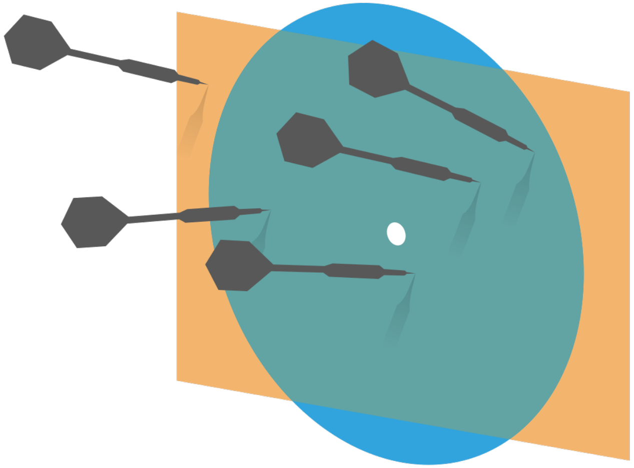
Probability and Statistics for Computer Science



$$cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$
$$= E[XY] - E[X]E[Y]$$

Covariance is coming back in matrix!

Credit: wikipedia

Last time

- Review of Maximum likelihood Estimation (MLE) $L(\theta)$

Likelihood func. is Probability, NOT a distri!!

$\int_\theta L(\theta)d\theta \neq 1$

- Bayesian Inference (MAP)

$$\boxed{P(\theta|D) = \frac{\overbrace{P(D|\theta)}^{L(\theta)}P(\theta)}{P(D)}}$$

Bayesian Posterior IS a distri.!!

θ is considered R.V. in Bayesian Inference.

Objectives

Recap of Bayesian Inference
Conjugate priors

Visualize & Summarize high dimensional data sets

Covariance Matrix

Beta distribution

- A distribution is Beta distribution if it has the following pdf:

$\alpha > 0 \quad \beta > 0$

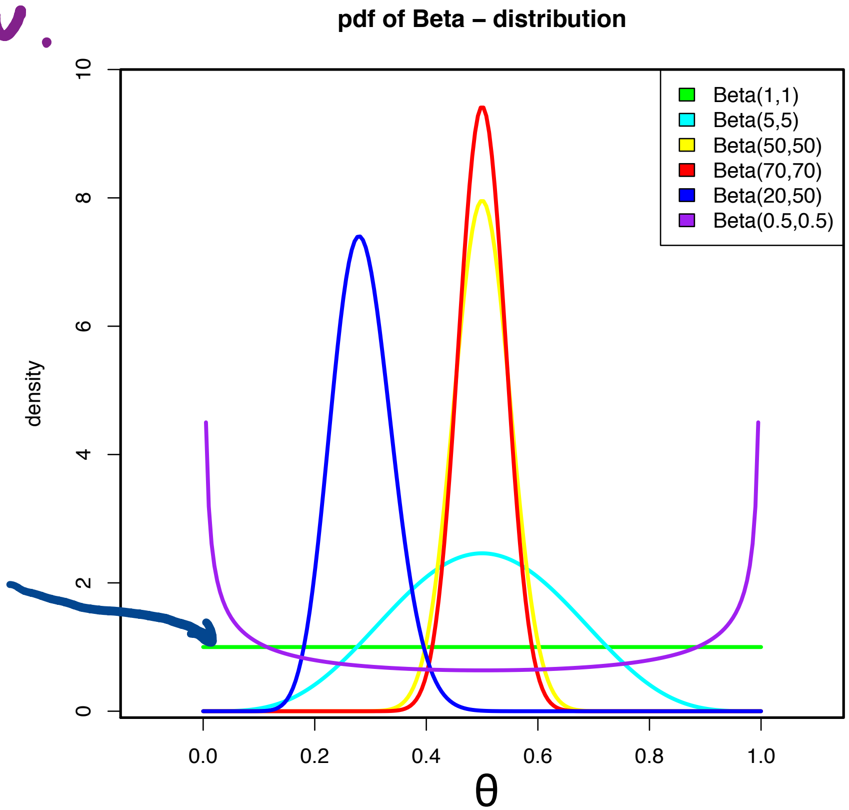
$\theta \in (0, 1)$

$$P(\theta) = \begin{cases} K(\alpha,\beta)\theta^{\alpha-1}(1-\theta)^{\beta-1} \\ 0 \qquad \text{o.w.} \end{cases}$$

$$K(\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

- Is an expressive family of distributions $E[\theta] = \frac{\alpha}{\alpha+\beta}$

