Lecture 19: Polar and singular value decompositions; generalized eigenspaces; the decomposition theorem (1)

Travis Schedler

Thurs, Nov 17, 2011 (version: Thurs, Nov 17, 1:00 PM)
Goals (2)

- Polar decomposition and singular value decomposition
- Generalized eigenspaces and the decomposition theorem

Read Chapter 7, begin Chapter 8, and do PS 9.
Warm-up exercise (3)

(a) Let $T$ be an invertible operator on a f.d. i.p.s. and set $P := \sqrt{T^*T}$ and $S := TP^{-1}$. Show that $S$ is an isometry. Recall $P$ is positive, so

$$T = SP$$

is a polar decomposition (i.e., $S$ is an isometry and $P$ positive).

(b) Now suppose $T = 0$. Show that polar decompositions $T = SP$ are exactly $T = S0$ for every isometry $S$, i.e., we have always $P = 0$ but $S$ can be anything.

One-dimensional analogue: Either $z \in \mathbb{C}$ is invertible, in which case $z = (z/|z|)|z| = sp$ or else $z$ is zero, in which case $z = s \cdot 0$ for any $s$ of absolute value one.
Solution to warm-up exercise (4)

(a) \( S^*S = (TP^{-1})^* TP^{-1} = (P^{-1})^* T^* TP^{-1} = P^{-1}P^2P^{-1} = I. \) Here we used that \( P^* = P \) and hence \( (P^{-1})^* = P^{-1} \) as well.

(b) Since isometries are invertible, \( 0 = SP \) for \( S \) an isometry implies \( P = S^{-1}0 = 0. \) On the other hand clearly \( S0 = 0 \) for all \( S. \)
Polar decomposition and SVD (5)

Proposition: every complex number $z$ is expressible as $z = r \cdot e^{i\theta}$, where $r \geq 0$ and $\theta \in [0, 2\pi)$. (Unique if $z$ nonzero).
Polar decomposition and SVD (5)

Proposition: every complex number \( z \) is expressible as \( z = r \cdot e^{i\theta} \), where \( r \geq 0 \) and \( \theta \in [0, 2\pi) \). (Unique if \( z \) nonzero).

Equivalently: \( z = s \cdot r \), for \( |s| = 1 \) and \( r = |z| = \sqrt{\bar{z} \cdot z} \geq 0 \).
Polar decomposition and SVD (5)

Proposition: every complex number $z$ is expressible as $z = r \cdot e^{i\theta}$, where $r \geq 0$ and $\theta \in [0, 2\pi)$. (Unique if $z$ nonzero).
Equivalently: $z = s \cdot r$, for $|s| = 1$ and $r = |z| = \sqrt{\bar{z} \cdot z} \geq 0$.

Theorem

Let $V$ be a f.d. i.p.s. and $T \in \mathcal{L}(V)$. Then there is an expression $T = SP$, for $S$ an isometry and $P$ positive.
P is unique and $P = \sqrt{T^* T}$.
Moreover, $S$ is unique iff $T$ is invertible.
Polar decomposition and SVD (5)

Proposition: every complex number $z$ is expressible as $z = r \cdot e^{i\theta}$, where $r \geq 0$ and $\theta \in [0, 2\pi)$. (Unique if $z$ nonzero).
Equivalently: $z = s \cdot r$, for $|s| = 1$ and $r = |z| = \sqrt{\bar{z} \cdot z} \geq 0$.

Theorem
Let $V$ be a f.d. i.p.s. and $T \in \mathcal{L}(V)$. Then there is an expression $T = SP$, for $S$ an isometry and $P$ positive.
$P$ is unique and $P = \sqrt{T^* T}$.
Moreover, $S$ is unique iff $T$ is invertible.

Corollary (Singular Value Decomposition (SVD))
There exists orthonormal bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ of $V$ such that $T e_i = s_i f_i$, for $s_i \geq 0$ the singular values.
Moreover, $(e_1, \ldots, e_n)$ is an orthonormal eigenbasis of $T^* T$ with eigenvalues $s_i^2$. 
Polar decomposition and SVD (5)

Proposition: every complex number $z$ is expressible as $z = r \cdot e^{i\theta}$, where $r \geq 0$ and $\theta \in [0, 2\pi)$. (Unique if $z$ nonzero).
Equivalently: $z = s \cdot r$, for $|s| = 1$ and $r = |z| = \sqrt{\bar{z} \cdot z} \geq 0$.

