Diagonalization

Prerequisites:

- Eigenvectors and Eigenvalues
- Matrix Similarity

A **diagonal** matrix like

$$\boxed{\mathsf{D}} = \begin{bmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_N \end{bmatrix}$$

is especially easy to work with. One can see immediately that the **eigenvalues** are $a_1, \dots, a_N$, and the corresponding **eigenvectors** are $\mathbf{e}_1, \dots, \mathbf{e}_N$.

Given an arbitrary matrix $\boxed{\mathsf{A}}$, it would be nice if we could show that $\boxed{\mathsf{A}}$ was **similar** to a diagonal matrix like $\boxed{\mathsf{D}}$. This process is called *diagonalization*.

Definition 1: *Diagonalizable, Diagonalization*

Let $\boxed{\mathsf{A}}$ be an $N \times N$ matrix. $\boxed{\mathsf{A}}$ is called **diagonalizable** if it is similar to a diagonal matrix $\boxed{\mathsf{D}}$.

Finding an invertible matrix $\boxed{\mathsf{B}}$ and a diagonal matrix $\boxed{\mathsf{D}}$ so that

$$\boxed{\mathsf{A}} = \boxed{\mathsf{B}} \cdot \boxed{\mathsf{D}} \cdot \boxed{\mathsf{B}}^{-1}$$

is called **diagonalizing** $\boxed{\mathsf{A}}$.

Example 2:

$$\text{Suppose } \boxed{\mathsf{A}} = \begin{bmatrix} -6 & -6 & -28 \\ & 3 & -\frac{5}{2} \\ & & 5 \end{bmatrix}$$

$$\text{and let } \boxed{\mathsf{B}} = \begin{bmatrix} 1 & 2 & 3 \\ & 4 & 5 \\ & & 6 \end{bmatrix} \text{ so that } \boxed{\mathsf{B}}^{-1} = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{12} \\ & \frac{1}{4} & -\frac{5}{24} \\ & & \frac{1}{6} \end{bmatrix}$$

(blank spaces indicate zeroes). Then

$$\boxed{\mathrm{B}} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}}^{-1} = \begin{bmatrix} -6 & -12 & -18 \\ & 12 & 15 \\ & & 30 \end{bmatrix} \cdot \boxed{\mathrm{B}}^{-1} = \begin{bmatrix} -6 & & \\ & 3 & \\ & & 5 \end{bmatrix}.$$

Remark 3: Suppose $\mathbb{V}$ was a finite-dimensional vector space, and $f : \mathbb{V} \longrightarrow \mathbb{V}$ was a linear transformation. Suppose that $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_N\}$ was a basis for $\mathbb{V}$, so that, with respect to $\mathcal{B}$, f had **matrix representation**

$$\begin{bmatrix} a_1 & 0 & 0 & \ldots & 0 \\ 0 & a_2 & 0 & \ldots & 0 \\ 0 & 0 & a_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_N \end{bmatrix}$$

Thus, it is clear $\mathbf{b}_1, \ldots, \mathbf{b}_N$ are all **eigenvectors** of f, with corresponding **eigenvalues** $a_1, \ldots, a_N$. We say that $\mathbb{V}$ has a **basis of eigenvectors of** f.

Theorem 4: *Diagonalization and Bases of Eigenvectors*

Suppose $\mathbb{V}$ is a finite-dimensional vector space, and $f : \mathbb{V} \longrightarrow \mathbb{V}$ is a linear transformation. Suppose that f has **matrix representation** $\boxed{\mathrm{F}}$ with respect to some basis $\mathcal{A}$ of $\mathbb{V}$. Then:

$$\left(\begin{array}{l}\text{There is a basis of } \mathbb{V} \text{ consisting} \\ \text{of } \textbf{eigenvectors} \text{ for } f\end{array}\right) \iff \left(\boxed{\mathrm{F}} \text{ is } \textbf{diagonalizable}\right).$$

Proof:

Proof of "$\Longrightarrow$": Suppose $\mathcal{B}$ is a basis of eigenvectors for f. Let $\boxed{\mathrm{B}}$ be the change-of-basis matrix from $\mathcal{A}$ to $\mathcal{B}$. Thus,

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

is the **matrix representation** of f with respect to $\mathcal{B}$; hence, $\boxed{\mathrm{D}}$ must be a diagonal matrix, hence, $\boxed{\mathrm{F}}$ must be diagonalizable.

Proof of "$\Longleftarrow$": Let $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ be the basis of $\mathbb{V}$ so that $\boxed{\mathrm{F}}$ is the representation of f with respect to $\mathcal{A}$.

