Orthogonal Diagonalization: The Spectral Theorem

Prerequisites:

- Orthogonal transformations
- Diagonalization

Definition 1: *Orthogonal Diagonalization*

Let $\boxed{\mathsf{B}}$ be an $N \times N$ matrix. We say that $\boxed{\mathsf{B}}$ is **orthogonally diagonalizable** if there is an *orthogonal matrix* $\boxed{\mathsf{G}}$ so that

$$\boxed{\mathsf{G}}^{-1} \cdot \boxed{\mathsf{B}} \cdot \boxed{\mathsf{G}} \qquad \left(\text{which is equal to } \boxed{\mathsf{G}}^{t} \cdot \boxed{\mathsf{B}} \cdot \boxed{\mathsf{G}}\right)$$

is a *diagonal* matrix.

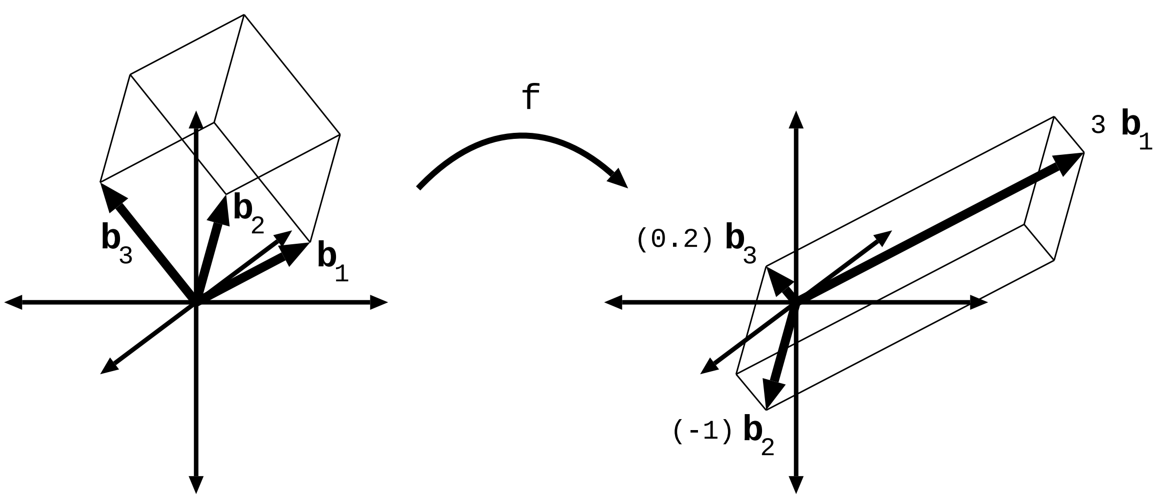

Figure 1: f has **orthonormal** eigenvectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_3$, with eigenvalues 3, -1, and 0.2, respectively

Recall that diagonalization is useful because it reveals the existence of a *basis of eigenvectors* for a transformation. *Orthogonal* diagonalization is even better: it reveals the existence of a *orthonormal* basis of eigenvectors for the transformation. (see Figure 1).

First we will prove the following partial result

Theorem 2: *Triangulation Theorem*

Let $\boxed{\mathsf{F}}$ be a $N \times N$ matrix whose characteristic polynomial **factors completely** —that is:

$$c_{\boxed{F}}(x) \;=\; (x-\lambda_1)\cdot(x-\lambda_2)\cdot\ldots\cdot(x-\lambda_N),$$

where $\lambda_1, \ldots, \lambda_N$ are real numbers. Then there is an **orthogonal matrix** $\boxed{\mathsf{P}}$ so that

$$\boxed{\mathsf{P}}\cdot\boxed{\mathsf{F}}\cdot\boxed{\mathsf{P}}^{-1}$$

is **upper triangular**.

Proof: We will prove this by induction on N.

Base Case ($N = 1$): A 1×1 matrix is automatically upper triangular, so this is trivial.

Induction: Suppose, inductively, that the theorem is true for $\mathbb{R}^{N-1}$.

Let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be the linear map: $f(\mathbf{x}) = \boxed{\mathrm{F}}\cdot\mathbf{x}$. Let λ_1 be the first eigenvalue for f. Let $\mathbf{b}_1$ be a corresponding eigenvector. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ be an **orthonormal basis** for $\mathbb{R}^N$, with $\mathbf{b}_1$ as its first element. Let $\boxed{\mathrm{B}}$ be the **change-of-basis matrix** from the standard basis into $\mathcal{B}$. Then we know:

- $\boxed{\mathrm{B}}$ is an orthogonal matrix.
- $\widetilde{\boxed{\mathrm{F}}} = \boxed{\mathrm{B}}\cdot\boxed{\mathrm{F}}\cdot\boxed{\mathrm{B}}^{-1}$ is the matrix representation of f relative to $\mathcal{B}$.