Theorem
Let $V$ be a f.d. i.p.s. and $T \in \mathcal{L}(V)$. Then there is an expression $T = SP$, for $S$ an isometry and $P$ positive.
$P$ is unique and $P = \sqrt{T^*T}$.
Moreover, $S$ is unique iff $T$ is invertible.

Corollary (Singular Value Decomposition (SVD))
There exists orthonormal bases $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_n)$ of $V$ such that $Te_i = s_i f_i$, for $s_i \geq 0$ the singular values.
Moreover, $(e_1, \ldots, e_n)$ is an orthonormal eigenbasis of $T^*T$ with eigenvalues $s_i^2$.
Proof: Let $(e_1, \ldots, e_n)$ be an orthonormal eigenbasis of $T^*T$ and $s_1, \ldots, s_n$ the square roots of the eigenvalues. When $s_i \neq 0$, set $f_i := s_i^{-1} Te_i$. Then extend the resulting $f_i$ to an orthonormal eigenbasis.
Uniqueness of polar decomposition (6)

- If $T = SP$, then $T^* T = P^* S^* S P = P^* P = P^2$, so $P = \sqrt{T^* T}$. Thus $P$ is unique (positive operators have unique positive square roots; see the slides for Lecture 18 or Axler).

- If $T$ is invertible, $S = T P^{-1}$ so $S$ is unique.

- Conversely, if $T$ is not invertible, neither is $P$, and we can replace $S$ by $S S'$ where $S'$ is an isometry such that $S' v = v$ for all eigenvectors $v$ of nonzero eigenvalue. So then $S$ is not unique.
Existence of polar decomposition (7)

- Set $P := \sqrt{T^* T}$.
- $\text{range}(P)$ is $P$-invariant and $P$ is an isomorphism there (it has an eigenbasis with nonzero eigenvalues).
  Define thus $P|_{\text{range}(P)}^{-1} : \text{range}(P) \to \text{range}(P)$.
  Consider $S_1 := TP|_{\text{range}(P)}^{-1} : \text{range}(P) \to \text{range}(T)$.
  $S_1^* S_1 = I$, so $\langle u, v \rangle = \langle S_1 u, S_1 v \rangle$ for all $u, v \in \text{range}(P)$.
- Recall: null($P$) = null($T$). So dim range($P$) = dim range($T$).
  Thus $S_1$ takes an on. basis $(e_1, \ldots, e_m)$ of range($P$) to an on. basis $(f_1, \ldots, f_m)$ of range($T$).
- Extend $(e_i)$ and $(f_i)$ to on. bases of $V$ and extend $S_1$ to $S \in \mathcal{L}(V)$ by $S(e_i) = f_i$ when $i > m$.
- So $S$ takes an on. basis to another on. basis, i.e., it is an isometry.
- Finally, $T = SP$, since it is true on range($P$) and null($T$) = null($P$) = range($P$)$^\perp = \text{Span}(e_{m+1}, \ldots, e_n)$. 
Nonuniqueness of SVD (8)

- Note: the SVD is *not unique*, even if $T$ is invertible: the orthonormal eigenbasis $(e_i)$ of $T^*T$ is not unique. (e.g., one can reorder them and multiply by $\pm 1$, at the least.)
Note: the SVD is not unique, even if $T$ is invertible: the orthonormal eigenbasis $(e_i)$ of $T^* T$ is not unique. (e.g., one can reorder them and multiply by $\pm 1$, at the least.)

On the other hand, the polar decomposition is unique iff $T$ is invertible.

Example: $T = T A$, $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$. We can guess that $A = SP = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. So this is the answer (unique since $A$, equivalently $P$, is invertible).

For SVD we could have $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $s_1 = 1$, $s_2 = 2$, $f_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$. But we could also swap everything: $s_1$ with $s_2$, $e_1$ with $e_2$, and $f_1$ with $f_2$. Or we could take $e_1$ to $-e_1$ (hence $f_1$ to $-f_1$) and/or $e_2$ to $-e_2$ (hence $f_2$ to $-f_2$).
Nonuniqueness of SVD (8)

- Note: the SVD is \textit{not unique}, even if $T$ is invertible: the orthonormal eigenbasis $(e_i)$ of $T^*T$ is not unique. (e.g., one can reorder them and multiply by $\pm 1$, at the least.)