- $Beta(\alpha = 1, \beta = 1)$ is uniform



Beta distribution as the conjugate prior for Binomial likelihood

- The likelihood is Binomial (N, k)

$$P(D|\theta) = \binom{N}{k} \theta^k (1-\theta)^{N-k} \quad \sim \theta^{\hat{\alpha}-1} (1-\theta)^{\hat{\beta}-1}$$

- The Beta distribution is used as the prior

$$P(\theta) = K(\alpha, \beta)\theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\hat{\alpha} = \alpha + k$$

$$\hat{\beta} = \beta + N - k$$

- So $P(\theta|D) \propto \theta^{\alpha+k-1}(1-\theta)^{\beta+N-k-1}$

- Then the posterior is $Beta(\boxed{\alpha + k,} \boxed{\beta + N - k)}$

$$P(\theta|D) = \underbrace{K(\alpha + k, \beta + N - k)\theta^{\alpha+k-1}(1-\theta)^{\beta+N-k-1}}_{\int_0^1 \quad d\theta = 1}$$

The update of Bayesian posterior

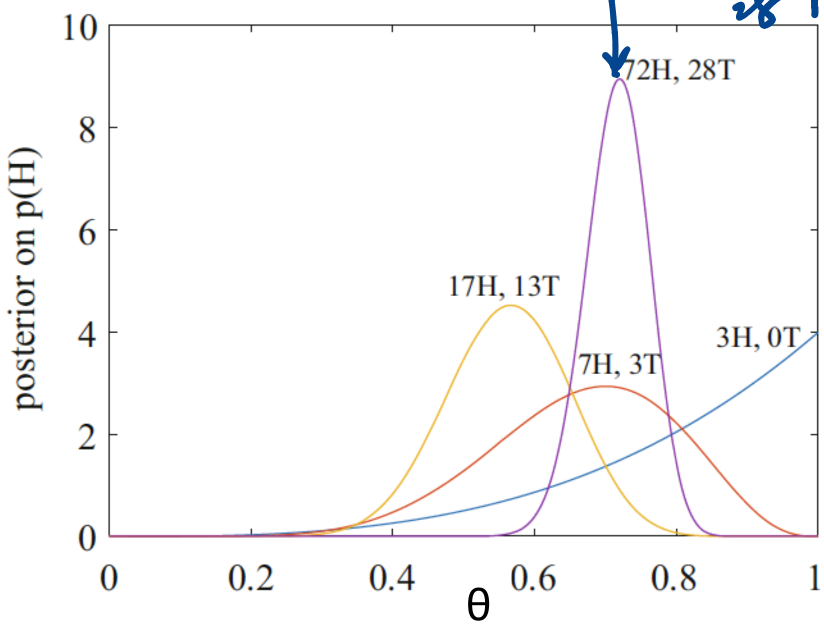
- Since the posterior is in the same family as the conjugate prior, the posterior can be used as a new prior if more data is observed.

- Suppose we start with a uniform prior on the probability θ of heads

$p(\theta \mid D)$

N	k	$\hat{\alpha}$	$\hat{\beta}$
		1	1
3	0	1	4
10	7	8	7
30	17	25	20
100	72	97	48

D : 72H 28T



Maximize the Bayesian posterior (MAP)

- The posterior of the previous example is $E[\theta|D]$ $=?$

$$P(\theta|D) = K(\alpha+k, \beta+N-k)\theta^{\alpha+k-1}(1-\theta)^{\beta+N-k-1}$$

$$\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} = \frac{\alpha+k}{\alpha+\beta+N}$$

$$\frac{d\,P(\theta|D)}{d\,\theta} = 0$$

- Differentiating and setting to 0 gives the MAP estimate

$$\hat{\theta} = \frac{\alpha - 1 + k}{\alpha + \beta - 2 + N}$$

if $\alpha = \beta = 1$

$$\hat{\theta} = \frac{k}{N}$$

$= \text{MLE}$

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\ P(\theta|D)$$

Table of conjugate prior for different likelihood functions

$L(\theta)$
$= P(D|\theta)$

Likelihood	Conjugate Prior
Bernoulli Geometric Binomial	Beta distri.
Poisson Exponential	Gamma distri.
Normal with known σ^2	Normal distri.

$P(\theta)$

Which distri. is the posterior?

if the likelihood is Geometric and we use the corresponding conjugate prior.

A) Binomial

$\boxed{\text{B)}}$ Beta

C) Poisson

D) Bernoulli

E) Normal

How many dimensions do you consider high?

