Similar matrices

The matrix of a linear map f on V is not unique, since it depends on the basis. Matrices of the same map, but w.r.t. different bases shall have the same eigenvalues.

$$[f]_{XX} = [id]_{YX}[f]_{YY}[id]_{XY}$$

$$\begin{aligned}
[f(u)]_X &= [f]_{XX}[u]_X \\
&= [id]_{YX}[f(u)]_Y = [id]_{YX}[f]_{YY}[u]_Y \\
&= [id]_{YX}[f]_{YY}[id]_{XY}[u]_X
\end{aligned}$$

Note that $[id]_{YX} = [id]_{XY}^{-1}$

Similar matrices

The matrix of a linear map f on V is not unique, since it depends on the basis. Matrices of the same map, but w.r.t. different bases shall have the same eigenvalues.

$$[f]_{XX} = [id]_{YX}[f]_{YY}[id]_{XY}$$

Definition Matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{K}^{n\times n}$ are *similar* if there exists a regular matrix $\boldsymbol{R}$ such that $\boldsymbol{A} = \boldsymbol{R}^{-1}\boldsymbol{B}\boldsymbol{R}$.

Observation: If $\boldsymbol{A}$ is similar to $\boldsymbol{B}$, i.e. $\boldsymbol{B} = \boldsymbol{R}\boldsymbol{A}\boldsymbol{R}^{-1}$,
and an eigenvalue λ corresponds to an eigenvector $\boldsymbol{x}$ in $\boldsymbol{A}$,
then λ is also an eigenvalue of $\boldsymbol{B}$ and corresponds here to $\boldsymbol{R}\boldsymbol{x}$.

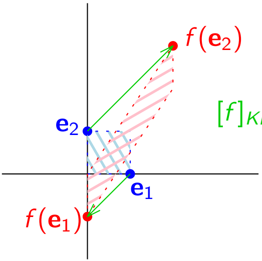
Proof: For $\boldsymbol{y} = \boldsymbol{R}\boldsymbol{x}$ holds: $\boldsymbol{B}\boldsymbol{y} = \boldsymbol{R}\boldsymbol{A}\boldsymbol{R}^{-1}\boldsymbol{R}\boldsymbol{x} = \boldsymbol{R}\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{R}\boldsymbol{x} = \lambda\boldsymbol{y}$.

Observation: If $\boldsymbol{B} = \boldsymbol{R}\boldsymbol{A}\boldsymbol{R}^{-1}$ then $p_{\boldsymbol{B}}(t) = p_{\boldsymbol{A}}(t)$.

Proof: $p_{\boldsymbol{B}}(t) = \det(\boldsymbol{B} - t\boldsymbol{I}) = \det(\boldsymbol{R}\boldsymbol{A}\boldsymbol{R}^{-1} - \boldsymbol{R}(t\boldsymbol{I})\boldsymbol{R}^{-1}) =$
$\det(\boldsymbol{R}(\boldsymbol{A} - t\boldsymbol{I})\boldsymbol{R}^{-1}) = \det(\boldsymbol{R})\det(\boldsymbol{A} - t\boldsymbol{I})\det(\boldsymbol{R}^{-1}) = p_{\boldsymbol{A}}(t)$

Example — a linear map in the plane

Does the following linear map have a better description?



$$[f]_{KK} = \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix}$$

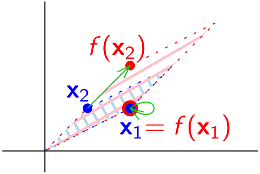
Characteristic polynomial:

$$p_{[f]_{KK}}(t) = \begin{vmatrix} -t & 2 \\ -1 & 3-t \end{vmatrix} = t^2 - 3t + 2 = (t-1)(t-2)$$

The eigenvalue $\lambda_1 = 1$ has eigenvector $\boldsymbol{x}_1 = (2,1)^T$, and
the eigenvalue $\lambda_2 = 2$ has eigenvector $\boldsymbol{x}_2 = (1,1)^T$.

