

Reminder

Multiple Linear Regression

(Chapters 12-13 in
Montgomery, Runger)

12-1: Multiple Linear Regression Model

12-1.1 Introduction

- Many applications of regression analysis involve situations in which there are more than one regressor variable X_k used to predict Y .
- A regression model then is called a **multiple regression model**.

Multiple Linear Regression Model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + \varepsilon$$

One can also use powers and products of other variables or even non-linear functions like $\exp(x_i)$ or $\log(x_i)$

instead of x_3, \dots, x_k .

Example: the general two-variable quadratic regression has 6 constants:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 (x_1)^2 + \beta_4 (x_2)^2 + \beta_5 (x_1 x_2) + \varepsilon$$

12-1: Multiple Linear Regression Model

12-1.3 Matrix Approach to Multiple Linear Regression

Suppose the model relating the regressors to the response is

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \quad i = 1, 2, \dots, n$$

In matrix notation this model can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (12-6)$$

12-1: Multiple Linear Regression Model

12-1.3 Matrix Approach to Multiple Linear Regression

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

12-1.3 Matrix Approach to Multiple Linear Regression

We wish to find the vector $\hat{\beta}$ that minimizes the sum of squares of error terms:

$$L = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (\mathbf{y} - \mathbf{X}\beta)' (\mathbf{y} - \mathbf{X}\beta)$$

$$0 = \frac{\partial L}{2\partial \beta} = -\mathbf{X}' (\mathbf{y} - \mathbf{X}\beta) = -\mathbf{X}' \mathbf{y} + (\mathbf{X}' \mathbf{X}) \beta$$

The resulting least squares estimate is

$$\hat{\beta} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{y} \quad (12-7)$$

Analog of $\frac{1}{\text{Var}(x)}$

Analog of $\text{Cov}(x, y)$

Multiple Linear Regression Model

$$\hat{\beta} = (X'X)^{-1} X'y$$

H is an idempotent matrix

$$\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y,$$

$$\hat{y} = Hy, \quad \text{and} \quad e = (I - H)y.$$



$$H = H^2; \quad H^2 = X \underbrace{(X'X)^{-1} X' X (X'X)^{-1}}_I X = X(X'X)^{-1} X' = H$$

Vectors \hat{y} & e are orthogonal since

$$\hat{y}'e = y'H(I-H)y = 0 \quad \text{since}$$

$$H(I-H) = H - H^2 = H - H = 0.$$

12-1: Multiple Linear Regression Models

12-1.4 Properties of the Least Squares Estimators

Unbiased estimators:

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon})] \\ &= E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\epsilon}] \\ &= \boldsymbol{\beta} \end{aligned}$$

Covariance Matrix of Estimators:

$$\mathbf{C} = (\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} C_{00} & C_{01} & C_{02} \\ C_{10} & C_{11} & C_{12} \\ C_{20} & C_{21} & C_{22} \end{bmatrix}$$

12-1: Multiple Linear Regression Models

12-1.4 Properties of the Least Squares Estimators

Individual variances and covariances:

$$V(\hat{\beta}_j) = \sigma^2 C_{jj}, \quad j = 0, 1, 2$$
$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = \sigma^2 C_{ij}, \quad i \neq j$$

In general,

$$\text{cov}(\hat{\beta}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} = \sigma^2 \mathbf{C}$$

12-1: Multiple Linear Regression Models

Estimating error variance σ_ε^2

An unbiased estimator of error variance σ_ε^2 is

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n e_i^2}{n-p} = \frac{SS_E}{n-p} \quad (12-16)$$

Here $p=k+1$ for k -variable multiple linear regression

R² and Adjusted R²

The **coefficient of multiple determination R²**

$$R^2 = \frac{SS_R}{SS_T} = 1 - \frac{SS_E}{SS_T}$$

The **adjusted R²** is

$$R_{\text{adj}}^2 = 1 - \frac{SS_E/(n-p)}{SS_T/(n-1)} \quad (12-23)$$

Handwritten red annotations: An arrow points from the denominator of the fraction to the term $\frac{\sigma^2}{\sum y}$ written below it. A horizontal line is drawn under the fraction, and a vertical line is drawn under the denominator.

- The adjusted R² statistic penalizes **adding terms** to the MLR model.
- It can help guard against **overfitting** (including regressors that are not really useful)

How to know where to stop adding variables?

- Adding new variables x_i to MLR
watch the adjusted R^2
- Once the adjusted R^2
no longer increases = stop.
Now you did the best you can.

