Introduction to data science & artificial intelligence (INF7100)

Arthur Charpentier

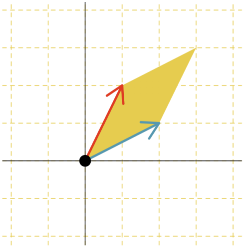
#422 Linear Algebra

été 2020

Determinant of square matrices (2×2)

Let $\boldsymbol{M} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, then

$$\det(\boldsymbol{M}) = |\boldsymbol{M}| = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$$



Note that $|\det(\boldsymbol{M})|$ is the area of the parallelogram, $\vec{\boldsymbol{u}} = \begin{pmatrix} a \\ b \end{pmatrix}$ and $\vec{\boldsymbol{v}} = \begin{pmatrix} c \\ d \end{pmatrix}$

Thus, $\det(\boldsymbol{M}) = 0$ if and only if $\vec{\boldsymbol{u}} /\!/ \vec{\boldsymbol{v}}$
Let θ denote the angle $(\vec{\boldsymbol{u}}, \vec{\boldsymbol{v}})$,

$$\cos(\theta) = \frac{\langle \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \rangle}{\|\vec{\boldsymbol{u}}\| \cdot \|\vec{\boldsymbol{v}}\|} \text{ and } \sin(\theta) = \frac{\det(M)}{\|\vec{\boldsymbol{u}}\| \cdot \|\vec{\boldsymbol{v}}\|}, \text{ where } M = (\vec{\boldsymbol{u}} \ \vec{\boldsymbol{v}})$$

Inverse of square matrices (2×2)

If $\det(\boldsymbol{M}) \neq 0$, then $\boldsymbol{M}$ has an inverse, denoted $\boldsymbol{M}^{-1}$,

$$\boldsymbol{M}\boldsymbol{M}^{-1} = \boldsymbol{M}^{-1}\boldsymbol{M} = \mathbb{I}$$

Example: $\boldsymbol{M} = \begin{pmatrix} 1 & 2 \\ -2 & 0 \end{pmatrix}$ then $\boldsymbol{M}^{-1} = \dfrac{1}{4}\begin{pmatrix} 0 & -1 \\ 2 & 1 \end{pmatrix}$

Important for linear systems: find $\boldsymbol{x}$ such that $\boldsymbol{M}\boldsymbol{x} = \boldsymbol{a}$,
then the unique solution is $\boldsymbol{x} = \boldsymbol{M}^{-1}\boldsymbol{a}$

Example: $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ then $\boldsymbol{\Sigma}^{-1} = \dfrac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$, if $|\rho| < 1$.
(here $\boldsymbol{\Sigma}$ is a variance matrix)

A projection matrix cannot be inverted

Eigenvalues and Eigenvectors of square matrices

Let $\boldsymbol{M} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, $\vec{\boldsymbol{u}} = \begin{pmatrix} x \\ y \end{pmatrix}$ and λ such that $\boldsymbol{M}\vec{\boldsymbol{u}} = \lambda\vec{\boldsymbol{u}}$,

$\vec{\boldsymbol{u}}$ is the eigenvector associated with eigenvalue λ.

$\begin{pmatrix} a-\lambda & c \\ b & d-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ implies $\det\begin{pmatrix} a-\lambda & c \\ b & d-\lambda \end{pmatrix} = 0$.

i.e. solve (in λ) $\det(M - \lambda\mathbb{I}) = 0$.
This can be extended in higher dimension

If $\det(\boldsymbol{M}) = 0$ then $\lambda = 0$ is an eigenvalue.

If $\boldsymbol{M}$ is symmetric, eigenvectors are orthogonal.

$\boldsymbol{M}$ is said to be positive if all eigenvalues are positive, and then $\boldsymbol{z}^\top\boldsymbol{M}\boldsymbol{z} \geq 0$, $\forall\boldsymbol{z}$.