With respect to the new basis $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2\} = \{(2,1)^T, (1,1)^T\}$ the matrix of *the same* linear map f is *diagonal*:

$$[f]_{XX} = [id]_{KX}[f]_{KK}[id]_{XK} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$



$$[f]_{XX} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Less formally: the plane is fixed along the line through $\boldsymbol{x}_1$ and twice stretched along the line through $\boldsymbol{x}_2$.

Observe that the eigenvalues and eigenvectors are preserved.

Algebraic and geometric multiplicity

Observation: If a basis X contains an eigenvector $\boldsymbol{x}$ of f, then the coordinate corresponding to $\boldsymbol{x}$ is scaled by λ under f.
In matrix terms: $[f]_{XX}$ contains in the column corresponding to $\boldsymbol{x}$ only λ at the diagonal and otherwise zeroes.

Proof: When an eigenvector $\boldsymbol{u}$ is the i-th vector of a basis X, then the i-th column of $[f]_{XX}$ is $[f(\boldsymbol{u})]_X = [\lambda \boldsymbol{u}]_X = \lambda [\boldsymbol{u}]_X = \lambda \boldsymbol{e}^i$.

Algebraic and geometric multiplicity

Observation: If a basis X contains an eigenvector $\boldsymbol{x}$ of f, then the coordinate corresponding to $\boldsymbol{x}$ is scaled by λ under f.
In matrix terms: $[f]_{XX}$ contains in the column corresponding to $\boldsymbol{x}$ only λ at the diagonal and otherwise zeroes.

Theorem: The geometric multiplicity of an eigenvalue λ of a matrix $\boldsymbol{A}$ is smaller or equal to its algebraic multiplicity.

Proof: View $\boldsymbol{A} \in \mathbb{K}^{n \times n}$ as the matrix of a linear map $f : \mathbb{K}^n \to \mathbb{K}^n$ w.r.t. the standard basis K, i.e. $\boldsymbol{A} = [f]_{K,K}$.
Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ be a basis of the space of eigenvectors of λ, i.e. k is its geometric multiplicity.

Extend this basis to a basis X of $\mathbb{K}^n$.
Then $[f]_{X,X} = [id]_{X,K}^{-1} \boldsymbol{A} [id]_{X,K}$ is similar to $\boldsymbol{A}$. Also $[f]_{X,X}$ has on the first k columns λ at the diagonal and otherwise zeroes.

Hence $(\lambda - t)^k$ divides $p_{[f]_{X,X}}(t)$. Since $\boldsymbol{A}$ and $[f]_{X,X}$ have equal characteristic polynomials, λ has algebraic multiplicity at least k.

Example

$$\boldsymbol{A} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}$$

$p_{\boldsymbol{A}}(t) = -t^3 + 5t^2 - 8x + 4 = (t-2)^2(t-1)$
eigenvalues are: 2 of *algebraic multiplicity* 2 and 1 of algebraic multiplicity 1.

$$\boldsymbol{A} - 2\boldsymbol{I}_3 = \begin{pmatrix} -3 & 7 & -5 \\ -2 & 5 & -4 \\ -1 & 3 & -3 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

The eigenvalue 2 has in $\boldsymbol{A}$ *geometric multiplicity* only 1.
We extend the eigenvector $(3,2,1)^T$ for 2 to a basis X,
e.g. $X = \{(3,2,1)^T, (2,2,1)^T, (1,1,1)^T\}$.

The matrix $\boldsymbol{A}$ is similar to $[id]_{X,K}^{-1}\boldsymbol{A}[id]_{X,K} =$

$$= \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Example

$$\boldsymbol{A} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}$$

$p_{\boldsymbol{A}}(t) = -t^3 + 5t^2 - 8x + 4 = (t-2)^2(t-1)$
eigenvalues are: 2 of *algebraic multiplicity* 2 and 1 of algebraic multiplicity 1.

$$\boldsymbol{A} - 2\boldsymbol{I}_3 = \begin{pmatrix} -3 & 7 & -5 \\ -2 & 5 & -4 \\ -1 & 3 & -3 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

The eigenvalue 2 has in $\boldsymbol{A}$ *geometric multiplicity* only 1.