Credit: XKCD
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

QUESTIONS

FOUND IN GOOGLE AUTOCOMplete



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

WHY AREN'T ECONOMISTS RICH

WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS THERE PHLEGM

WHY DO AMERICANS CALL IT SOCCER

WHY IS PSYCHIC WEAK TO BUG

WHY ARE MY EARS RINGING

WHY DO CHILDREN GET CANCER

WHY ARE THERE SO MANY AVENGERS

WHY IS POSEIDON ANGRY WITH ODYSSEUS

WHY ARE THE AVENGERS FIGHTING THE X MEN

WHY IS THERE ICE IN SPACE

WHY IS WOLVERINE NOT IN THE AVENGERS

WHY ARE THERE ANTS IN MY LAPTOP

WHY IS EARTH TILTED

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL IN MY BACKYARD

WHY IS SPACE BLACK

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS OUTER SPACE SO COLD

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE PYRAMIDS ON THE MOON

WHY ARE THERE GHOSTS

WHY DO OWLS ATTACK PEOPLE

WHY IS NASA SHUTTING DOWN

WHY ARE THERE GHOSTS

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE GHOSTS

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE THERE GODS

WHY DO SPIDERS COME INSIDE

WHY ARE THERE GHOSTS

WHY ARE THERE TWO SPOCKS

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY IS LIFE SO BORING

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE CIGARETTES LEGAL

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY ARE THERE DUCKS IN MY POOL