- On the other hand, the polar decomposition is unique iff $T$ is invertible.

- Example: $T = T_A$, $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$. 

- We can guess that $A = SP = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. So this is the answer (unique since $A$, equivalently $P$, is invertible).

- For SVD we could have $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $s_1 = 1$, $s_2 = 2$, $f_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

- But we could also swap everything: $s_1$ with $s_2$, $e_1$ with $e_2$, and $f_1$ with $f_2$. Or we could take $e_1$ to $-e_1$ (hence $f_1$ to $-f_1$) and/or $e_2$ to $-e_2$ (hence $f_2$ to $-f_2$).
Nonuniqueness of SVD (8)

- **Note:** the SVD is *not unique*, even if $T$ is invertible: the orthonormal eigenbasis $(e_i)$ of $T^* T$ is not unique. (e.g., one can reorder them and multiply by $\pm 1$, at the least.)
- On the other hand, the polar decomposition is unique iff $T$ is invertible.
- **Example:** $T = T_A$, $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$.
- We can guess that $A = SP = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. So this is the answer (unique since $A$, equivalently $P$, is invertible).
Nonuniqueness of SVD (8)

- Note: the SVD is not unique, even if $T$ is invertible: the orthonormal eigenbasis $(e_i)$ of $T^* T$ is not unique. (e.g., one can reorder them and multiply by $\pm 1$, at the least.)
- On the other hand, the polar decomposition is unique iff $T$ is invertible.
- Example: $T = T_A$, $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$.
- We can guess that $A = SP = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. So this is the answer (unique since $A$, equivalently $P$, is invertible).
- For SVD we could have $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $s_1 = 1$, $s_2 = 2$, $f_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $f_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.
Nonuniqueness of SVD (8)

Note: the SVD is not unique, even if \( T \) is invertible: the orthonormal eigenbasis \( (e_i) \) of \( T^* T \) is not unique. (e.g., one can reorder them and multiply by \( \pm 1 \), at the least.)

On the other hand, the polar decomposition is unique iff \( T \) is invertible.

Example: \( T = T_A, \ A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \).

We can guess that \( A = SP = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \). So this is the answer (unique since \( A \), equivalently \( P \), is invertible).

For SVD we could have \( e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, s_1 = 1, s_2 = 2, \)
\( f_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, f_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \).

But we could also swap everything: \( s_1 \) with \( s_2 \), \( e_1 \) with \( e_2 \), and \( f_1 \) with \( f_2 \). Or we could take \( e_1 \) to \( -e_1 \) (hence \( f_1 \) to \( -f_1 \)) and/or \( e_2 \) to \( -e_2 \) (hence \( f_2 \) to \( -f_2 \)).
Computing SVD and polar decomposition (9)

- The best way to compute these is to do SVD first; then let $P$ be the operator with eigenvectors $(e_i)$ and eigenvalues $(s_i)$, and let $S$ be the isometry $Se_i = f_i$ for all $i$. 

To compute SVD, given $T$, compute first $T^*T$. 

Then find the eigenvalues of $T^*T$ (2 × 2 case: characteristic polynomial: for $A = M(T)$ in an orthonormal basis, these are the roots of $x^2 - \text{tr}(\bar{A}^t\bar{A})x + \det(\bar{A}^t\bar{A})$). 

Find the eigenspaces and an orthonormal eigenbasis $(e_i)$ of $T^*T$. 

Next, set $P := \sqrt{T^*T}$, by taking the nonnegative square root of the eigenvalues. These eigenvalues are the $s_i$. 

Finally, let $f_i := s_i^{-1}Te_i$ for the nonzero $s_i$; for the remaining $f_i$ just extend the ones we get to an orthonormal basis.
Computing SVD and polar decomposition (9)

- The best way to compute these is to do SVD first; then let $P$ be the operator with eigenvectors $(e_i)$ and eigenvalues $(s_i)$, and let $S$ be the isometry $Se_i = f_i$ for all $i$.
- To compute SVD, given $T$, compute first $T^*T$. 
The best way to compute these is to do SVD first; then let $P$ be the operator with eigenvectors $(e_i)$ and eigenvalues $(s_i)$, and let $S$ be the isometry $Se_i = f_i$ for all $i$.