Compare with,

$$\boldsymbol{B} = \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & -2 \\ 0 & 1 & 0 \end{pmatrix}$$

has the same characteristic polynomial
$p_{\boldsymbol{B}}(t) = -t^3 + 5t^2 - 8x + 4 = (t-2)^2(t-1)$
and the same eigenvalues, i.e. 2 of algebraic multiplicity 2 and 1 of algebraic multiplicity 1.

$$\boldsymbol{B} - 2\boldsymbol{I}_3 = \begin{pmatrix} 0 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 1 & -2 \end{pmatrix} \sim\sim \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

the eigenvalue 2 has in $\boldsymbol{B}$ geometric multiplicity 2.

Example

$$\boldsymbol{A} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}$$

$p_{\boldsymbol{A}}(t) = -t^3 + 5t^2 - 8x + 4 = (t-2)^2(t-1)$
eigenvalues are: 2 of *algebraic multiplicity* 2 and 1 of algebraic multiplicity 1.

$$\boldsymbol{A} - 2\boldsymbol{I}_3 = \begin{pmatrix} -3 & 7 & -5 \\ -2 & 5 & -4 \\ -1 & 3 & -3 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

The eigenvalue 2 has in $\boldsymbol{A}$ *geometric multiplicity* only 1.

Compare with,

$$\boldsymbol{B} = \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & -2 \\ 0 & 1 & 0 \end{pmatrix}$$

has the same characteristic polynomial
$p_{\boldsymbol{B}}(t) = -t^3 + 5t^2 - 8x + 4 = (t-2)^2(t-1)$
and the same eigenvalues, i.e. 2 of algebraic multiplicity 2 and 1 of algebraic multiplicity 1.

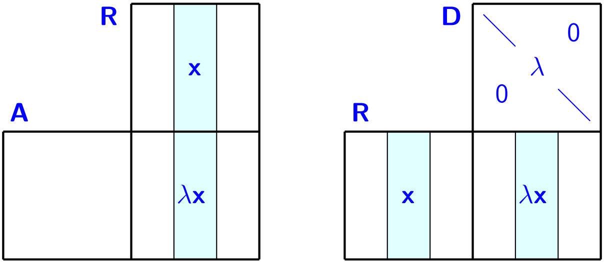
W.r.t. (by coincidence the same) basis X we get $[id]_{X,K}^{-1}\boldsymbol{B}[id]_{X,K} =$

$$= \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 1 & -2 \\ 0 & 3 & -2 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Diagonalization

Observation: A matrix $\boldsymbol{A} \in \mathbb{K}^{n \times n}$ is similar to a diagonal matrix if and only if $\mathbb{K}^n$ has a basis consisting of eigenvectors of $\boldsymbol{A}$.

Proof: $\boldsymbol{AR} = \boldsymbol{RD}$ with diagonal matrix $\boldsymbol{D}$, iff for every i there exists a vector $\boldsymbol{x}$ (the i-th column of $\boldsymbol{R}$) such that $\boldsymbol{Ax} = \lambda \boldsymbol{x} = d_{ii}\boldsymbol{x}$.



$$\mathbf{A} = \mathbf{RDR}^{-1} \quad \Longleftrightarrow \quad \mathbf{AR} = \mathbf{RD} \quad \Longleftrightarrow \quad \mathbf{R}^{-1}\mathbf{AR} = \mathbf{D}$$

Diagonalization

Observation: A matrix $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ is similar to a diagonal matrix if and only if $\mathbb{K}^n$ has a basis consisting of eigenvectors of $\boldsymbol{A}$.

Proof: $\boldsymbol{AR} = \boldsymbol{RD}$ with diagonal matrix $\boldsymbol{D}$, iff for every i there exists a vector $\boldsymbol{x}$ (the i-th column of $\boldsymbol{R}$) such that $\boldsymbol{Ax} = \lambda \boldsymbol{x} = d_{ii}\boldsymbol{x}$.

Definition: A matrix similar to a diagonal matrix is *diagonalizable*.

Corollary: If a square matrix of order n has n distinct eigenvalues, then it is diagonalizable.