Example: Let $\boldsymbol{X}$ be a $n \times m$ matrix, then $\boldsymbol{X}^\top\boldsymbol{X}$ and $\boldsymbol{X}\boldsymbol{X}^\top$ are symmetric and positive matrices.
(think of $\boldsymbol{X}^\top\boldsymbol{X}$ as a covariance matrix).

Quadratic Forms

Consider $\boldsymbol{M} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$,

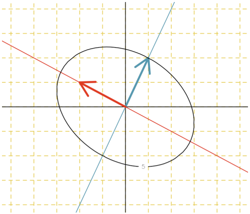
and function $\boldsymbol{z} \mapsto \boldsymbol{z}^\top \boldsymbol{M} \boldsymbol{z}$, i.e.

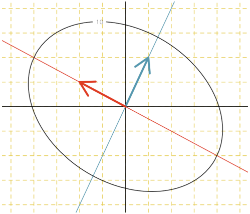
$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

or $ax^2 + 2bxy + cy^2$ is a quadratic form.

If $\det(\boldsymbol{M}) > 0$, points $\boldsymbol{z} = (x, y)$ such that $\boldsymbol{z}^\top \boldsymbol{M} \boldsymbol{z} = \gamma$, for some $\gamma > 0$, are on an ellipse (centered on $\boldsymbol{0}$)

Let $\lambda_1 \geq \lambda_2 > 0$ denote the eigenvalues of $\boldsymbol{M}$ and $\vec{\boldsymbol{v}}_1$ and $\vec{\boldsymbol{v}}_2$ denote the eigenvectors.





Quadratic Forms

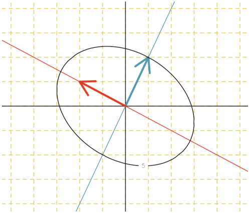
On the picture, $\boldsymbol{M} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.9 \end{pmatrix}$

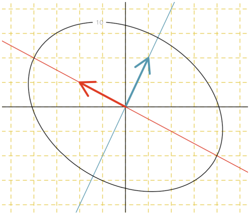
in Python

```
> import numpy as np
> from numpy import linalg
> M = np.array([[.6,.2], [.2, .9]])
> l, v = linalg.eig(M)
> print(l)
[ 0.5   1. ]
> print(v)
[[-0.89442719 -0.4472136 ]
 [ 0.4472136  -0.89442719]]
```

i.e. $\lambda_1 = 1/2$ and $\lambda_2 = 1$, and

$$\vec{\boldsymbol{v}}_1 = \begin{pmatrix} -2\sqrt{5} \\ \sqrt{5} \end{pmatrix} = \sqrt{5}\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \ \vec{\boldsymbol{v}}_2 = \sqrt{5}\begin{pmatrix} -1 \\ -2 \end{pmatrix}$$





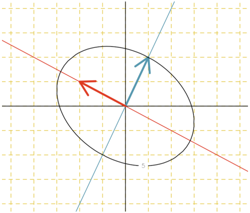
Quadratic Forms

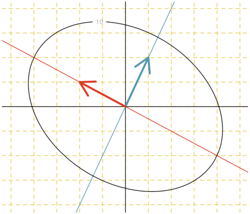
On the picture, $\boldsymbol{M} = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.9 \end{pmatrix}$

in R

```
> M=matrix(c(.6,.2,.2,.9),2,2)
> eigen(M)
eigen() decomposition
$values
[1] 1.0 0.5
$vectors
          [,1]       [,2]
[1,] 0.4472136 -0.8944272
[2,] 0.8944272  0.4472136
```

Note that $\|\vec{\boldsymbol{v}}_1\| = \|\vec{\boldsymbol{v}}_2\| = 1$ and $\vec{\boldsymbol{v}}_1 \perp \vec{\boldsymbol{v}}_2$