To compute SVD, given $T$, compute first $T^*T$.

Then find the eigenvalues of $T^*T$ (2 × 2 case: characteristic polynomial: for $A = M(T)$ in an orthonormal basis, these are the roots of $x^2 - \text{tr}(\bar{A}^tA)x + \det(\bar{A}^tA)$.)
Computing SVD and polar decomposition (9)

- The best way to compute these is to do SVD first; then let $P$ be the operator with eigenvectors $(e_i)$ and eigenvalues $(s_i)$, and let $S$ be the isometry $Se_i = f_i$ for all $i$.
- To compute SVD, given $T$, compute first $T^*T$.
- Then find the eigenvalues of $T^*T$ ($2 \times 2$ case: characteristic polynomial: for $A = \mathcal{M}(T)$ in an orthonormal basis, these are the roots of $x^2 - \text{tr}(A^tA)x + \det(A^tA)$.)
- Find the eigenspaces and an orthonormal eigenbasis $(e_i)$ of $T^*T$. 


Computing SVD and polar decomposition (9)

- The best way to compute these is to do SVD first; then let $P$ be the operator with eigenvectors ($e_i$) and eigenvalues ($s_i$), and let $S$ be the isometry $Se_i = f_i$ for all $i$.
- To compute SVD, given $T$, compute first $T^*T$.
- Then find the eigenvalues of $T^*T$ ($2 \times 2$ case: characteristic polynomial: for $A = M(T)$ in an orthonormal basis, these are the roots of $x^2 - \text{tr}(\bar{A}^tA)x + \det(\bar{A}^tA)$.)
- Find the eigenspaces and an orthonormal eigenbasis ($e_i$) of $T^*T$.
- Next, set $P := \sqrt{T^*T}$, by taking the nonnegative square root of the eigenvalues. These eigenvalues are the $s_i$. 
The best way to compute these is to do SVD first; then let $P$ be the operator with eigenvectors $(e_i)$ and eigenvalues $(s_i)$, and let $S$ be the isometry $Se_i = f_i$ for all $i$.

To compute SVD, given $T$, compute first $T^*T$.

Then find the eigenvalues of $T^*T$ (2 × 2 case: characteristic polynomial: for $A = M(T)$ in an orthonormal basis, these are the roots of $x^2 - \text{tr}(\overline{A}^tA)x + \det(\overline{A}^tA)$.)

Find the eigenspaces and an orthonormal eigenbasis $(e_i)$ of $T^*T$.

Next, set $P := \sqrt{T^*T}$, by taking the nonnegative square root of the eigenvalues. These eigenvalues are the $s_i$.

Finally, let $f_i := s_i^{-1}Te_i$ for the nonzero $s_i$; for the remaining $f_i$ just extend the ones we get to an orthonormal basis.
Example (10)

Example from before: $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$.
Example (10)

Example from before: $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$.

First, $\bar{A}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.

Next, an eigenbasis is $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ with eigenvalues 1 and 4.

So $P = \sqrt{\bar{A}^t A}$ has the same eigenbasis, with eigenvalues $s_1 = \sqrt{1} = 1$ and $s_2 = \sqrt{4} = 2$.

Then $f_1 = s_1^{-1} A e_1 = 1 - 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Also $f_2 = s_2^{-1} A e_2 = 2 - 1 \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

Now $P = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, as desired.

In general: $P = (e_1 e_2)\begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix}(e_1 e_2)^{-1}$ and $S = (f_1 f_2)(e_1 e_2)^{-1}$.
Example (10)

- Example from before: $A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$.

- First, $\bar{A}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$.

- Next, an eigenbasis is $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ with eigenvalues 1 and 4.
Example (10)

- Example from before: \( A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \).
- First, \( \bar{A}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \).
- Next, an eigenbasis is \( e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and \( e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) with eigenvalues 1 and 4.
- So \( P = \sqrt{\bar{A}^t A} \) has the same eigenbasis, with eigenvalues \( s_1 = \sqrt{1} = 1 \) and \( s_2 = \sqrt{4} = 2 \).
Example (10)

- Example from before: \( A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \).