Corollary: When $p_{\boldsymbol{A}}(t) = \prod_i (t-\lambda_i)^{r_i}$, then:

$$\boldsymbol{A} \text{ is diagonalizable} \quad \Longleftrightarrow \quad \dim(\operatorname{Ker}(\boldsymbol{A} - \lambda_i \boldsymbol{I})) = r_i$$

Corollary: If $\boldsymbol{A} = \boldsymbol{R}^{-1}\boldsymbol{DR}$, then for any $k : \boldsymbol{A}^k = \boldsymbol{R}^{-1}\boldsymbol{D}^k\boldsymbol{R}$.
$\boldsymbol{A}^k = (\boldsymbol{R}^{-1}\boldsymbol{DR})^k = \boldsymbol{R}^{-1}\boldsymbol{D}\boldsymbol{R}\boldsymbol{R}^{-1}\boldsymbol{D}\boldsymbol{R}\boldsymbol{R}^{-1}\cdots\boldsymbol{R}^{-1}\boldsymbol{DR} = \boldsymbol{R}^{-1}\boldsymbol{D}^k\boldsymbol{R}$.

Jordan normal form

Example: The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable in any field.

Proof: It has eigenvalue 1 of multiplicity two, hence could only be similar to $\boldsymbol{I}_2$. But for any regular $\boldsymbol{R}$: $\boldsymbol{R}^{-1}\boldsymbol{I}_2\boldsymbol{R} = \boldsymbol{I}_2 \neq \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Jordan normal form

Example: The matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is not diagonalizable in any field.

Definition: A *Jordan block* is a square matrix of the form

$$\boldsymbol{J}_\lambda = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

Theorem: Every square complex matrix $\boldsymbol{A}$ is similar to a matrix $\boldsymbol{J}$ in the so called *Jordan normal form*

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_{\lambda_1} & & \\ & \ddots & \\ & & \boldsymbol{J}_{\lambda_k} \end{pmatrix}$$

Each Jordan block $\boldsymbol{J}_{\lambda_i}$ corresponds to an eigenvalue λ_i of $\boldsymbol{A}$. A λ_i may yield several Jordan blocks, indeed of various sizes.

Fact: For each λ, the number of blocks and their sizes are uniquely determined by $\boldsymbol{A}$. Hence the Jordan normal form of $\boldsymbol{A}$ is unique upto a permutation of the Jordan blocks on the diagonal.

Observation: A diagonalizable matrix has Jordan blocks of size one.

Generalized eigenvectors

When $\boldsymbol{A}$ is diagonalizable, i.e. $\boldsymbol{AR} = \boldsymbol{RD}$,

then the columns of $\boldsymbol{R}$ are eigenvectors of $\boldsymbol{A}$.

What can we say about matrices that are not diagonalizable?

Proposition: Let $\boldsymbol{AR} = \boldsymbol{RJ}_\lambda$.

If $\boldsymbol{x}_i$ is the i-th column of $\boldsymbol{R}$, then it satisfies $(\boldsymbol{A} - \lambda \boldsymbol{I})^i \boldsymbol{x}_i = \mathbf{0}$.

Proof:

$$
\begin{array}{cccc|cccc}
 & & & & \lambda & 1 & & \\
 & \boldsymbol{RJ}_\lambda & & & & \lambda & \dots & 1 \\
 & & & & & & & \lambda \\
\hline
\boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n & \lambda \boldsymbol{x}_1 & \boldsymbol{x}_1 + \lambda \boldsymbol{x}_2 & \dots & \boldsymbol{x}_{n-1} + \lambda \boldsymbol{x}_n
\end{array}
$$

$$
\begin{array}{lclcl}
\boldsymbol{A}\boldsymbol{x}_1 = \lambda \boldsymbol{x}_1 & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_1 = \mathbf{0} & & \\
\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{x}_1 + \lambda \boldsymbol{x}_2 & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_2 = \boldsymbol{x}_1 & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})^2 \boldsymbol{x}_2 = \mathbf{0} \\
 & \vdots & & & \\
\boldsymbol{A}\boldsymbol{x}_n = \boldsymbol{x}_{n-1} + \lambda \boldsymbol{x}_n & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x}_n = \boldsymbol{x}_{n-1} & \Rightarrow & (\boldsymbol{A} - \lambda \boldsymbol{I})^n \boldsymbol{x}_n = \mathbf{0}
\end{array}
$$

Generalized eigenvectors

When $\boldsymbol{A}$ is diagonalizable, i.e. $\boldsymbol{AR} = \boldsymbol{RD}$,
then the columns of $\boldsymbol{R}$ are eigenvectors of $\boldsymbol{A}$.