Thus, since $\mathbf{b}_1$ is an eigenvector of f with eigenvalue λ_1, the matrix $\widetilde{\boxed{\mathrm{F}}}$ must have the form:

$$\widetilde{\boxed{\mathrm{F}}} = \left[\begin{array}{c|cccc} \lambda_1 & * & * & \ldots & * \\ \hline 0 & & & & \\ 0 & & & & \\ \vdots & & \mathbf{F}_1 & & \\ 0 & & & & \end{array}\right],$$

where $\boxed{\mathbf{F}_1}$ is an $(N-1)\times(N-1)$, matrix, having eigenvalues $\lambda_2, \lambda_3, \ldots, \lambda_N$. Thus, by the induction hypothesis, there is an $(N-1)\times(N-1)$ orthogonal matrix $\boxed{\mathbf{C}_1}$ so that

$$\boxed{\nabla} \;=\; \boxed{\mathbf{C}_1}\cdot\boxed{\mathbf{F}_1}\cdot\boxed{\mathbf{C}_1}^{-1}$$

is an $(N-1)\times(N-1)$, upper triangular matrix.

Now define

$$\boxed{\mathrm{C}} = \left[\begin{array}{c|ccc} 1 & 0 & \ldots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbf{C}_1 & \\ 0 & & & \end{array}\right].$$

Then $\boxed{\mathrm{C}}$ is also an orthogonal matrix (**Why?**), and

$$\boxed{\mathrm{C}} \cdot \widetilde{\boxed{\mathrm{F}}} \cdot \boxed{\mathrm{C}}^{-1} =$$

is an upper-triangular matrix. But of course,

$$\begin{aligned} \boxed{\mathrm{C}} \cdot \widetilde{\boxed{\mathrm{F}}} \cdot \boxed{\mathrm{C}}^{-1} &= \boxed{\mathrm{C}} \cdot \boxed{\mathrm{B}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{C}}^{-1} \\ &= \left(\boxed{\mathrm{C}} \cdot \boxed{\mathrm{B}}\right) \cdot \boxed{\mathrm{F}} \cdot \left(\boxed{\mathrm{C}} \cdot \boxed{\mathrm{B}}\right)^{-1} \end{aligned}$$

and $\left(\boxed{\mathrm{C}} \cdot \boxed{\mathrm{B}}\right)$ is the product of two orthogonal matrices, therefor itself orthogonal, so this constitutes an **orthogonal upper-triangulation** of $\boxed{\mathrm{F}}$.

□ [Theorem 2]

Theorem 3: *Spectral Theorem*[1] *for Symmetric Matrices*

Let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be a linear transformation, equivalent to multiplication by the matrix $\boxed{\mathrm{F}}$. The following are equivalent:

1. $\mathbb{R}^N$ has an **orthonormal basis** given by *eigenvectors* of f.
2. $\boxed{\mathrm{F}}$ is **orthogonally diagonalizable**.
3. $\boxed{\mathrm{F}}$ is a **symmetric** matrix.

Proof:

Proof of "(1)⟹(2)": Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ is an orthonormal basis of $\mathbb{R}^N$. If $\boxed{\mathrm{B}} = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_N \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$, then we know that the *matrix representation* of f relative to $\mathcal{B}$ is given by

[1] Also known as the "Orthogonal Diagonalization Theorem", or the "Principal Axis Theorem"

$$\boxed{\mathrm{G}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}$$

which is equal to $\boxed{\mathrm{B}}^{t} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}$, since $\boxed{\mathrm{B}}$ is an orthogonal matrix. But if $\mathbf{b}_1, \ldots, \mathbf{b}_N$ are all **eigenvectors** of f, then we know that $\boxed{\mathrm{G}}$ must be diagonal.

Proof of "(2)⟹(1)": Suppose that $\boxed{\mathrm{B}}$ is an orthogonal matrix such that $\boxed{\mathrm{G}} = \boxed{\mathrm{B}}^{t} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}$ is diagonal. Suppose that $\boxed{\mathrm{B}} = \begin{bmatrix} \uparrow & \uparrow & \ldots & \uparrow \\ \mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_N \\ \downarrow & \downarrow & \ldots & \downarrow \end{bmatrix}$, and let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$. Then $\mathcal{B}$ is an orthonormal basis, and $\boxed{\mathrm{G}}$ is the matrix representation of f relative to $\mathcal{B}$. The fact that $\boxed{\mathrm{G}}$ is diagonal means that $\mathbf{b}_1, \ldots, \mathbf{b}_N$ must be eigenvectors of f.