- First, \( \bar{A}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \).

- Next, an eigenbasis is \( e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and \( e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) with eigenvalues 1 and 4.

- So \( P = \sqrt{\bar{A}^t A} \) has the same eigenbasis, with eigenvalues \( s_1 = \sqrt{1} = 1 \) and \( s_2 = \sqrt{4} = 2 \).

- Then \( f_1 = s_1^{-1} A e_1 = 1^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \). Also \( f_2 = s_2^{-1} A e_2 = 2^{-1} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \).
Example (10)

Example from before: \( A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \).

First, \( \bar{A}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \).

Next, an eigenbasis is \( e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and \( e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) with eigenvalues 1 and 4.

So \( P = \sqrt{\bar{A}^t A} \) has the same eigenbasis, with eigenvalues \( s_1 = \sqrt{1} = 1 \) and \( s_2 = \sqrt{4} = 2 \).

Then \( f_1 = s_1^{-1} A e_1 = 1^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \). Also

\( f_2 = s_2^{-1} A e_2 = 2^{-1} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \).

Now \( P = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \) and \( S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \), as desired.
Example (10)

- Example from before: \( A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} \).

- First, \( \bar{A}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \).

- Next, an eigenbasis is \( e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and \( e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) with eigenvalues 1 and 4.

- So \( P = \sqrt{\bar{A}^t A} \) has the same eigenbasis, with eigenvalues \( s_1 = \sqrt{1} = 1 \) and \( s_2 = \sqrt{4} = 2 \).

- Then \( f_1 = s_1^{-1} A e_1 = 1^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \). Also \( f_2 = s_2^{-1} A e_2 = 2^{-1} \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \).

- Now \( P = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \) and \( S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \), as desired.

- In general: \( P = (e_1 e_2) \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} (e_1 e_2)^{-1} \) and \( S = (f_1 f_2)(e_1 e_2)^{-1} \).
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over \( F = \mathbb{C} \) they always admit a basis of \textit{generalized eigenvectors}.
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over $\mathbb{F} = \mathbb{C}$ they always admit a basis of \textit{generalized eigenvectors}.

Definition

A generalized eigenvector $v$ of $T$ eigenvalue $\lambda$ is one such that, for some $m \geq 1$, $(T - \lambda I)^m v = 0$.
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over $\mathbb{F} = \mathbb{C}$ they always admit a basis of \textit{generalized eigenvectors}.

\textbf{Definition}

A generalized eigenvector $v$ of $T$ eigenvalue $\lambda$ is one such that, for some $m \geq 1$, $(T - \lambda I)^m v = 0$.

\textbf{Examples}:
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over $\mathbb{F} = \mathbb{C}$ they always admit a basis of generalized eigenvectors.

Definition
A generalized eigenvector $v$ of $T$ eigenvalue $\lambda$ is one such that, for some $m \geq 1$, $(T - \lambda I)^m v = 0$.

Examples:
- $m = 1$ above if and only if $v$ is an (ordinary) eigenvector.
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over $\mathbb{F} = \mathbb{C}$ they always admit a basis of generalized eigenvectors.

Definition
A generalized eigenvector $v$ of $T$ eigenvalue $\lambda$ is one such that, for some $m \geq 1$, $(T - \lambda I)^m v = 0$.

Examples:
- $m = 1$ above if and only if $v$ is an (ordinary) eigenvector.
- If $T$ is nilpotent, then all vectors are generalized eigenvectors of eigenvalue zero. So, even though it does not have an eigenbasis, every basis is a basis of generalized eigenvectors!
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over $\mathbf{F} = \mathbb{C}$ they always admit a basis of generalized eigenvectors.

Definition
A generalized eigenvector $v$ of $T$ eigenvalue $\lambda$ is one such that, for some $m \geq 1$, $(T - \lambda I)^m v = 0$.

Examples:
- $m = 1$ above if and only if $v$ is an (ordinary) eigenvector.
- If $T$ is nilpotent, then all vectors are generalized eigenvectors of eigenvalue zero. So, even though it does not have an eigenbasis, every basis is a basis of generalized eigenvectors!