What can we say about matrices that are not diagonalizable?

Proposition: Let $\boldsymbol{AR} = \boldsymbol{RJ}_\lambda$.
If $\boldsymbol{x}_i$ is the i-th column of $\boldsymbol{R}$, then it satisfies $(\boldsymbol{A} - \lambda\boldsymbol{I})^i\boldsymbol{x}_i = \mathbf{0}$.

Definition: *Generalized eigenvector* of a matrix $\boldsymbol{A}$ for an eigenvalue λ is any vector $\boldsymbol{x}$ satisfying $(\boldsymbol{A} - \lambda\boldsymbol{I})^k\boldsymbol{x} = \mathbf{0}$ for some $k \in \mathbb{N}$.

Can be ordered into *chains* $\dots, \boldsymbol{x}_2, \boldsymbol{x}_1, \mathbf{0}$, where $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x}_i = \boldsymbol{x}_{i-1}$.

Analogously, for a linear map f we get $f(\boldsymbol{x}_i) - \lambda\boldsymbol{x}_i = \boldsymbol{x}_{i-1}$.

In another notation: $\boldsymbol{x} \in \ker((\boldsymbol{A} - \lambda\boldsymbol{I})^k)$, or $\boldsymbol{x} \in \ker((f - \lambda id)^k)$.

Theorem: (equivalent version of Jordan's normal form theorem)
Each finitely generated space V over $\mathbb{C}$ and linear $f : V \to V$ has a basis from chains of generalized eigenvectors of the map f.

Note: Also holds for any $\mathbb{K}$, when eigenvalues have algebraic multiplicity $\dim(V)$, i.e. if $p_{[f]_{X,X}}(t)$ decomposes into linear terms.

Example

The matrix $\boldsymbol{A} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}$ is similar to a matrix in the Jordan normal form with two blocks $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, because

$$\boldsymbol{AR} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{RJ}$$

$(3,2,1)^T$ is an eigenvector for 2, i.e. $(\boldsymbol{A} - 2\boldsymbol{I}_3)(3,2,1)^T = \mathbf{0}$ and
$(1,1,1)^T$ is an eigenvector for 1, i.e. $(\boldsymbol{A} - 1\boldsymbol{I}_3)(1,1,1)^T = \mathbf{0}$.

The middle column of the matrix $\boldsymbol{R}$ however satisfies
$\boldsymbol{A} \cdot (2,2,1)^T = (3,2,1)^T + 2 \cdot (2,2,1)^T \quad \Longrightarrow$
$(\boldsymbol{A} - 2\boldsymbol{I}_3)\ (2,2,1)^T = (3,2,1)^T \quad \Longrightarrow$
$(\boldsymbol{A} - 2\boldsymbol{I}_3)^2 (2,2,1)^T = (\boldsymbol{A} - 2\boldsymbol{I}_3)(3,2,1)^T = \mathbf{0}$.

Proof of the theorem — Part 1

By induction on $\dim(V)$. For each eigenvalue λ we introduce the map $g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x}$. Fix some eigenvalue λ.

Consider $W = g_\lambda(V)$, the range of g_λ.

W is a subset of V, because $\forall \boldsymbol{x} \in V : g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x} \in V$.