A) ≥ 3

B) > 4

C) ≥ 4

D) others

A data set with high dimensions

- Seed data set from the UCI Machine Learning site: data frame in Python

features

	areaA	perimeterP	compactness	lengthKernel	widthKernel	asymmetry	lengthGroove	Label
1	15.26	14.84	0.871	5.763	3.312	2.221	5.22	1
2	14.88	14.57	0.8811	5.554	3.333	1.018	4.956	1
3	14.29	14.09	0.905	5.291	3.337	2.699	4.825	1
4	13.84	13.94	0.8955	5.324	3.379	2.259	4.805	1
5	16.14	14.99	0.9034	5.658	3.562	1.355	5.175	1
6	14.38	14.21	0.8951	5.386	3.312	2.462	4.956	1
7	14.69	14.49	0.8799	5.563	3.259	3.586	5.219	1

...

$d = 7$

Matrix format of a dataset in the textbook

Col 1 - - - - N

Row 1

2

d

← area A

← perimeter P

← compactness

$\boxed{d \times N}$

↓

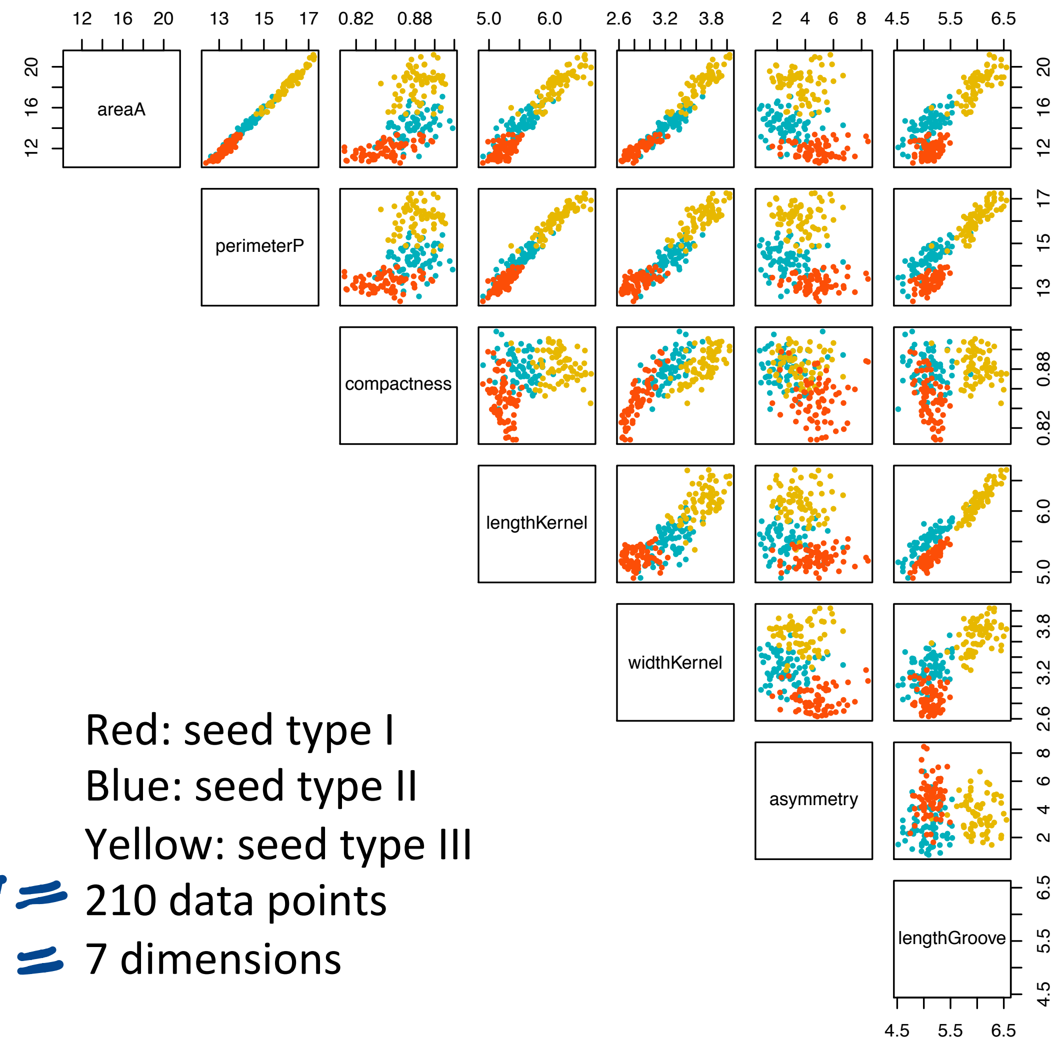
# of features

or # of components

# of dim = d

Scatterplot matrix

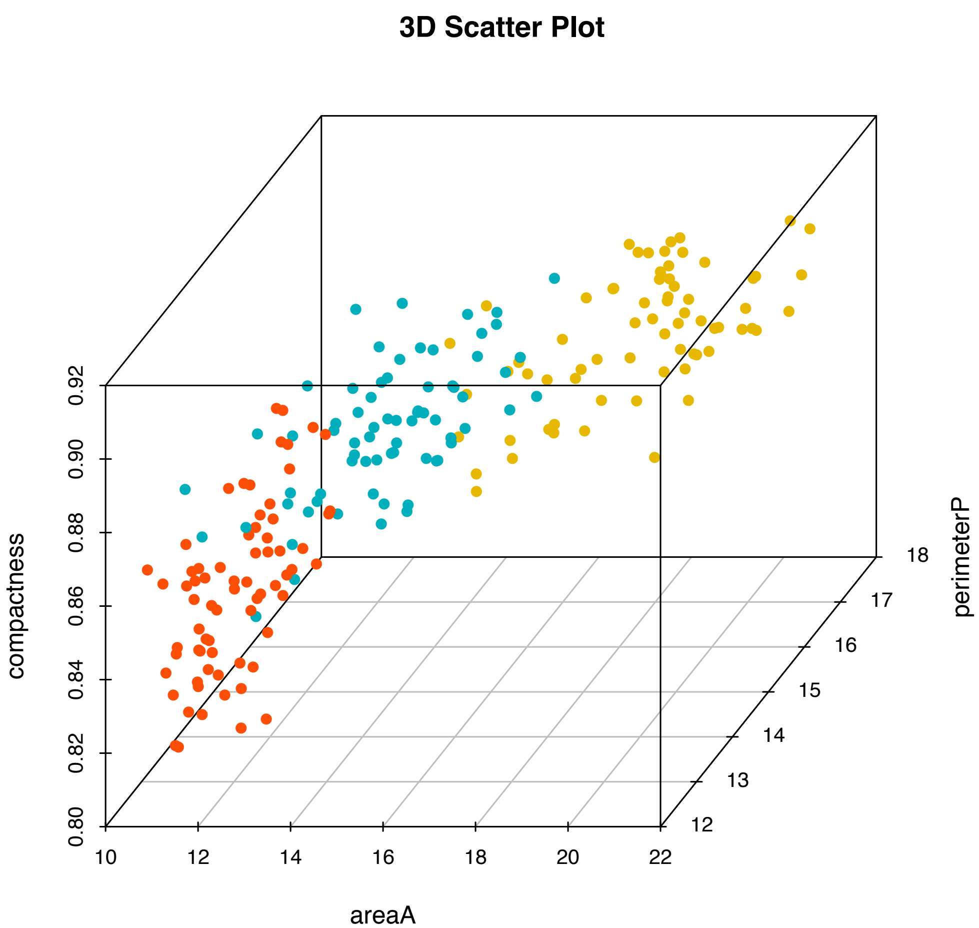
- Visualizing high dimensional data with scatter plot matrix
- Limited to small number of scatter plots



Red: seed type I
Blue: seed type II
Yellow: seed type III
$N =$ 210 data points
$d =$ 7 dimensions

3D scatter plot

- We can also view the data set in 3 dimensions
- But it's still limited in terms of number of dimensions we can see.



Summarizing multidimensional data

- Location and spread parameters of a data set

- Notation
  - Write {**x**} for a dataset consisting of N data items
  - Each item x_i is a **d**-dimensional vector; column
  - Write jth component of x_i as $x_i^{(j)}$; row
  - Matrix for the data set {**x**} is **d** by **N** dimension

Mean of a multidimensional data

- We compute the mean of {x} by computing the mean of each component separately and stacking them to a vector

$$\text{mean of jth component} = \frac{\sum_i x_i^{(j)}}{N}$$

- We write the mean of {x} as

$$mean(\{x\}) = \frac{\sum_i x_i}{N}$$

Example of mean of a multidimensional data set

areaA

	1	2	N = 3	Mean
feature 1	1	2	3	$\frac{6}{3} = 2$
	-1	0	1	0
	3	7	5	5

Mean-Centering a data matrix

Raw

			Mean
1	2	3	2
-1	0	1	0
3	7	5	5

Mean centered

-1	0	1
-1	0	1
-2	2	0

Covariance

- The **covariance** of random variables X and Y is

$$cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

- Note that

$$cov(X,X) = E[(X - E[X])^2] = var[X]$$

Correlation coefficient is normalized covariance

- The correlation coefficient is

$$corr(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

dot prod.