Let $\boxed{\mathrm{B}} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1N} \\ b_{21} & b_{22} & \dots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N1} & b_{N2} & \dots & b_{NN} \end{bmatrix}$ be a matrix so that $\boxed{\mathrm{D}} = \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{F}} \cdot \boxed{\mathrm{B}}$ is diagonal[1].

For all $k \in [1..N]$, define $\mathbf{b}_k = \sum_{n=1}^{N} b_{n,k} \cdot \mathbf{a}_n$. It is left as an **exercise** to check that:

- $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_N\}$ is a **basis** for $\mathbb{V}$,
- $\boxed{\mathrm{B}}$ is the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{A}$.
- For all $n \in [1...N]$, $\mathbf{b}_n$ is an **eigenvector** of f.

□ `[Theorem 4]`

Example 5:

$$\begin{aligned} \text{Suppose } \boxed{\mathrm{A}} &= \begin{bmatrix} \frac{3}{4} & \frac{5}{4} \\ \frac{5}{4} & \frac{3}{4} \end{bmatrix}, \\ \text{and } \mathbf{b}_1 &= (1,1) \\ \mathbf{b}_2 &= (-1,1) \end{aligned}$$

$$\text{thus } \boxed{\mathrm{A}} \cdot \mathbf{b}_1 = \begin{bmatrix} \frac{3}{4} & \frac{5}{4} \\ \frac{5}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2.\mathbf{b}_1$$

$$\text{and, similarly, } \boxed{\mathrm{A}} \cdot \mathbf{b}_2 = \frac{1}{2}\mathbf{b}_2$$

Thus, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ is a **basis of eigenvectors** for $\boxed{\mathrm{A}}$; the corresponding eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = \frac{1}{2}$.

Suppose $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$ is the "standard" basis for $\mathbb{R}^2$. Then

$$\mathbf{b}_1 = \mathbf{e}_1 + \mathbf{e}_2 \quad \text{and} \quad \mathbf{b}_2 = -\mathbf{e}_1 + \mathbf{e}_2$$

$$\text{so the change-of-basis matrix from } \mathcal{B} \text{ to } \mathcal{E} \text{ is } \boxed{\mathrm{B}} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

[1] Note that we have reversed the positions of "$\boxed{\mathrm{B}}$" and "$\boxed{\mathrm{B}}^{-1}$" here. Of course, this makes no difference to the definition of diagonalization.

$$\text{while} \quad \mathbf{e}_1 \;=\; \tfrac{1}{2}\mathbf{b}_1 - \tfrac{1}{2}\mathbf{b}_2 \quad \text{and} \quad \mathbf{e}_2 \;=\; \tfrac{1}{2}\mathbf{b}_1 + \tfrac{1}{2}\mathbf{b}_2$$

so the change-of-basis matrix from $\mathcal{E}$ to $\mathcal{B}$ is $\boxed{\mathrm{B}}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}$.

$$\text{Hence } \boxed{\mathrm{B}}^{-1} \cdot \boxed{\mathrm{A}} \cdot \boxed{\mathrm{B}} \;=\; \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} \frac{3}{4} & \frac{5}{4} \\ \frac{5}{4} & \frac{3}{4} \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \;=\; \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

Remark 6: *Consequences of Diagonalizability*
Suppose that $\boxed{\mathrm{A}}$ is a diagonalizable $N \times N$ matrix. Then

- Any matrix **similar** to $\boxed{\mathrm{A}}$ is diagonalizable.
- $\boxed{\mathrm{A}}^{-1}$ is diagonalizable.
- The characteristic polynomial of $\boxed{\mathrm{A}}$ *factors completely*:

$$c_{\boxed{A}}(x) \;=\; (x-\lambda_1)^{N_1} \cdot (x-\lambda_2)^{N_2} \cdot \ldots \cdot (x-\lambda_K)^{N_K}.$$

$$\text{and thus,} \quad \det\left(\boxed{\mathrm{A}}\right) \;=\; \lambda_1^{N_1} \cdot \lambda_2^{N_2} \cdot \ldots \lambda_K^{N_K}.$$

The proofs of these facts are **exercises.**
The following theorem provides useful tools for determining whether a matrix is diagonalizable —ie. whether the eigenvectors form a basis.

Theorem 7: Let $\mathbb{V}$ be an N-dimensional vector space, and $f : \mathbb{V} \longrightarrow \mathbb{V}$ be a linear transformation. Suppose that the **eigenvalues** of f are $\lambda_1, \lambda_2, \ldots, \lambda_K$, having **multiplicities** $n_1, n_2, \ldots, n_K$ and **eigenspaces** $\mathbb{E}_{\lambda_1}, \mathbb{E}_{\lambda_2}, \ldots, \mathbb{E}_{\lambda_K}$, respectively.