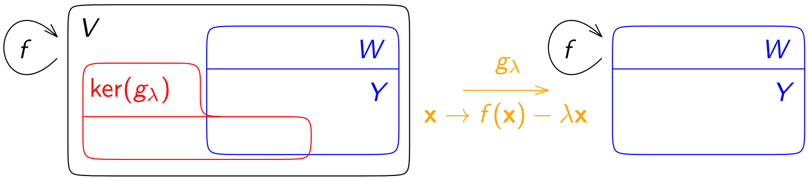
W is a subspace because for $\boldsymbol{u}, \boldsymbol{v} \in W, \alpha \in \mathbb{C}$ exist $\boldsymbol{x}, \boldsymbol{y} \in V$
s.t. $\boldsymbol{u} = g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x}$ and $\boldsymbol{v} = g_\lambda(\boldsymbol{y}) = f(\boldsymbol{y}) - \lambda\boldsymbol{y}$ and:
$\boldsymbol{u} + \boldsymbol{v} = g_\lambda(\boldsymbol{x}) + g_\lambda(\boldsymbol{y}) = f(\boldsymbol{x} + \boldsymbol{y}) - \lambda(\boldsymbol{x} + \boldsymbol{y}) = g_\lambda(\boldsymbol{x} + \boldsymbol{y}) \in W$.
$\alpha\boldsymbol{u} = \alpha g_\lambda(\boldsymbol{x}) = \alpha(f(\boldsymbol{x}) - \lambda\boldsymbol{x}) = f(\alpha\boldsymbol{x}) - \lambda(\alpha\boldsymbol{x}) = g_\lambda(\alpha\boldsymbol{x}) \in W$.

Next, $\dim(W) < \dim(V)$ because the eigenvector $\boldsymbol{u}$ for λ satisfies $g_\lambda(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda\boldsymbol{x} = \mathbf{0}$, i.e. $\dim(\ker(g_\lambda)) \geq 1$ and thus
$\dim(V) = \dim(g_\lambda(V)) + \dim(\ker(g_\lambda)) = \dim(W) + \dim(\ker(g_\lambda))$.

The mar f can be restricted to W, because for $g_\lambda(\boldsymbol{x}) \in W$ we have
$f(g_\lambda(\boldsymbol{x})) = f(f(\boldsymbol{x}) - \lambda\boldsymbol{x}) = f(f(\boldsymbol{x})) - \lambda f(\boldsymbol{x}) = g_\lambda(f(\boldsymbol{x})) \in W$.

According to the inductive hypothesis for f and W, the subspace W has a basis Y from chains of generalized eigenvectors of f.

Example for the first part of the proof



For $[f]_{K,K} = \begin{pmatrix} -1 & 7 & -5 \\ -2 & 7 & -4 \\ -1 & 3 & -1 \end{pmatrix}a\lambda = 2$ is $[g_2]_{K,K} = \begin{pmatrix} -3 & 7 & -5 \\ -2 & 5 & -4 \\ -1 & 3 & -3 \end{pmatrix} \sim\sim \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$

$Z = \{(3,2,1)^T\}$ is a basis of $\ker(g_2)$ so $\dim(W) = 3 - 1 = 2$.

When we extend Z by $\boldsymbol{e}^1, \boldsymbol{e}^2$ to a basis of V, we get $\{g_2(\boldsymbol{e}^1), g_2(\boldsymbol{e}^2)\} = \{(-3,-2,-1)^T, (7,5,3)^T\}$ as a basis of W.

Note that $W \cap \ker(g_2) \neq \emptyset$. This intersection has dimension 1.

There are two chains that form the basis Y of the subspace W: the first is $(3,2,1)^T$ for $\lambda = 2$ and the next is $(1,1,1)^T$ for $\lambda = 1$. (Both have length one, so they contain "ordinary" eigenvectors.)

Proof of theorem — Part 2

Denote $d = \dim(\ker(g_\lambda))$ and $d' = \dim(\ker(g_\lambda) \cap W)$.