$$= \sum \frac{\hat{x}}{\sqrt{N}} \cdot \frac{\hat{y}}{\sqrt{N}}$$

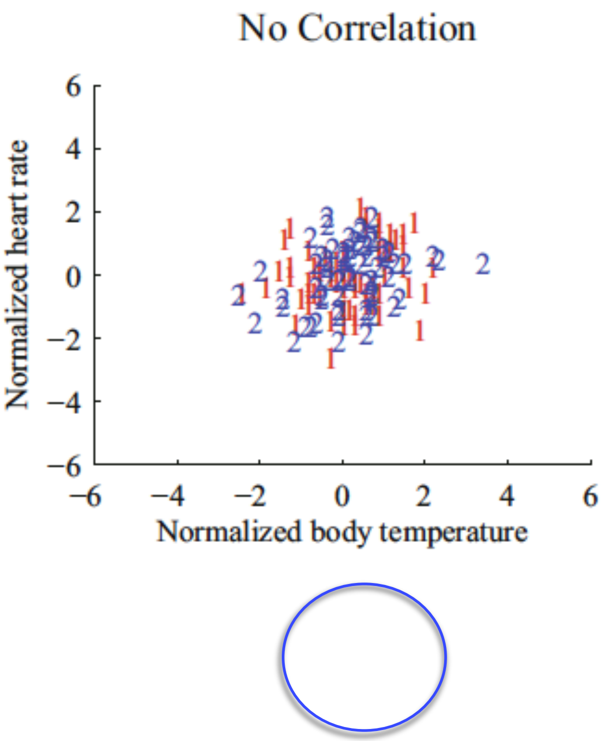
inner prod.

- When X, Y takes on values with equal probability to generate data sets $\{(x,y)\}$, the correlation coefficient will be as seen in Chapter 2.

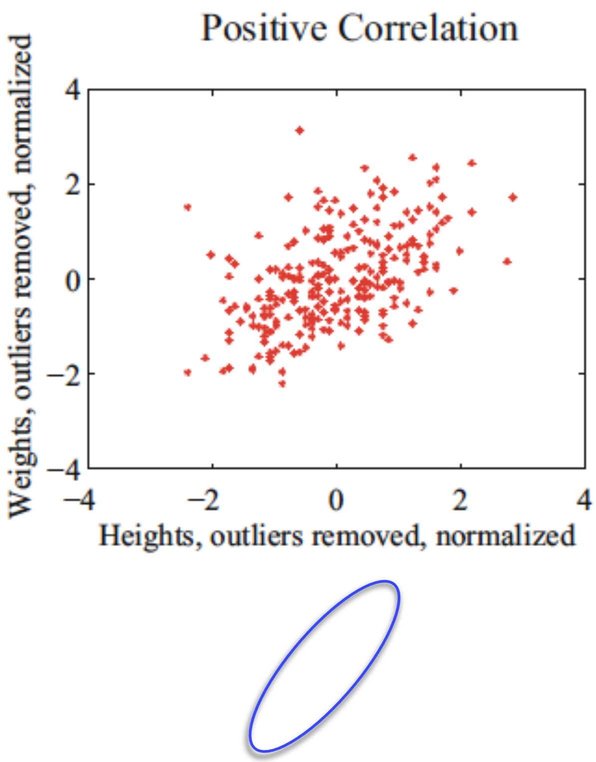
$$corr(\{(x,y)\}) = \frac{\sum \hat{x}\hat{y}}{N}$$

