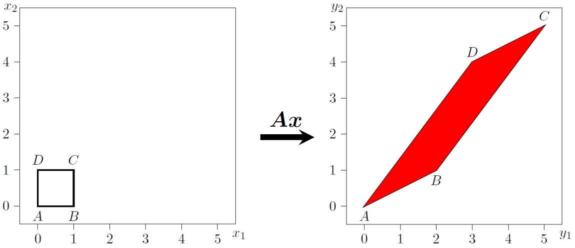
Statistical Sciences Honours
Matrix Methods
Lecture 6 – Eigenvalues, Eigenvectors and Singular Value Decomposition

Stefan S. Britz
stefan.britz@uct.ac.za

Department of Statistical Sciences
University of Cape Town

6.1 Basic ideas

In lecture 2 we saw that multiplying a matrix by a vector results in a linear transformation that (possibly) changes the vector.



$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

	Initial coordinates	Transformed coordinates
A	$\boldsymbol{x}' = [0\ \ 0]$	$\boldsymbol{y}' = [0\ \ 0]$
B	$\boldsymbol{x}' = [1\ \ 0]$	$\boldsymbol{y}' = [a_{11}\ \ a_{21}]$
C	$\boldsymbol{x}' = [1\ \ 1]$	$\boldsymbol{y}' = [a_{11} + a_{12}\ \ a_{21} + a_{22}]$
D	$\boldsymbol{x}' = [0\ \ 1]$	$\boldsymbol{y}' = [a_{12}\ \ a_{22}]$

6.1 Definition

- The question now arises: can we find vectors such that, for some matrix multiplication, the vector changes in such a way that it still lies on the same line?
- In other words, after this transformation it will either still point in the same direction or in the exact opposite direction.
- We can view this by saying that the effect of the matrix multiplication is that the vector is only multiplied by some constant.
- So if we consider some $n \times n$ matrix $\boldsymbol{A}$ and a non-zero vector $\boldsymbol{x}$, then we have the relationship
$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

6.1 Definition

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$$

- If a vector $\boldsymbol{x}$ and scalar λ satisfies this equation, then we refer to them respectively as an **eigenvector** and corresponding **eigenvalue** of $\boldsymbol{A}$.
- Note that the scaling of $\boldsymbol{x}$ is arbitrary, since if the above is true, then $\boldsymbol{A}(k\boldsymbol{x}) = \lambda(k\boldsymbol{x})$.
- It is therefore common to express the eigenvector in normalised form.

6.2 Evaluating eigenvalues and eigenvectors

- We can now rewrite the equation such that

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \lambda\boldsymbol{x} \\ \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{x} &= \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{I}\boldsymbol{x} &= \boldsymbol{0} \\ (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} &= \boldsymbol{0} \end{aligned}$$

6.2 Evaluating eigenvalues and eigenvectors

- We can now rewrite the equation such that

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \lambda\boldsymbol{x} \\ \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{x} &= \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x} - \lambda\boldsymbol{I}\boldsymbol{x} &= \boldsymbol{0} \\ (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} &= \boldsymbol{0} \end{aligned}$$

- From lecture 4 we know that a non-trivial solution will only exist if $(\boldsymbol{A} - \lambda\boldsymbol{I})$ is singular.

6.2 Evaluating eigenvalues and eigenvectors

- We can now rewrite the equation such that

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &= \lambda \boldsymbol{x} \\ \boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{x} &= \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{I}\boldsymbol{x} &= \boldsymbol{0} \\ (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} &= \boldsymbol{0} \end{aligned}$$

- From lecture 4 we know that a non-trivial solution will only exist if $(\boldsymbol{A} - \lambda \boldsymbol{I})$ is singular.

- Therefore, to find eigenvalues we can now determine for which values of λ the following holds:

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$$

6.2 Evaluating eigenvalues and eigenvectors

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$$

- This is known as the characteristic equation.
- It defines an n^{th}-order polynomial in λ, which we use to solve for n values of λ, although note that some values may repeat.
- For each solution we can the solve for $\boldsymbol{x}$ from the equation $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$.

Example

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} -1 & -2 & -2 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

Exercise

Exercise 6.1

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 0 & 1 & 2 \\ 2 & 3 & 0 \\ 0 & 4 & 5 \end{bmatrix}$$

6.3 Properties I

1. If

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

then

$$\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}(\lambda\boldsymbol{x}) = \lambda\boldsymbol{A}\boldsymbol{x} = \lambda^2\boldsymbol{x}$$

This can be expanded such that, in general,

$$\boldsymbol{A}^k\boldsymbol{x} = \lambda^k\boldsymbol{x}$$

2. If $\boldsymbol{A}$ is non-singular, all the eigenvalues will be non-zero.

   In this case, we can also extend the previous property such that $\left(\frac{1}{\lambda}, \boldsymbol{x}\right)$ is an eigenvalue-eigenvector pair for $\boldsymbol{A}^{-1}$.

   $\lambda = 0$ can be an eigenvalue of $\boldsymbol{A}$, but this implies that $|\boldsymbol{A}| = 0$.