Arrange the basis Y into r strings so that the first d' corresponds to λ and others correspond to the other eigenvalues $\lambda', \ldots, \lambda'^{\cdots\prime}$:

$$
\begin{array}{cccccccccc}
\boldsymbol{y}^1_{k_1} & \xrightarrow{g_\lambda} & \cdots\cdots\cdots & & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_2 & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
 & \boldsymbol{y}^2_{k_2} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^2_2 & \xrightarrow{g_\lambda} & \boldsymbol{y}^2_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
 & & & & & & & & & \vdots \\
 & & & \boldsymbol{y}^{d'}_{k_{d'}} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^{d'}_1 & \xrightarrow{g_\lambda} & \mathbf{0} \\
 & & & \boldsymbol{y}^{d'+1}_{k_{d'+1}} & \xrightarrow{g_{\lambda'}} & \cdots & \xrightarrow{g_{\lambda'}} & \boldsymbol{y}^{d'+1}_1 & \xrightarrow{g_{\lambda'}} & \mathbf{0} \\
 & & & & & & & & & \vdots \\
 & & & & & & \cdots & \boldsymbol{y}^r_1 & \xrightarrow{g_{\lambda'^{\cdots\prime}}} & \mathbf{0}
\end{array}
$$

As chains of Y are in W, we can extend each of the first d' chains by some $\boldsymbol{x}^i \in V$ so that $g_\lambda(\boldsymbol{x}^i) = \boldsymbol{y}^i_{k_i}$ for $i \in \{1, \ldots, d'\}$.

The vectors $\boldsymbol{y}^1_1, \ldots, \boldsymbol{y}^{d'}_1$ form the basis of the space $\ker(g_\lambda) \cap W$.

Complete them by $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^{d-d'}$ to a basis of $\ker(g_\lambda)$ (other than Z) and get $d - d'$ new chains of length 1 formed by $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^{d-d'}$.

That yields chains

$$
\begin{array}{ccccccccccc}
\boldsymbol{x}^1 & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_{k_1} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_2 & \xrightarrow{g_\lambda} & \boldsymbol{y}^1_1 & \xrightarrow{g_\lambda} & \mathbf{0}\\
 & & & & & & & & & & \vdots\\
 & & \boldsymbol{x}^{d'} & \xrightarrow{g_\lambda} & \boldsymbol{y}^{d'}_{k_{d'}} & \xrightarrow{g_\lambda} & \cdots & \xrightarrow{g_\lambda} & \boldsymbol{y}^{d'}_1 & \xrightarrow{g_\lambda} & \mathbf{0}\\
 & & & & \boldsymbol{y}^{d'+1}_{k_{d'+1}} & \xrightarrow{g_{\lambda'}} & \cdots & \xrightarrow{g_{\lambda'}} & \boldsymbol{y}^{d'+1}_1 & \xrightarrow{g_{\lambda'}} & \mathbf{0}\\
 & & & & & & & & & & \vdots\\
 & & & & & & \cdots & & \boldsymbol{y}^r_1 & \xrightarrow{g_{\lambda'\cdots'}} & \mathbf{0}\\
 & & & & & & & & \boldsymbol{z}^1 & \xrightarrow{g_\lambda} & \mathbf{0}\\
 & & & & & & & & & & \vdots\\
 & & & & & & & & \boldsymbol{z}^{d-d'} & \xrightarrow{g_\lambda} & \mathbf{0}
\end{array}
$$

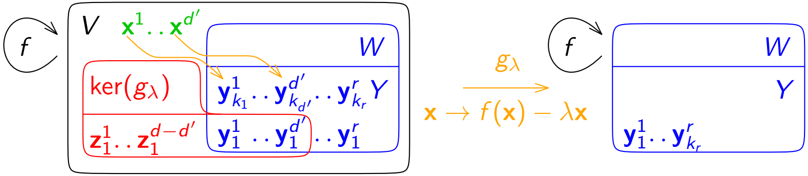
In our example:

$$
\begin{array}{ccccc}
(2,2,1)^T & \xrightarrow{g_2} & (3,2,1)^T & \xrightarrow{g_2} & \mathbf{0}\\
 & & (1,1,1)^T & \xrightarrow{g_1} & \mathbf{0}
\end{array}
$$

We have no $\boldsymbol{z}^i$ because $d = d' = 1$.

We added $d = \dim(\ker(g_\lambda))$ vectors to the basis of W, so in total we have as many as is the dimension of the space V.

We show that they are linearly independent and therefore they form a basis of the space V.