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY IS JESUS WHITE

WHY DO SPIDER BITES ITCH

WHY ARE THERE GHOSTS

WHY IS THERE LIQUID IN MY EAR

WHY IS DYING SO SCARY

WHY ARE THERE GHOSTS

WHY DO Q TIPS FEEL GOOD

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS

WHY AREN'T THERE DINOSAUR GHOSTS

WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS

WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY

WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT

WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES

WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS

WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD

WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS

WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO

WHY IS OHIO WEATHER SO WEIRD
WHY ARE THERE MALE AND FEMALE BIKES
WHY ARE THERE BRIDESMAIDS

WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KUNGONS DIFFERENT

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER

WHY DO IGUANAS DIE

WHY ARE THERE FEMALE MR NIMES

WHY IS GPS FREE

WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK

WHY IS THERE PHLEGM

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE GODS

WHY IS LIFE SO BORING

WHY ARE ULTRASOUNDS IMPORTANT

WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN

WHY IS PSYCHIC WEAK TO BUG

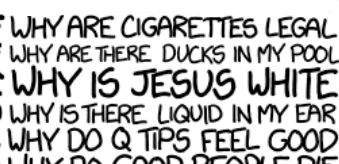
WHY DO CHILDREN GET CANCER

WHY IS THERE ICE IN SPACE