1. If $\mathbf{v}_1, \ldots, \mathbf{v}_K$ are eigenvectors of f, with $\mathbf{v}_1 \in \mathbb{E}_{\lambda_1}, \mathbf{v}_1 \in \mathbb{E}_{\lambda_2}, \ldots, \mathbf{v}_1 \in \mathbb{E}_{\lambda_K}$, respectively, then the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K\}$ is **linearly independent**.
2. For all k, $\mathbb{E}_{\lambda_k}$ has a basis $\mathcal{E}_k = \{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$ consisting of λ_k**-eigenvectors** of f.
3. For all k, $\dim[\mathbb{E}_{\lambda_k}] \;\le\; n_k$.

4. The following are equivalent:

   (a) $\dim[\mathbb{E}_{\lambda_1}] \; + \; \dim[\mathbb{E}_{\lambda_2}] \; + \; \ldots \; + \; \dim[\mathbb{E}_{\lambda_k}] \;\; = \;\; \dim[\mathbb{V}]$.

   (b) For all k, $\dim[\mathbb{E}_{\lambda_k}] \; = \; n_k$, and also $n_1 + n_2 + \ldots + n_k = N$.

   (c) $\mathbb{V}$ has a **basis** consisting of eigenvectors of f.

   (d) If $\boxed{\mathsf{F}}$ is any matrix representation of f, then $\boxed{\mathsf{F}}$ is diagonalizable.

Proof:

Proof of Part 1: We will prove, by induction, that, for all $J < K$, the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_K\}$ is linearly independent.

Base Case ($J = 1$): Clearly, the set $\{\mathbf{v}_1\}$ is linearly independent.

Induction: Suppose, inductively, that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{J-1}\}$ is linearly independent, and suppose that $r_1, \ldots, r_J$ are real coefficients so that

$$r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \ldots + r_J\mathbf{v}_J \;\; = \;\; 0 \qquad\qquad (A)$$

We want to show that $r_1, r_2, \ldots, r_J$ must all be equal to zero.
So, first, divide equation (A) by λ_J :

$$0 \;\; = \;\; \frac{r_1}{\lambda_J}\mathbf{v}_1 + \frac{r_2}{\lambda_J}\mathbf{v}_2 + \ldots + \frac{r_J}{\lambda_J}\mathbf{v}_J$$

and then act on it by f:

$$\begin{aligned}
0 \;\; &= \;\; f(0) \\
&= \;\; f\left(\frac{r_1}{\lambda_J}\mathbf{v}_1 + \frac{r_2}{\lambda_J}\mathbf{v}_2 + \ldots + \frac{r_J}{\lambda_J}\mathbf{v}_J\right) \\
&= \;\; \frac{r_1}{\lambda_J}f(\mathbf{v}_1) + \frac{r_2}{\lambda_J}f(\mathbf{v}_2) + \ldots + \frac{r_{J-1}}{\lambda_J}f(\mathbf{v}_{J-1}) + \frac{r_J}{\lambda_J}f(\mathbf{v}_J) \\
&= \;\; \frac{r_1}{\lambda_J}\lambda_1.\mathbf{v}_1 + \frac{r_2}{\lambda_J}\lambda_2.\mathbf{v}_2 + \ldots + \frac{r_{J-1}}{\lambda_J}\lambda_{J-1}.\mathbf{v}_{J-1} \;\; + \;\; r_J\mathbf{v}_J \qquad\qquad (B)
\end{aligned}$$

Now subtract (A) from (B), to get:

$$0 \;\; = \;\; \left(\frac{r_1}{\lambda_J}\lambda_1 - r_1\right)\mathbf{v}_1 + \left(\frac{r_2}{\lambda_J}\lambda_2 - r_2\right).\mathbf{v}_2 + \ldots + \left(\frac{r_{J-1}}{\lambda_J}\lambda_{J-1} - r_{J-1}\right)\mathbf{v}_{J-1}$$

Hence, by the induction hypothesis, we must conclude:

$$\begin{aligned}
\frac{r_1}{\lambda_J}\lambda_1 - r_1 &= 0\\
\frac{r_2}{\lambda_J}\lambda_2 - r_2 &= 0\\
&\vdots\\
\frac{r_{J-1}}{\lambda_J}\lambda_{J-1} - r_{J-1} &= 0
\end{aligned}$$

$$\text{In other words,}\quad \begin{aligned}
\frac{r_1}{\lambda_J} &= \frac{r_1}{\lambda_1}\\
\frac{r_2}{\lambda_J} &= \frac{r_2}{\lambda_2}\\
&\vdots\\
\frac{r_{J-1}}{\lambda_J} &= \frac{r_{J-1}}{\lambda_{J-1}}
\end{aligned}$$

but since λ_J is different from each of $\lambda_1, \ldots, \lambda_{J-1}$, this can only be true if

$$r_1 = r_2 = \ldots = r_{J-1} = 0.$$

Hence, the original equation (A) must read:

$$r_J \mathbf{v}_J = 0$$

which means that r_J must also be zero.