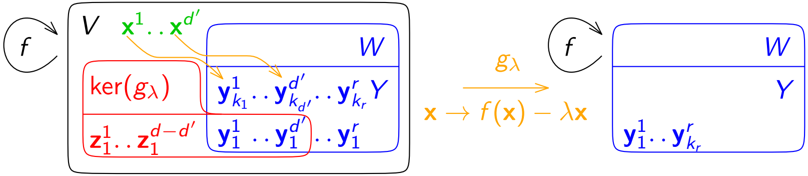
Consider a linear combination $\sum_i a_i \boldsymbol{x}^i + \sum_i b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i = \mathbf{0}$.

Since $\mathbf{0} = g_\lambda(\mathbf{0}) = g_\lambda\left(\sum_i a_i \boldsymbol{x}^i + \sum_{i,j} b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i\right) = \sum_{i,j} b'_{i,j} \boldsymbol{y}^i_j$,

where the vectors $\boldsymbol{y}^i_j$ are linearly independent, we must have $0 = b'_{i,j} = \begin{cases} a_i & \text{for } i \leq d', j = k_i \\ b_{i,j+1} & \text{for } i \leq d', j < k_i \\ (\lambda^* - \lambda) b_{i,j} & \text{for } i > d', j = k_i \\ (\lambda^* - \lambda) b_{i,j} + b_{i,j+1} & \text{for } i > d', j < k_i \end{cases}$

where $\lambda^* \neq \lambda$ matches the i-th chain.

It follows from $g_\lambda(\boldsymbol{x}^i) = \boldsymbol{y}^i_{k_i}$ and $g_\lambda(\boldsymbol{y}^i_j) = \boldsymbol{y}^i_{j-1}$ for $i \leq d'$; while for $i > d' : g_\lambda(\boldsymbol{y}^i_1) = f(\boldsymbol{y}^i_1) - \lambda \boldsymbol{y}^i_1 = \lambda^* \boldsymbol{y}^i_1 - \lambda \boldsymbol{y}^i_1 = (\lambda^* - \lambda) \boldsymbol{y}^i_1$ and for $j > 1$ also: $g_\lambda(\boldsymbol{y}^i_j) = f(\boldsymbol{y}^i_j) - \lambda \boldsymbol{y}^i_j = f(\boldsymbol{y}^i_j) - \lambda^* \boldsymbol{y}^i_j + (\lambda^* - \lambda) \boldsymbol{y}^i_j = g_{\lambda^*}(\boldsymbol{y}^i_j) + (\lambda^* - \lambda) \boldsymbol{y}^i_j = \boldsymbol{y}^i_{j-1} + (\lambda^* - \lambda) \boldsymbol{y}^i_j$.



Consider a linear combination $\sum_i a_i \boldsymbol{x}^i + \sum_i b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i = \mathbf{0}$.

Since $\mathbf{0} = g_\lambda(\mathbf{0}) = g_\lambda\left(\sum_i a_i \boldsymbol{x}^i + \sum_{i,j} b_{i,j} \boldsymbol{y}^i_j + \sum_i c_i \boldsymbol{z}^i\right) = \sum_{i,j} b'_{i,j} \boldsymbol{y}^i_j$,
where the vectors $\boldsymbol{y}^i_j$ are linearly independent, we must have $0 = b'_{i,j} = \begin{cases} a_i & \text{for } i \le d', j = k_i \\ b_{i,j+1} & \text{for } i \le d', j < k_i \\ (\lambda^* - \lambda) b_{i,j} & \text{for } i > d', j = k_i \\ (\lambda^* - \lambda) b_{i,j} + b_{i,j+1} & \text{for } i > d', j < k_i \end{cases}$
where $\lambda^* \neq \lambda$ matches the i-th chain.

The first case gives: $\forall i : a_i = 0$, the next: $\forall i \le d', \forall j > 1 : b_{i,j} = 0$ and the other two: $\forall i > d', \forall j : b_{i,j} = 0$. In the combination, only the coefficients $b_{i,1}$ for$i \le d'$ and c_i remain, but they are also zero, since the vectors $\boldsymbol{y}^1_1, \ldots, \boldsymbol{y}^{d'}_1, \boldsymbol{z}^1, \ldots, \boldsymbol{z}^{d-d'}$ form a basis of $\ker(g_\lambda)$.