$

$$\begin{bmatrix} \lambda_1 \vec{v}_1^T \vec{v}_1 & \vec{v}_1^T MR \\ \lambda_1 R^T \vec{v}_1 & R^T MR \end{bmatrix} = \begin{bmatrix} \lambda_1 & \vec{v}_1^T MR \\ \lambda_1 R^T \vec{v}_1 & R^T MR \end{bmatrix}$$

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

$$U^T U = I$$

$$\begin{bmatrix} \vec{u}_1^{\,T} \\ \vdots \\ \vec{u}_n^{\,T} \end{bmatrix} \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix} = \begin{bmatrix} \vec{u}_1^{\,T}\vec{u}_1 & & \vec{u}_n^{\,T}\vec{u}_1 \\ & \ddots & \\ \vec{u}_j^{\,T}\vec{u}_i & & \vec{u}_n^{\,T}\vec{u}_n \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}}_{I}$$

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

$$M - U \begin{bmatrix} \lambda_1 & \vec{a}^\top \\ \vec{0} & Q \end{bmatrix} U^\top \tag{2}$$

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

$$U^T M U = \begin{bmatrix} \lambda_1 & \vec{v}_1^T M R \\ \lambda_1 R^T v_1 & R^T M R \end{bmatrix} \Rightarrow M = U \begin{bmatrix} \lambda_1 & \vec{v}_1^T M R \\ \lambda_1 R^T M & R^T M R \end{bmatrix} U^T$$

$$Q = R^T M R$$

$$\vec{a}^T = \vec{v}_1^T M R$$

$$R \perp \vec{v}_1 \Rightarrow R^T \vec{v}_1 = \vec{0}$$

$$\Rightarrow M = U \begin{bmatrix} \lambda_1 & \vec{a}^T \\ \vec{0} & Q \end{bmatrix} U^T$$

Q

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

$$Q = \overset{U_2}{\begin{bmatrix}\vec{v}_2 & Y\end{bmatrix}} \overset{T_2}{\begin{bmatrix}\lambda_2 & \vec{b}^\top \\ \vec{0} & P\end{bmatrix}} \begin{bmatrix}\vec{v}_2 & Y\end{bmatrix}^\top \tag{3}$$

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

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:

$$Q = R^T M R$$

$$M = \begin{bmatrix}\vec{v}_1 & R\vec{v}_2 & RY\end{bmatrix} \begin{bmatrix}\lambda_1 & \breve{a}_1 & \breve{\vec{a}}_{\text{rest}}^\top \\ 0 & \lambda_2 & \vec{b}^\top \\ \vec{0} & \vec{0} & P\end{bmatrix} \begin{bmatrix}\vec{v}_1 & R\vec{v}_2 & RY\end{bmatrix}^\top \tag{4}$$

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

$$U_2 T_2 U_2^T = R^T M R$$

$$\Rightarrow T_2 = U_2^T R^T M R U_2$$

$$M = U\begin{bmatrix}\lambda_1 & \vec{a}^T \\ \vec{0} & U_2 T_2 U_2^T\end{bmatrix}U^T$$

$$M = U\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2^T\end{bmatrix}\begin{bmatrix}\lambda_1 & \vec{a}^T \\ \vec{0} & U_2 T_2 U_2^T\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2^T\end{bmatrix}U^T$$

$$= U\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix}\begin{bmatrix}\lambda_1 & \vec{a}^T U_2 \\ \vec{0} & T_2\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2^T\end{bmatrix}U^T$$

$$= \begin{bmatrix}V_1 & R\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2\end{bmatrix}\begin{bmatrix}\lambda_1 & \breve{\vec{a}}^T \\ \vec{0} & T_2\end{bmatrix}\begin{bmatrix}1 & 0 \\ 0 & U_2^T\end{bmatrix}\begin{bmatrix}\vec{V}_1 & R^T\end{bmatrix}$$

$$= \begin{bmatrix}V_1 & RU_2\end{bmatrix}\begin{bmatrix}\lambda_1 & \breve{\vec{a}}^T \\ \vec{0} & T_2\end{bmatrix}\begin{bmatrix}V_1 & RU_2\end{bmatrix}^T$$

$$\begin{bmatrix}\vec{v}_1 & R\vec{v}_2 & RY\end{bmatrix}\begin{bmatrix}\lambda_1 & \vec{x}^T \\ \vec{0} & T_2\end{bmatrix}\begin{bmatrix}\vec{v}_1 & R\vec{v}_2 & RY\end{bmatrix}^T$$

$$= \quad '' \quad \begin{bmatrix}\lambda_1 & \vec{a}_1 & \vec{a}_{rest}^T \\ 0 & \lambda_2 & \vec{b}^T \\ \vec{0} & \vec{0} & P\end{bmatrix} \quad ''$$

$$V_2 = \begin{bmatrix} v_2 & Y \end{bmatrix}$$

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

What did the 16B student say when asked to decompose a matrix?

Schur!

Feedback:

https://tinyurl.com/manav16b