Proof of "(2)⟹(3)": Suppose that $\boxed{\mathrm{B}}$ is an orthogonal matrix such that $\boxed{\mathrm{G}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}$ is diagonal. Then

$$\begin{aligned} \boxed{\mathrm{F}} &= \boxed{\mathrm{B}} \cdot \boxed{\mathrm{G}} \cdot \boxed{\mathrm{B}}^{-1} \\ &= \boxed{\mathrm{B}} \cdot \boxed{\mathrm{G}} \cdot \boxed{\mathrm{B}}^{t} \\ \text{therefore, } \boxed{\mathrm{F}}^{t} &= \boxed{\mathrm{B}} \cdot \boxed{\mathrm{G}}^{t} \cdot \boxed{\mathrm{B}}^{t} \\ &=_{(1)} \boxed{\mathrm{B}} \cdot \boxed{\mathrm{G}} \cdot \boxed{\mathrm{B}}^{t} \\ &= \boxed{\mathrm{F}} \end{aligned}$$

(1) $\boxed{\mathrm{G}}$ is diagonal, therefore symmetric.

hence $\boxed{\mathrm{F}}$ is symmetric.

Proof of "(3)⟹(2)":

Claim 1: The characteristic polynomial of $\boxed{\mathrm{F}}$ factors completely.

Proof: *The proof of this claim involves the use of* **complex numbers**, *and hence, is not covered in this course. A sketch is as follows:*

If $c_{\boxed{F}}(x)$ is the characteristic polynomial of $\boxed{F}$, then we know that $c_{\boxed{F}}(x)$ **factors completely** *over the complex numbers; in other words,*

$$c_{\boxed{F}}(x) = (x - \lambda_1) \cdot (x - \lambda_2) \cdot \ldots \cdot (x - \lambda_N),$$

where $\lambda_1, \ldots, \lambda_N$ are complex numbers. These numbers are then **complex eigenvalues** *of $\boxed{F}$. (It turns out that, for a symmetric matrix, all these eigenvalues will be real, but we don't know this yet).*

The proof which follows can then be carried out using these complex eigenvalues. We can therefor diagonalize $\boxed{F}$ *into the matrix*

$$\boxed{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_N \end{bmatrix}.$$

But $\boxed{F}$ *is a real-valued matrix, and* $\boxed{\Lambda} = \boxed{B}^{-1} \cdot \boxed{\Lambda} \cdot \boxed{B}$ *for some (real-valued) matrix* $\boxed{B}$*; hence,* $\boxed{\Lambda}$ *must also be a real-valued matrix, which means that* $\lambda_1, \dots, \lambda_N$ *must be* **real** *numbers.*

.... □ [Claim 1]

Now, to prove "(3)$\Longrightarrow$(2)", use the Triangulation Theorem to find an an **orthogonal matrix** $\boxed{\mathrm{P}}$ so that

$$\boxed{\mathrm{P}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{P}}^{-1}$$

is **upper triangular**. But $\boxed{\mathrm{P}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{P}}^{-1} = \boxed{\mathrm{P}} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{P}}^{t}$ is also **symmetric** (why?); thus, if it is upper triangular, it must actually be **diagonal**. Hence, this constitutes an orthogonal diagonalization of $\boxed{\mathrm{F}}$, so we're done.

□ [Theorem 3]

Example 4: *(wantonly plagiarised from Nicholson)*

If $\boxed{\mathrm{A}} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$, then $\boxed{\mathrm{A}}$ has characteristic polynomial

$$c_{\boxed{A}}(x) = \det \begin{bmatrix} 1-x & 0 & -1 \\ 0 & 1-x & 2 \\ -1 & 2 & 5-x \end{bmatrix} = x(x-1)(x-6),$$

hence, eigenvalues 0, 1 and 6.

The corresponding eigenvectors (normalized to have all have norm 1) are:

$$\mathbf{b}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \qquad \mathbf{b}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \qquad \mathbf{b}_3 = \frac{1}{\sqrt{30}} \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix},$$

and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is an **orthonormal basis** (check). Hence, if we define

$$\boxed{\mathrm{B}} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{0}{\sqrt{30}} \\ \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & \frac{5}{\sqrt{30}} \end{bmatrix},$$

Then $\boxed{\mathrm{B}}$ is an orthogonal matrix, and

$$\boxed{\mathrm{B}}^t \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}} = \begin{bmatrix} 0 & & \\ & 1 & \\ & & 6 \end{bmatrix}.$$

is a diagonal matrix.