Covariance seen from scatter plots

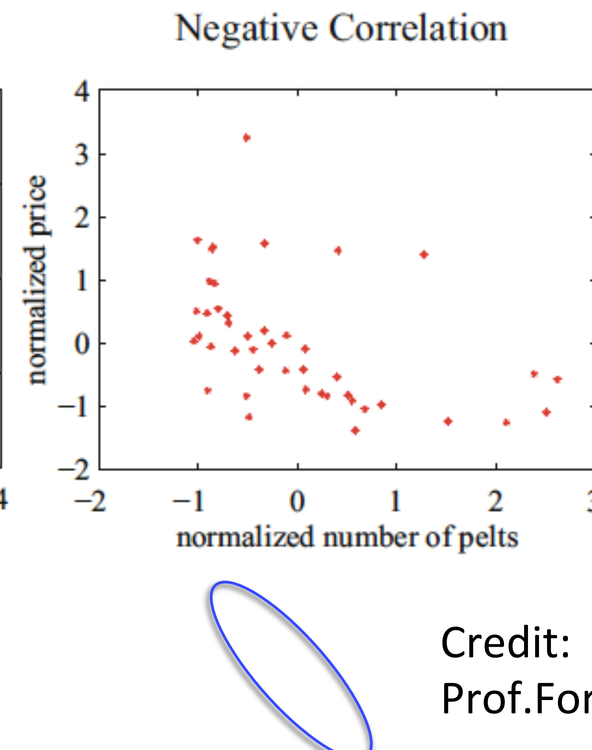
Zero
Covariance



Positive
Covariance



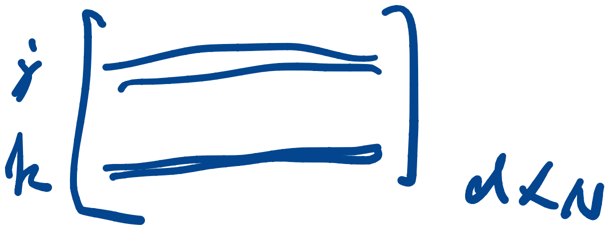
Negative
Covariance



Credit:
Prof.Forsyth

Covariance for a pair of components in a data set

- For the jth and kth components of a data set {x}



$$\frac{cov(\{x\};j,k)}{\sigma_j \, \sigma_k} = \frac{\sum_i (x_i^{(j)} - mean(\{x^{(j)}\}))(x_i^{(k)} - mean(\{x^{(k)}\}))}{\sigma_j \quad N \quad \sigma_k}$$

$LHS = corr(\{x\}; j, k)$

$$RHS = \frac{\sum \hat{x}_i^{(j)} (\hat{x}_i^{(k)})}{N}$$

Covariance of a pair of components

Data set $\{\mathbf{x}\}$ 7×8

$d=7$
$N=8$

$cov(\{\mathbf{x}\}; 3, 5)$

	1	2	3	4	5	6	7	8
1	*	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*	*
4	*	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*	*

Take each row (component) of a pair and subtract it by the row mean, then do the inner product of the two resulting rows and divide by the number of columns

Covariance of a pair of components

Data set $\{\mathbf{x}\}$ 7×8

$d=7$ $N=8$

$cov(\{\mathbf{x}\}; 3, 5)$

	1	2	3	4	5	6	7	8
1	*	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*	*
4	*	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*	*

area

length

How many pairs of rows are there for which we can compute the covariance?

A) 49
B) 64
C) 56

7×7

$d=7$

pair of rows

$=7\times7$

Covariance matrix

Data set $\{\mathbf{x}\}$ 7×8

Covmat($\{\mathbf{x}\}$) 7×7

$cov(\{\mathbf{x}\}; 3, 5)$

	1	2	3	4	5	6	7	8
1	*	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*	*
4	*	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*	*

	1	2	3	4	5	6	7
1	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*
4	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*

$cov(\{x\}, j.k) = \underline{\sigma_x \sigma_y} \cdot corr(j,k)$

$k \quad j)$ $\qquad (k,j)$

Properties of Covariance matrix

$$cov(\{x\}; j, j) = var(\{x^{(j)}\})$$

Covmat$(\{\mathbf{x}\})$ 7×7

- The diagonal elements of the covariance matrix are just variances of each jth components
- The off diagonals are covariance between different components

	1	2	3	4	5	6	7
1	* σ_1^2	*	*	*	*	*	*
2	*	* σ_2^2	*	*	*	*	*
3	*	*	* σ_3^2	*	*	*	*
4	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*

Properties of Covariance matrix

$$cov(\{x\}; j, k) = cov(\{x\}; k, j)$$

Covmat($\{\mathbf{x}\}$) 7×7

- The covariance matrix is **symmetric**!
- And it's **positive semi-definite**, that is all $\lambda_i \geq 0$
- Covariance matrix is diagonalizable

	1	2	3	4	5	6	7
1	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*
4	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*

Properties of Covariance matrix

- If we define $\mathbf{x}_c$ as the mean centered matrix for dataset {x}

$$Covmat(\{x\}) = \frac{X_c X_c^T}{N}$$

$u \cdot u^T =$ inner prod

$(u, u^T) = |u|^2$

- The covariance matrix is a d×d matrix

$\mathrm{Covmat}(\{\mathbf{x}\})$ 7×7

	1	2	3	4	5	6	7
1	*	*	*	*	*	*	*
2	*	*	*	*	*	*	*
3	*	*	*	*	*	*	*
4	*	*	*	*	*	*	*
5	*	*	*	*	*	*	*
6	*	*	*	*	*	*	*
7	*	*	*	*	*	*	*

$d \times d$

d =7

Example: covariance matrix of a data set

(I) $N=5$ $d=2$

$$A_0 = \begin{array}{cc} \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix} & \begin{array}{c} X^{(1)} \\ X^{(2)} \end{array} \end{array}$$

$d=2$

2×2

What are the dimensions of the covariance matrix of this data?