Spectral & Singular Value Decomposition

Let $\boldsymbol{M}$ denote a symmetric $d \times d$, with eigenvalues $\lambda_1, \cdots, \lambda_d$ and eigenvectors $\vec{\boldsymbol{u}}_1, \cdots, \vec{\boldsymbol{u}}_d$ with $\|\vec{\boldsymbol{u}}_j\| = 1$,

$$\boldsymbol{M} = \underbrace{\begin{pmatrix} | & | & & | \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_d \\ | & | & & | \end{pmatrix}}_{d\times d} \underbrace{\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix}}_{d\times d} \underbrace{\begin{pmatrix} -\boldsymbol{u}_1^\top - \\ -\boldsymbol{u}_2^\top - \\ \vdots \\ -\boldsymbol{u}_d^\top - \end{pmatrix}}_{d\times d}$$

i.e. $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$.

Let $\boldsymbol{M}$ denote $p \times q$, with $p \leq q$, we can write

$$\boldsymbol{M} = \underbrace{\begin{pmatrix} | & | & & | \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_p \\ | & | & & | \end{pmatrix}}_{p\times p} \underbrace{\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}}_{p\times p} \underbrace{\begin{pmatrix} -\boldsymbol{v}_1^\top - \\ -\boldsymbol{v}_2^\top - \\ \vdots \\ -\boldsymbol{v}_p^\top - \end{pmatrix}}_{p\times q}$$

i.e. $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^\top$, with $\boldsymbol{u} \in \mathbb{R}^p$ and $\boldsymbol{v} \in \mathbb{R}^q$

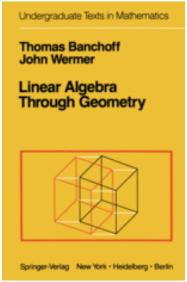
Approximations

Consider $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^\top$ et $\tilde{\boldsymbol{M}} = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}\boldsymbol{V}^\top$ where $\tilde{\boldsymbol{\Lambda}} = \text{diag}(\lambda_1\lambda_2, \boldsymbol{0})$.

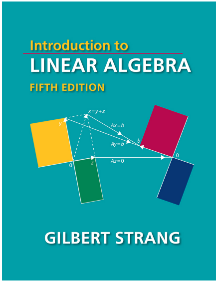
$$\begin{pmatrix} 91 & 107 & 136 & 4 \\ 91 & 82 & 86 & 81 \\ 31 & 33 & -52 & 300 \\ 105 & 86 & 97 & 28 \\ 84 & 80 & 78 & 76 \\ 96 & 97 & 124 & 15 \end{pmatrix}^\top \simeq \begin{pmatrix} 103.5 & 101.4 & 130.5 & 2.2 \\ 87.9 & 85.7 & 85.6 & 80.9 \\ 33.5 & 31.7 & -53.0 & 299.7 \\ 91.1 & 89.1 & 105.6 & 30.8 \\ 82.0 & 80.0 & 79.6 & 76.5 \\ 99.3 & 97.2 & 121.3 & 14.1 \end{pmatrix}^\top$$

The old dataset $\boldsymbol{M}$ is $n = 6$ obsevations, in dimension 4
The new dataset $\tilde{\boldsymbol{M}}$ is $n = 6$ obsevations, in (real) dimension 2
(see principal components #271)

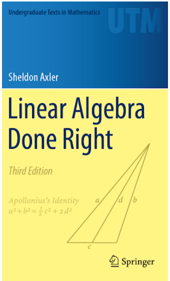
References



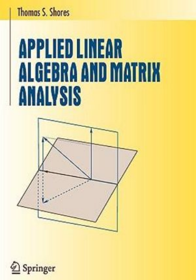
Undergraduate Texts in Mathematics
Thomas Banchoff
John Wermer
Linear Algebra Through Geometry
Springer-Verlag New York · Heidelberg · Berlin



Introduction to
LINEAR ALGEBRA
FIFTH EDITION
GILBERT STRANG



Undergraduate Texts in Mathematics
UTM
Sheldon Axler
Linear Algebra Done Right
Third Edition
Apollonius's Identity
Springer



Thomas S. Shores
APPLIED LINEAR ALGEBRA AND MATRIX ANALYSIS
Springer