Proof of Part 2: Exercise.

Proof of Part 3: Suppose that $\dim[\mathbb{E}_{\lambda_k}] = d$, and let $\mathcal{A}_k = \{\mathbf{a}_1, \ldots, \mathbf{a}_d\}$ be a basis for $\mathbb{E}_{\lambda_k}$ (consisting of of λ_k-eigenvectors). Recall that we can "complete" $\mathcal{A}_k$ to a basis for all of $\mathbb{V}$, by adding linearly independent vectors $\mathbf{b}_{d+1}, \mathbf{b}_{d+2}, \ldots, \mathbf{b}_N$, so that the set $\mathcal{B} = \{\mathbf{a}_1, \ldots, \mathbf{a}_d,\ \mathbf{b}_{d+1}, \mathbf{b}_{d+2}, \ldots, \mathbf{b}_N\}$ is a basis of $\mathbb{V}$. Let $\widetilde{\boxed{\mathrm{F}}}$ be the **matrix representation** of f relative to $\mathcal{B}$.

Since $\widetilde{\boxed{\mathrm{F}}}$ is **similar** to $\boxed{\mathrm{F}}$, the two matrices have the same characteristic polynomial; hence, all eigenvalues appear with the same multiplicities. So, consider the characteristic polynomial of $\widetilde{\boxed{\mathrm{F}}}$. Since $\mathbf{a}_1, \ldots, \mathbf{a}_d$ are all λ_k-**eigenvectors** of f, it is clear that the matrix $\widetilde{\boxed{\mathrm{F}}}$ must be of the form:

$$\widetilde{\boxed{\mathrm{F}}} = \left[\begin{array}{ccccc|ccccc}
\lambda_k & 0 & 0 & \ldots & 0 & * & * & * & \ldots & * \\
0 & \lambda_k & 0 & \ldots & 0 & * & * & * & \ldots & * \\
0 & 0 & \lambda_k & \ldots & 0 & * & * & * & \ldots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & \lambda_k & * & * & * & \ldots & * \\
\hline
0 & 0 & 0 & \ldots & 0 & * & * & * & \ldots & * \\
0 & 0 & 0 & \ldots & 0 & * & * & * & \ldots & * \\
0 & 0 & 0 & \ldots & 0 & * & * & * & \ldots & * \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 0 & * & * & * & \ldots & * \\
\multicolumn{5}{c}{\underbrace{\hphantom{0000000000}}_{d}} & \multicolumn{5}{c}{\underbrace{\hphantom{0000000000}}_{(N-d)}}
\end{array}\right]$$

Hence, the characteristic polynomial of $\widetilde{\boxed{\mathrm{F}}}$ is of the form $(x-\lambda_k)^d \cdot p(x)$, where $p(x)$ is some polynomial of degree $(N-d)$; hence the **multiplicity** of the eigenvalue λ_k is at *least* d. In other words, $n_k \geq d$.

Proof of Part 4: "$(a) \Longrightarrow (b)$": The characteristic polynomial of $\boxed{\mathrm{F}}$ has degree N (why?), and thus, $n_1 + n_2 + \ldots + n_k \leq N$ (why?) But by **Part 3**, we know that $\dim[\mathbb{E}_{\lambda_k}] \leq n_k$ for all k. Thus,

$$\begin{aligned}
\dim[\mathbb{E}_{\lambda_1}] \;+\; \dim[\mathbb{E}_{\lambda_2}] \;+\; \ldots \;+\; \dim[\mathbb{E}_{\lambda_k}] \;&\leq\; n_1 + n_2 + \ldots + n_k, \\
&\leq\; N \\
&=\; \dim[\mathbb{V}]
\end{aligned}$$

But if (a) is true, then the first and last expression are *equal.* The two sides of the equation can only be equal if $\dim[\mathbb{E}_{\lambda_k}] = n_k$ for all k, and also $n_1 + n_2 + \ldots + n_k = N$.

"$(b) \Longrightarrow (a)$" This is immediate.

"$(a) \iff (c)$" This follows from **Part 2**

"$(c) \iff (d)$" This is just a restatement of the previous theorem.

□ [Theorem 7]

Corollary 8: Let $f : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ be the linear transformation corresponding multiplication by matrix $\boxed{\mathrm{F}}$.

If f has N *different* eigenvalues, then $\boxed{\mathrm{F}}$ is diagonalizable.

Proof: Exercise

□ [Corollary 8]