Definition
Let $V(\lambda)$ be the generalized eigenspace of eigenvalue $\lambda$: the span of all generalized eigenvectors of eigenvalue $\lambda$. 
Generalized eigenvectors (11)

Goal: Although not all f.d. vector spaces admit an eigenbasis, over $\mathbb{F} = \mathbb{C}$ they always admit a basis of generalized eigenvectors.

Definition
A generalized eigenvector $v$ of $T$ eigenvalue $\lambda$ is one such that, for some $m \geq 1$, $(T - \lambda I)^m v = 0$.

Examples:
- $m = 1$ above if and only if $v$ is an (ordinary) eigenvector.
- If $T$ is nilpotent, then all vectors are generalized eigenvectors of eigenvalue zero. So, even though it does not have an eigenbasis, every basis is a basis of generalized eigenvectors!

Definition
Let $V(\lambda)$ be the generalized eigenspace of eigenvalue $\lambda$: the span of all generalized eigenvectors of eigenvalue $\lambda$.

Note that $V(\lambda)$ is $T$-invariant, since $(T - \lambda I)^m v = 0$ implies $(T - \lambda I)^m T v = T(T - \lambda I)^m v = 0$. 
The decomposition theorem (12)

Theorem

Let $V$ be f.d., $F = \mathbb{C}$, and $T \in \mathcal{L}(V)$. Then $V$ is the direct sum of its generalized eigenspaces: $V = \bigoplus \lambda V(\lambda)$. 

First step: 
Lemma 

Suppose that $\lambda \neq \mu$. Then $V(\lambda) \cap V(\mu) = \{0\}$.

Proof. 

Suppose $v \in V(\lambda) \cap V(\mu)$ is nonzero. Let $(T - \lambda I)^m v = 0$ but not $(T - \lambda I)^{m-1} v$. So $u := (T - \lambda I)^{m-1} v$ is a nonzero eigenvector of eigenvalue $\lambda$.

Now let $m' \geq 1$ be such that $(T - \mu I)^{m'} v = 0$. Then also $(T - \mu I)^{m'} u = (T - \mu I)^{m'} (T - \lambda I)^{m-1} v = 0$.

But, $(T - \mu I)^{m'} u = (\lambda - \mu)^{m'} u \neq 0$, since $\lambda \neq \mu$. This is a contradiction.
The decomposition theorem (12)

**Theorem**

Let $V$ be f.d., $\mathbf{F} = \mathbf{C}$, and $T \in \mathcal{L}(V)$. Then $V$ is the direct sum of its generalized eigenspaces: $V = \bigoplus \lambda V(\lambda)$.

**First step:**

**Lemma**

Suppose that $\lambda \neq \mu$. Then $V(\lambda) \cap V(\mu) = \{0\}$. 
The decomposition theorem (12)

Theorem

Let \( V \) be f.d., \( F = \mathbb{C} \), and \( T \in \mathcal{L}(V) \). Then \( V \) is the direct sum of its generalized eigenspaces: \( V = \bigoplus_{\lambda} V(\lambda) \).

First step:

Lemma

Suppose that \( \lambda \neq \mu \). Then \( V(\lambda) \cap V(\mu) = \{0\} \).

Proof.

- Suppose \( v \in V(\lambda) \cap V(\mu) \) is nonzero. Let \((T - \lambda I)^m v = 0 \) but not \((T - \lambda I)^{m-1} v \). So \( u := (T - \lambda I)^{m-1} v \) is a nonzero eigenvector of eigenvalue \( \lambda \).

- Now let \( m' \geq 1 \) be such that \((T - \mu I)^{m'} v = 0 \). Then also \((T - \mu I)^{m'} u = (T - \mu I)^{m'} (T - \lambda I)^{m-1} v = (T - \lambda I)^{m-1} (T - \mu I)^{m'} v = 0 \).

- But, \((T - \mu I)^{m'} u = (\lambda - \mu)^{m'} u \neq 0 \), since \( \lambda \neq \mu \). This is a contradiction.
Lemma on generalized eigenspaces (13)

**Lemma**

\[ \mathcal{V}(\lambda) = \text{null}(T - \lambda I)^{\dim V}. \]

I.e., if \( v \) is a generalized eigenvector of eigenvalue \( \lambda \), we can take \( m = \dim V \) before: \( (T - \lambda I)^{\dim V} v = 0 \).
Lemma on generalized eigenspaces (13)

Lemma
\[ V(\lambda) = \text{null}(T - \lambda I)^{\dim V}. \]
I.e., if \( v \) is a generalized eigenvector of eigenvalue \( \lambda \), we can take \( m = \dim V \) before: \( (T - \lambda I)^{\dim V} v = 0. \)

Proof.