6.3 Properties II

3. The rank of $\boldsymbol{A}$ is the number of non-zero eigenvalues of $\boldsymbol{A}$ (counting repeated roots as many times as they occur).

4. The sum of a matrix's eigenvalues equals its trace

$$\sum_{i=1}^{n} \lambda_i = tr(\boldsymbol{A})$$

   The product of a matrix's eigenvalues equals its determinant

$$\prod_{i=1}^{n} \lambda_i = |\boldsymbol{A}|$$

6.3 Properties III

5. For each of the n eigenvalues of $\boldsymbol{A}$, $\lambda_1, \lambda_2, \ldots, \lambda_n$ (some of which may be repeated), there are corresponding eigenvectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$.

   For repeated eigenvalues, we may or may not have repeated eigenvectors.

   By constructing the $n \times n$ matrix $\boldsymbol{U}$ having the eigenvectors as columns, we have that:

   $$\boldsymbol{A}\boldsymbol{U} = \boldsymbol{A}[\boldsymbol{u}_1 \quad \cdots \quad \boldsymbol{u}_n] = [\lambda_1 \boldsymbol{u}_1 \quad \cdots \quad \lambda_n \boldsymbol{u}_n] = \boldsymbol{U}\boldsymbol{D}$$

   where $\boldsymbol{D} = diag(\lambda_i)$.

   If $\boldsymbol{U}$ is non-singular, then $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{D}$ and $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1}$

6.3 Properties IV

Combining the results from properties 2 and 5, we have that:

$$\boldsymbol{A}^k = \boldsymbol{U}\boldsymbol{D}^k\boldsymbol{U}^{-1}$$

and

$$\boldsymbol{A}^{-1} = \boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^{-1}$$

where

$\boldsymbol{D}^k = diag\left(\lambda_i^k\right)$ and $\boldsymbol{D}^{-1} = diag\left(\frac{1}{\lambda_i}\right)$

6.4 Properties of symmetric matrices

- For **symmetric matrices** some interesting properties emerge.
- Firstly, all eigenvalues are real.
- Secondly, suppose we have two eigenvalues λ_k and λ_l of the symmetric matrix $\boldsymbol{A}$ such that $\lambda_k \neq \lambda_l$. Let $\boldsymbol{u}_k$ and $\boldsymbol{u}_l$ be the corresponding eigenvectors. Then

$$\begin{aligned}
\lambda_k \boldsymbol{u}_l' \boldsymbol{u}_k &= \boldsymbol{u}_l' \lambda_k \boldsymbol{u}_k \\
&= \boldsymbol{u}_l' \boldsymbol{A} \boldsymbol{u}_k \\
&= \boldsymbol{u}_k' \boldsymbol{A}' \boldsymbol{u}_l \\
&= \boldsymbol{u}_k' \boldsymbol{A} \boldsymbol{u}_l \\
&= \boldsymbol{u}_k' \lambda_l \boldsymbol{u}_l \\
&= \lambda_l \boldsymbol{u}_k' \boldsymbol{u}_l \\
\lambda_k \boldsymbol{u}_l' \boldsymbol{u}_k &= \lambda_l \boldsymbol{u}_l' \boldsymbol{u}_k
\end{aligned}$$

6.4 Properties of symmetric matrices

$$\lambda_k \boldsymbol{u}_l' \boldsymbol{u}_k = \lambda_l \boldsymbol{u}_l' \boldsymbol{u}_k$$

- Now, since $\lambda_k \neq \lambda_l$, this implies that $\boldsymbol{u}_l' \boldsymbol{u}_k = 0$.
- In other words, symmetric matrices with unique eigenvalues have orthogonal eigenvectors.
- It can also be shown that even with repeated eigenvalues, orthogonal sets of eigenvectors can be found (with the Gram-Schmidt process) when $\boldsymbol{A}$ is symmetric.

6.4 Properties of symmetric matrices

- Thus it follows that if $\boldsymbol{A}$ is symmetric, then

$$\begin{aligned} \boldsymbol{U}'\boldsymbol{U} &= \boldsymbol{I} \\ \boldsymbol{U}' &= \boldsymbol{U}^{-1} \end{aligned}$$

- From Property 5 we can see that

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}'$$

- This can be used to define spectral decomposition for a symmetric matrix.

6.5 Spectral decomposition

The spectral decomposition of a symmetric matrix $\boldsymbol{A}$ follows directly from

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}'$$

6.5 Spectral decomposition

The spectral decomposition of a symmetric matrix $\boldsymbol{A}$ follows directly from

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}'$$

$$\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1' \\ \vdots \\ \boldsymbol{u}_n' \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 \boldsymbol{u}_1 & \cdots & \lambda_n \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1' \\ \vdots \\ \boldsymbol{u}_n' \end{bmatrix} \\
&= \sum_{k=1}^{n} \lambda_k \boldsymbol{u}_k \boldsymbol{u}_k'
\end{aligned}$$

6.5 Spectral decomposition

The spectral decomposition of a symmetric matrix $\boldsymbol{A}$ follows directly from

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}'$$

$$\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1' \\ \vdots \\ \boldsymbol{u}_n' \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 \boldsymbol{u}_1 & \cdots & \lambda_n \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1' \\ \vdots \\ \boldsymbol{u}_n' \end{bmatrix} \\
&= \sum_{k=1}^{n} \lambda_k \boldsymbol{u}_k \boldsymbol{u}_k'
\end{aligned}$$

Therefore, spectral decomposition expresses a symmetric matrix $\boldsymbol{A}$ as the sum of n rank 1 matrices.