A) 2 by 2
B) 5 by 5
C) 5 by 2
D) 2 by 5

Example: covariance matrix of a data set

Mean centering

(I)

$$A_0 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$$

Mean

3

0

$$A_1 = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$$

Example: covariance matrix of a data set

(I)

Mean centering

$$A_0 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$$

(II) $A_2 = A_1 A_1^T$

Inner product of each pairs:

A_2 [1,1] = 10

A_2 [2,2] = 4

A_2 [1,2] = 0

Example: covariance matrix of a data set

(I)

Mean centering

$$A_0 = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$$

(II) $A_2 = A_1 A_1^T$

Inner product of each pairs:

A_2 [1,1] = 10

A_2 [2,2] = 4

A_2 [1,2] = 0

(III)

Divide the matrix with N – the number of items

$$\text{Covmat}(\{\mathbf{x}\}) = \frac{1}{N} A_2 = \frac{1}{5} \begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix} = \boxed{\begin{bmatrix} 2 & 0 \\ 0 & 0.8 \end{bmatrix}}$$

Translation properties of mean and covariance matrix

- Translating the data set translates the mean

$$mean(\{x\} + c) = mean(\{x\}) + c$$

- Translating the data set leaves the covariance matrix unchanged

$$Covmat(\{x\} + c) = Covmat(\{x\})$$

Translation properties of covariance matrix

- Proof:

$$\text{Covmat}(\{x\}) = \frac{X_c X_c^T}{N}$$

$X_c \rightarrow$ doesn't change

if $\{x\}$ is translated!

$\because$ $x + c - \text{mean}(\{x+c\})$

$= x - \text{mean}(\{x\}) = X_c$

Linear transformation properties of mean and covariance matrix

- Linearly transforming the data set linearly transforms the mean

$$mean(\{A\mathbf{x}\}) = A\ mean(\{\mathbf{x}\})$$

if $mean(\{x\}) = 0$
$mean(\{Ax\}) = 0$

- Linearly transforming the data set linearly changes the covariance matrix quadratically

$\to m \times m$

$$Covmat(\{A\mathbf{x}\}) = A\ Covmat(\{\mathbf{x}\})A^T$$

AX $m \times N$

$X \to d \times N$

$A \to m \times d$

$var(cX) = c^2 var(X)$

$A\ X_{d \times N}$ # of col of $A = d$

Proof of linear transformation of covariance matrix

$$\text{Covmat}(\{x\}) = \frac{X_c X_c^T}{N}$$

Suppose $X = X_c$

data matrix X is centered

$AX = AX_c$

$$\text{covmat}(\{Ax\}) = \frac{(AX)_c[(AX)_c]^T}{N}$$

if X_c is centered

AX_c is centered

$$= \frac{AX_c(AX_c)^T}{N}$$

$(BC)^T$

$= C^T B^T$

$$= \frac{A X_c \cdot X_c^T A^T}{N}$$

$= A \cdot \text{covmat}(\{x\}) \cdot A^T$

Dimension Reduction

- In stead of showing more dimensions through visualization, it's a good idea to do dimension reduction in order to see the major features of the data set.
- For example, principal component analysis help find the major components of the data set.
- PCA is essentially about finding eigenvectors of covariance matrix

Why linear algebra?

- We are entering into part **IV** of the course. The contents will be basic machine learning techniques.
- Linear algebra is essential for a lot of machine Learning methods!

Eigenvalues and eigenvectors review

- If A is an **n×n** square matrix, an eigenvalue λ and its corresponding eigenvector v (of dimension n×1) satisfy $Av = \lambda v$.

- To solve for λ, we solve the characteristic equation

  $$|A - \lambda I| = 0$$

- Given a value of λ, we solve v by solving

  $$(A - \lambda I)\, v = 0$$

- Note if v is an eigenvector, then so is any multiple kv.