- Let \( U_i := (T - \lambda I)^i V(\lambda). \)
- Since \( V(\lambda) \) is \( T \)-invariant (hence \( (T - \lambda I) \)-invariant), \( U_0 \supseteq U_1 \supseteq \cdots. \)
- However, if \( U_i = U_{i+1} \), then \( (T - \lambda I) \) is injective on \( U_i \). Since \( (T - \lambda I) \) is nilpotent, this implies \( U_i = \{0\}. \)
- So \( U_0 \supsetneq U_1 \supsetneq \cdots \supsetneq U_m = \{0\} \), and \( \dim U_i \leq \dim V(\lambda) - i \). Hence \( m \leq \dim V(\lambda) \leq \dim V \), and \( (T - \lambda I)^{\dim V} v = 0 \) for all \( v \in V(\lambda). \)
One more lemma (14)

Lemma

\[ V = (T - \lambda I)^{\dim V} \oplus \text{range}(T - \lambda I)^{\dim V} = V(\lambda) \oplus \text{range}(T - \lambda I)^{\dim V}. \]
One more lemma (14)

Lemma

\[ V = (T - \lambda I)^{\dim V} \oplus \text{range}(T - \lambda I)^{\dim V} = V(\lambda) \oplus \text{range}(T - \lambda I)^{\dim V}. \]

Proof.

- Since the dimensions are equal, we need to show just that \((T - \lambda I)^{\dim V} \cap \text{range}(T - \lambda I)^{\dim V} = \{0\} \).
- Let \( v \in (T - \lambda I)^{\dim V} \cap \text{range}(T - \lambda I)^{\dim V} \). Write \( v = (T - \lambda I)^{\dim V}u \).
- Since \( (T - \lambda I)^{2\dim V}u = (T - \lambda I)^{\dim V}v = 0 \), also \( u \) is a generalized eigenvector of eigenvalue \( \lambda \).
- But, by the last lemma, then \( (T - \lambda I)^{\dim V}u = 0 \), so \( v = 0 \).

\[ \square \]
Proof of the decomposition theorem (15)

Theorem
Let $V$ be f.d., $F = \mathbb{C}$, and $T \in \mathcal{L}(V)$. Then $V$ is the direct sum of its generalized eigenspaces: $V = \bigoplus \lambda V(\lambda)$. 
Proof of the decomposition theorem (15)

Theorem
Let $V$ be f.d., $F = \mathbb{C}$, and $T \in \mathcal{L}(V)$. Then $V$ is the direct sum of its generalized eigenspaces: $V = \bigoplus \lambda V(\lambda)$.

Proof: By induction on $\dim V$. Let $\lambda$ be an eigenvalue of $T$, so $V(\lambda) \neq \{0\}$.

Write $V = V(\lambda) \oplus \text{range}(T - \lambda I)^{\dim V}$.

Since $\dim \text{range}(T - \lambda I)^{\dim V} < \dim V$, the ind. hyp. shows that $\text{range}(T - \lambda I)$ is the direct sum of the generalized eigenspaces of $T|_{\text{range}(T-\lambda I)^{\dim V}}$.

To conclude, we claim that for $\mu \neq \lambda$, $V(\mu) \subseteq \text{range}(T - \lambda I)^{\dim V}$. Thus $V(\mu)$ is a generalized eigenspace of $T|_{\text{range}(T-\lambda I)^{\dim V}}$.

For this, we show that $(T - \lambda I)^{\dim V} V(\mu) = V(\mu)$.

First, $V(\mu)$ is $T$-invariant, so $(T - \lambda I)^{\dim V} V(\mu) \subseteq V(\mu)$. We only need to show $(T - \lambda I)^{\dim V}$ is injective on $V(\mu)$.

This means that $V(\mu) \cap \text{null}(T - \lambda I)^{\dim V} = \{0\}$. But this is $V(\mu) \cap V(\lambda) = \{0\}$, by the lemmas.