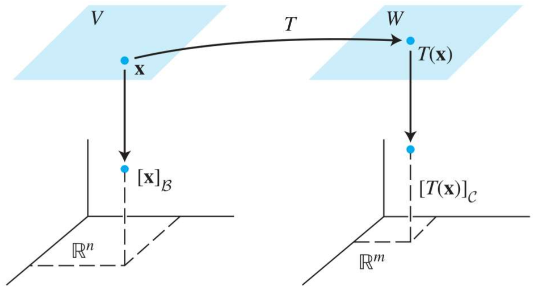
Eigenvectors and linear transformations

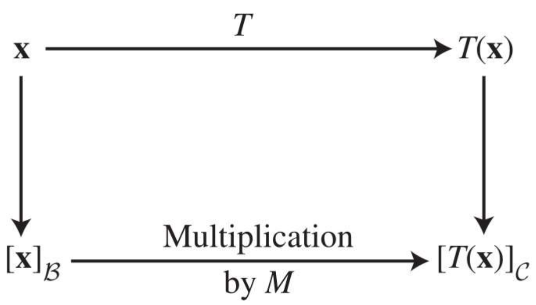
- Let $T: V \to W$ be a linear transformation with $\dim(V) = n, \dim(W) = m$, let $\mathscr{B}$ and $\mathscr{C}$ be bases for V and W respectively. In terms of the coordinates with respect to the bases, T is described by a matrix M.



$$[\,T(\mathbf{x})\,]_{\mathcal{C}} = M[\,\mathbf{x}\,]_{\mathcal{B}}$$

$$M = \left[\,[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} \quad [\,T(\mathbf{b}_2)\,]_{\mathcal{C}} \quad \cdots \quad [\,T(\mathbf{b}_n)\,]_{\mathcal{C}}\,\right]$$

- The matrix M is a matrix representation of T, called the matrix for T relative to the bases $\mathscr{B}$ and $\mathscr{C}$.



- In case that $V = W$, and $\mathscr{B} = \mathscr{C}$, the matrix M is called the matrix for T relative to $\mathscr{B}$.

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication}", "\text{by } {[T]_{\mathcal{B}}}"'] & {[T(\mathbf{x})]_{\mathcal{B}}}
\end{tikzcd}$$

- Let A be a $n \times n$ matrix, it defines a linear transformation $f : R^n \to R^n$ sending x to Ax. Let $\mathscr{B} = \{b_1, \cdots, b_n\}$ be another basis of R^n, let $P = [b_1 \ \cdots \ b_n]$, then the transition matrix $P_{\mathscr{E} \to \mathscr{B}}$ equals P^{-1}. The matrix of f relative to $\mathscr{B}$ equals $C = P^{-1}AP$. In other words, $A = PCP^{-1}$, which can be expressed as the commutative diagram:

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "\text{Multiplication}", "\text{by } A"'] \arrow[d, "\text{Multiplication}"', "\text{by } P^{-1}"] & A\mathbf{x} \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplication}", "\text{by } C"'] & {[A\mathbf{x}]_{\mathcal{B}}} \arrow[u, "\text{Multiplication}"', "\text{by } P"]
\end{tikzcd}$$

- Theorem

Diagonal Matrix Representation

Suppose $A = PDP^{-1}$, where D is a diagonal $n \times n$ matrix. If $\mathcal{B}$ is the basis for $\mathbb{R}^n$ formed from the columns of P, then D is the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Diagonalization of complex matrices

- Recall that the matrix $A = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$ can not be diagonalized as a real matrix, i.e. there is no real matrices B and D with B invertible and D diagonal such that $B^{-1}AB = D$.
- The reason is that its characteristic equation $\det(A - \lambda I) = \lambda^2 + 1$ has no real roots.
- But it does have complex roots, this suggests that we should consider the complex matrices B and D.

- Example: Diagonalize $A = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$ if possible.

Solution The characteristic equation of A is

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$$

It has roots $\lambda = \pm i$.

We need to calculate its eigenvectors:

1) For $\lambda = i$, solve the system

$$\begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow x_1 + i x_2 = 0 \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = c \cdot \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

2) For $\lambda = -i$, need to solve the system

$$\begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow x_1 - i x_2 = 0 \Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = c' \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

These two eigenvector equation can be written together as

$$A \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}\begin{bmatrix} i & \\ & -i \end{bmatrix}$$

$$\Rightarrow A = \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}\begin{bmatrix} i & \\ & -i \end{bmatrix}\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}^{-1}$$ is diagonalizable.

$\boxed{S}$

- Let A be an $n \times n$ complex matrix. Following the same procedure as before, we can determine whether A is diagonalizable.
- The difference with the real case is that the characteristic equation will always have n roots if counted with the multiplicity, i.e.

$$\det(A - \lambda I) = \prod_{i=1}^{r} (\lambda_i - \lambda)^{m_i}, \quad \text{with } \sum_{i=1}^{r} m_i = n\,.$$

- This is due to the fundamental theorem of algebra: Any polynomial of degree n has n complex roots if counted with multiplicity.

- We use the same notation as before: Let m_λ be the multiplicity of a root λ of the characteristic equation $\det(A - \lambda I) = 0$, let d_λ be the dimension of the eigenspace V_λ.
- Theorem: Let A be a $n \times n$ complex matrix, then $m_\lambda \geq d_\lambda$ for any complex eigenvalue λ. The matrix A is diagonalizable if and only if $m_\lambda = d_\lambda$ for all the complex roots of the equation $\det(A - \lambda I) = 0$.
- Corollary: If the roots of $\det(A - \lambda I) = 0$ appear with multiplicity 1, then A must be diagonalizable.