Eigenvalues and eigenvectors example

- Find the eigenvalues and eigenvectors

$$A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

What's special of this A?

symmetric

$$|A - \lambda I| = 0$$

$$\begin{vmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{vmatrix} = 0$$

$$(5-\lambda)(5-\lambda) - 9 = 0$$

$$(\lambda - 8)(\lambda - 2) = 0$$

$$\lambda_1 = 8$$

$$\lambda_2 = 2$$

positive definite

$$\lambda_i > 0$$

Eigenvalues and eigenvectors example

- Find the eigenvectors

$$A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

$$(A - \lambda I)v = 0$$

$$\lambda_1 = 8 \qquad \begin{bmatrix} 5-8 & 3 \\ 3 & 5-8 \end{bmatrix} v = 0$$

$$\begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} v = 0$$

$$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 2 \qquad \begin{bmatrix} 5-2 & 3 \\ 3 & 5-2 \end{bmatrix} v = 0$$

$$\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} v = 0$$

$$v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$u_1 = \frac{1}{\|v_1\|} v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$u_2 = \frac{1}{\|v_2\|} v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvalues and eigenvectors example (2)

- Find the eigenvalues and eigenvectors of

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$ 2×row 1

$|A - \lambda I| = 0$

$\left| \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0$

$\left| \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} \right| = 0$

$(1-\lambda)(4-\lambda) - 4 = 0$

$(\lambda - 5) \cdot \lambda = 0$

$\lambda_1 = 5, \quad \lambda_2 = 0$

What's special of this A?
Symmetric & singular

$\det(A) = \prod \lambda_i = 0$

$\lambda_i \geq 0$

positive semi-definite

Eigenvalues and eigenvectors example

- Find the eigenvectors of

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$\lambda_1 = 5$

$(A - \lambda_1 I)v_1 = 0$

$(A - 5I)v_1 = 0 \Rightarrow \begin{bmatrix} 1-5 & 2 \\ 2 & 4-5 \end{bmatrix} v_1 = 0$

$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow u_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\lambda_2 = 0$

$A v_2 = 0$

$v_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \Rightarrow u_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

Diagonalization of a symmetric matrix

- If A is an n×n symmetric square matrix, the eigenvalues are real.
- If the eigenvalues are also distinct, their eigenvectors are orthogonal
- We can then scale the eigenvectors to unit length, and place them into an orthogonal matrix $U = [\mathbf{u}_1\ \mathbf{u}_2\ \dots .\ \mathbf{u}_n]$
- We can write the diagonal matrix $\Lambda = U^T A U$ such that the diagonal entries of Λ are $\lambda_1, \lambda_2 \dots \lambda_n$ in that order.

why do we do this?

Diagonalization example

- For

$$A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

$\lambda_1 = 8 \qquad u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = 2 \qquad u_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$\underbrace{\begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}}_{\Lambda} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{U^T} \underbrace{\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}}_{A} \underbrace{\begin{array}{cc} u_1 & u_2 \\ \downarrow & \downarrow \\ \end{array}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}}_{U}$$

Q. Are these two vectors orthogonal?

$V_1 = [3\ 6]$, $V_2 = [-2\ 1]$

$\boxed{\text{A. Yes}}$

B. No

$3 \times (-2) + 6 \times 1 = 0$

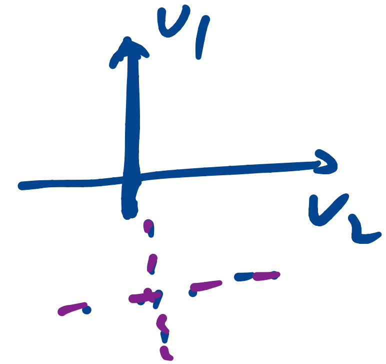
$\sum V_{1i} \cdot V_{2i} = 0$

$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \& \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ orthogonal

$\underline{\text{dot prod}\ (V_1, V_2) = 0}$

Q. Is this true?

When two zero-mean vectors of data are orthogonal, they are uncorrelated



$\text{mean}[v_i] = 0$

A. Yes

B. No

$$\sum (x - \text{mean}(\{x\}))(y - \text{mean}(\{y\})) = 0$$

$\hat{x}$ $\hat{y}$

$$\frac{\sum \hat{x}\hat{y}}{N}$$

Q. Is this true?

When two zero-mean vectors of data are orthogonal, they are uncorrelated

A. Yes

B. No

Assignments

- Read Chapter 10 of the textbook
- Next time: PCA

Additional References

- Robert V. Hogg, Elliot A. Tanis and Dale L. Zimmerman. “Probability and Statistical Inference”
- Morris H. Degroot and Mark J. Schervish "Probability and Statistics”

See you next time

See You!