6.6 Singular value decomposition

We will now generalise this idea to non-square matrices

- Let $\boldsymbol{X}: n \times p$ be any matrix of rank k where $k \leq p < n$.
- The matrix $\boldsymbol{X}$ can be expressed in the form

$$\boldsymbol{X} = \boldsymbol{U}^* \boldsymbol{D}^* \boldsymbol{V}'$$

where $\boldsymbol{U}^*: n \times n$ and $\boldsymbol{V}: p \times p$ are orthogonal matrices, and

$$\boldsymbol{D}^*: n \times p = \begin{bmatrix} \boldsymbol{D}: k \times k & \mathbf{0}: k \times (p-k) \\ \mathbf{0}: (n-k) \times k & \mathbf{0}: (n-k) \times (p-k) \end{bmatrix}$$

where $\boldsymbol{D} = diag(d_i) = diag(\sqrt{\lambda_i})$

6.6 Singular value decomposition

We will now generalise this idea to non-square matrices

- Let $\boldsymbol{X}: n \times p$ be any matrix of rank k where $k \leq p < n$.
- The matrix $\boldsymbol{X}$ can be expressed in the form

$$\boldsymbol{X} = \boldsymbol{U}^* \boldsymbol{D}^* \boldsymbol{V}'$$

where $\boldsymbol{U}^*: n \times n$ and $\boldsymbol{V}: p \times p$ are orthogonal matrices, and

$$\boldsymbol{D}^*: n \times p = \begin{bmatrix} \boldsymbol{D}: k \times k & \mathbf{0}: k \times (p-k) \\ \mathbf{0}: (n-k) \times k & \mathbf{0}: (n-k) \times (p-k) \end{bmatrix}$$

where $\boldsymbol{D} = diag(d_i) = diag(\sqrt{\lambda_i})$

- The positive quantities $\lambda_1, \ldots, \lambda_k$ are the non-zero eigenvalues of both the symmetric matrices $\boldsymbol{X}'\boldsymbol{X}: p \times p$ and $\boldsymbol{X}\boldsymbol{X}': n \times n$.

6.6 Singular value decomposition

$\boldsymbol{X} = \boldsymbol{U}^*\boldsymbol{D}^*\boldsymbol{V}'$ is often called the full version of the SVD, since it can also be written as the **compact SVD**:

$$ {}_n\boldsymbol{X}_p = {}_n\boldsymbol{U}_p \boldsymbol{D}_p \boldsymbol{V}'_p $$

where $\boldsymbol{U}$ consists of the first p columns of $\boldsymbol{U}^*$

- Note that since ${}_n\boldsymbol{U}_p$ is not square, its columns are mutually orthonormal, but not the rows:

$$ {}_p\boldsymbol{U}'_n\boldsymbol{U}_p = \boldsymbol{I}_p \quad ; \quad {}_n\boldsymbol{U}_p\boldsymbol{U}'_n \neq \boldsymbol{I}_n $$

6.6 Singular value decomposition

- The columns of the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are called the left and right **singular vectors** of $\boldsymbol{X}$ respectively.
- The matrix $\boldsymbol{D}$ is a diagonal matrix containing the singular values on the diagonal.
- It is assumed unless stated otherwise that the singular values are ordered in decreasing order.
- Without loss of generality, the singular values are always positive.
- One practical problem in applying the SVD is that eigenvectors are defined in an arbitrary directional sense – if $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$, then so is $-\boldsymbol{x}$. One needs to check that the directions of the eigenvectors in $\boldsymbol{U}$ and in $\boldsymbol{V}$ are consistently defined.

6.6 Singular value decomposition

If $\boldsymbol{X} = \boldsymbol{UDV}'$, then

- $\boldsymbol{U}$ is the eigenvectors of $\boldsymbol{XX}'$, since $(\boldsymbol{XX}')\boldsymbol{U} = \boldsymbol{UD}^2$
- $\boldsymbol{V}$ is the eigenvectors of $\boldsymbol{X'X}$, since $(\boldsymbol{XX}')\boldsymbol{V} = \boldsymbol{VD}^2$
- $\boldsymbol{D}$ contains the **singular values** of $\boldsymbol{X}$ on the diagonal.
- $\boldsymbol{D}^2$ contains the eigenvalues of $\boldsymbol{XX'}$ and $\boldsymbol{X'X}$ on the diagonal.

This video by Michael Greenacre summarises the (wonderful) SVD methodology melodically.