- The theorem has applications to the real matrices as well
- Let A be a 2×2 real matrix with a complex but non-real eigenvalue λ, let $v \in \mathbb{C}^2$ be the eigenvector, i.e. $Av = \lambda v$. Take complex conjugate, we get $A\bar{v} = \bar{\lambda}\bar{v}$. Hence A has another eigenvalue $\bar{\lambda}$ with eigenvector $\bar{v}$.
- We write this relation as

$$A[v\ \bar{v}] = [v\ \bar{v}]\begin{bmatrix} \lambda & \\ & \bar{\lambda} \end{bmatrix}.$$

- To go back to the real matrices, consider the real part $\mathrm{Re}(v)$ and imaginary part $\mathrm{Im}(v)$ of v:

$$v = \mathrm{Re}(v) + i\mathrm{Im}(v) \quad \text{with } \mathrm{Re}(v), \mathrm{Im}(v) \in \mathbb{R}^2.$$

- Take complex conjugate, get $\bar{v} = \mathrm{Re}(v) - i\mathrm{Im}(v)$.
- The above two equations can be written as

$$[v\ \bar{v}] = [\mathrm{Re}(v)\ \mathrm{Im}(v)]\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}.$$

- Plug it into the eigenvector equation, we get

$$A \cdot [\mathrm{Re}(v)\ \mathrm{Im}(v)]\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} = [\mathrm{Re}(v)\ \mathrm{Im}(v)]\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \cdot \begin{bmatrix} \lambda & \\ & \bar{\lambda} \end{bmatrix}.$$

- Let $\lambda = a - bi$. Move the factor $\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ to the right, the above equation becomes

$$A \cdot [\mathrm{Re}(v)\ \mathrm{Im}(v)] = [\mathrm{Re}(v)\ \mathrm{Im}(v)] \cdot \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

- The above discussion can be summarized as the theorem:

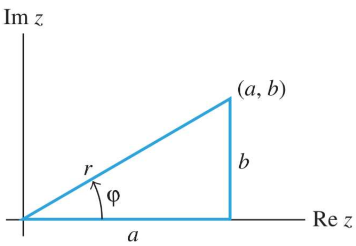
Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ $(b \neq 0)$ and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$. Then

$$A = PCP^{-1}, \quad \text{where} \quad P = [\operatorname{Re}\mathbf{v} \quad \operatorname{Im}\mathbf{v}] \quad \text{and} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

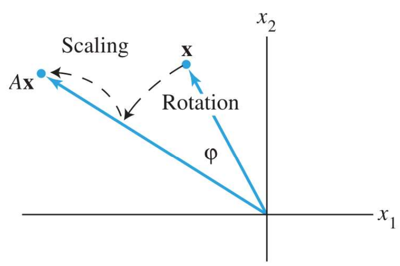
- Geometric meaning: Let $r = |\lambda| = \sqrt{a^2 + b^2}$, then $\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = 1$

and we can set $a/r = \cos\varphi, b/r = \sin\varphi$. Then

$$C = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$



- Recall that the matrix $\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ acts on $\mathbb{R}^2$ as rotation around the origin by the angle φ.
- The theorem states that we can find a basis of $\mathbb{R}^2$with respect to which the action of A is the composition of a rotation by angle φ and a scaling by $|\lambda|$.

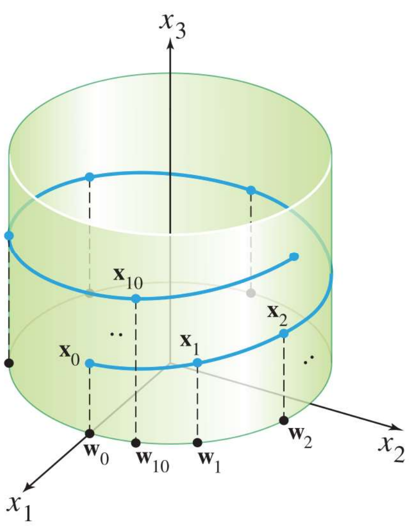


- With the same idea, we can prove:
- Theorem: Let A be a real 3×3 matrix with a complex eigenvalue $\lambda = a - bi, b \neq 0$, then there exists an invertible matrix B such that

$$A = B \begin{bmatrix} a & -b & \\ b & a & \\ & & c \end{bmatrix} B^{-1}, \quad \text{with } c \in \mathbb{R}.$$

- Geometrically, this means that we can find a basis of $\mathbb{R}^3$ such that A acts on the new x_3-ax by scaling and on the new (x_1, x_2)-plane by a composition of scaling and rotation.

- Iterating the action of A on a point $x_0 \in \mathbb{R}^3$, get a picture like:



Jordan form

- It is not true that any matrix can be diagonalized, even as a complex matrix. Nonetheless, we have the theorem:

(Jordan form) If A has s independent eigenvectors, it is similar to a matrix J that has s Jordan blocks $J_1 \ldots, J_s$ on its diagonal. Some matrix B puts A into Jordan form:

Jordan form
$$B^{-1}AB = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix} = J. \tag{3}$$

Each block J_i has one eigenvalue λ_i, one eigenvector, and 1's just above the diagonal:

Jordan block
$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & & \lambda_i \end{bmatrix}. \tag{4}$$

Matrices are similar if they share the same Jordan form J—not otherwise.