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TWO SPOCKS



WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND

WHY IS SEX SO IMPORTANT



WHY ARE THERE SQUIRRELS



WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL

WHY IS THERE HELL IF GOD FORGIVES

WHY IS THERE HELL IF GOD FORGIVES

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

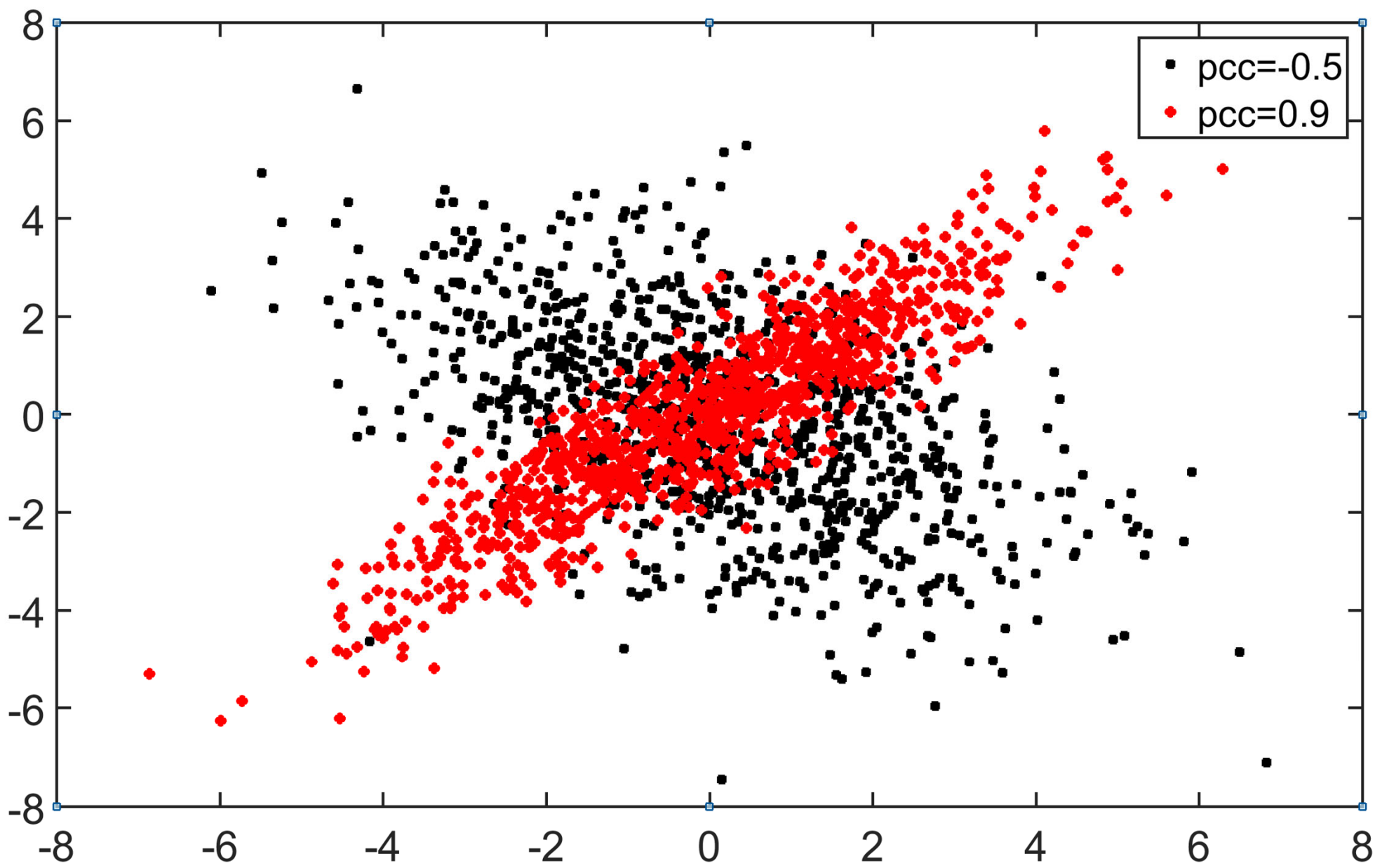
WHY IS GPS FREE

WHY IS GPS FREE

WHY IS GPS FREE

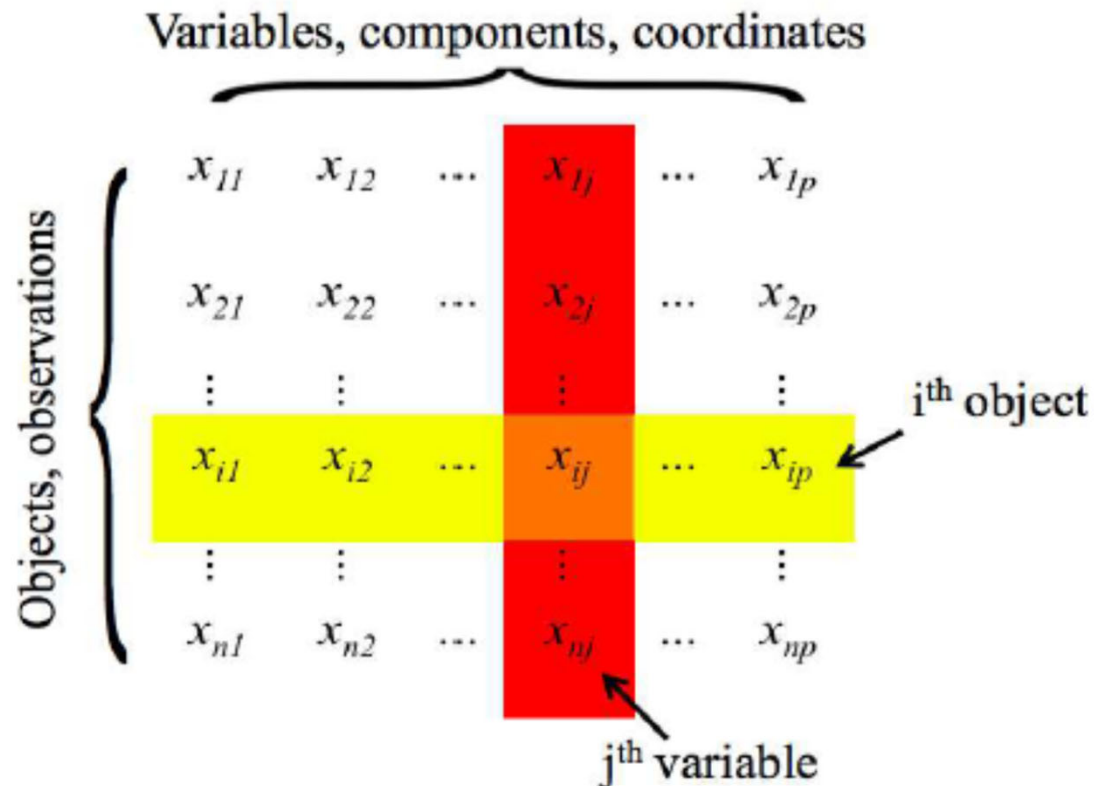
WHY IS GPS FREE

Principal Component Analysis



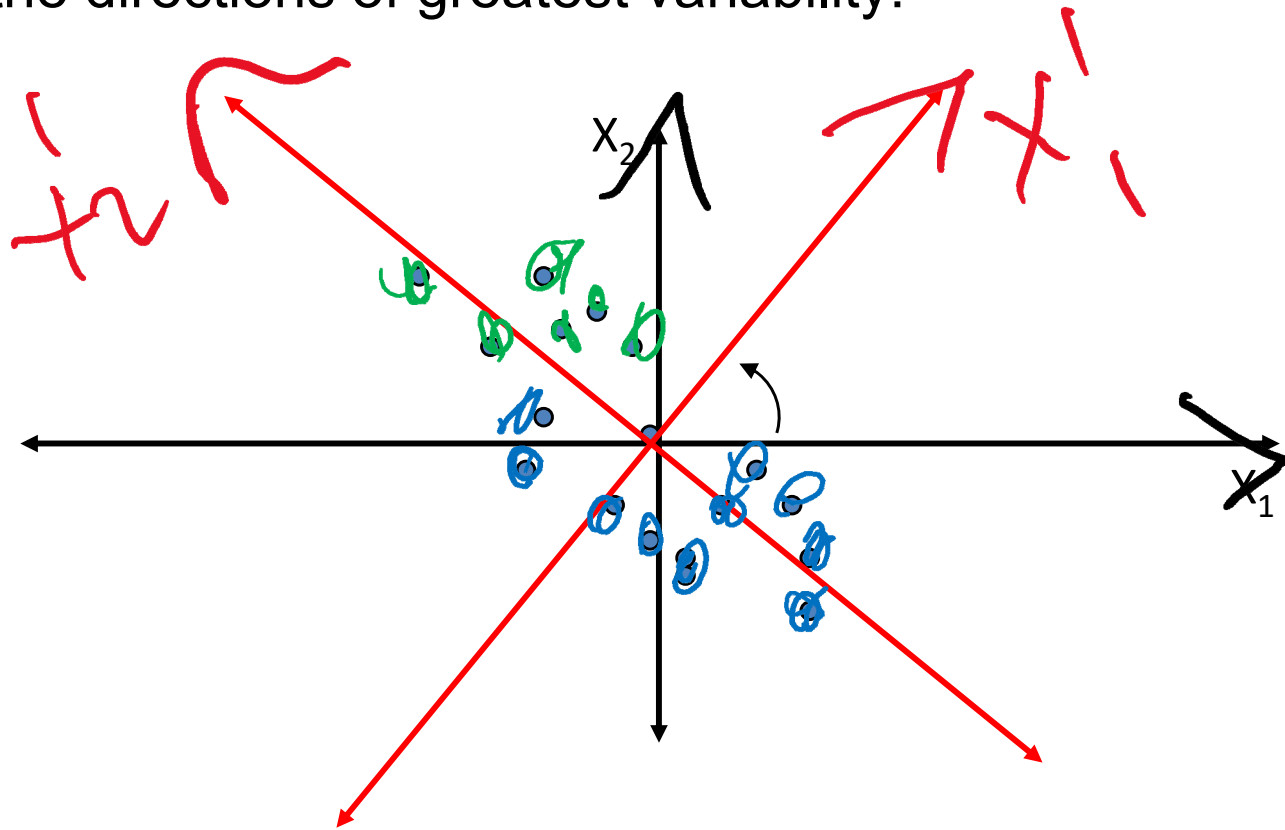
Multivariable statistics and Principal Component Analysis (PCA)

- A table of n observations in which p variables were measured



Trick: Rotate Coordinate Axes

Suppose we have a population measured on p random variables X_1, \dots, X_p . Note that these random variables represent the p -axes of the Cartesian coordinate system in which the population resides. Our goal is to develop a new set of p axes (linear combinations of the original p axes) in the directions of greatest variability:



This is accomplished by rotating the axes.

Applications of PCA

- Uses:
 - Data Visualization
 - Dimensional Reduction
 - Data Classification
- Examples:
 - How many unique “sub-sets” are in the sample?
 - How are they similar / different?
 - What are the underlying factors that most influence the samples?
 - Which measurements are best to differentiate between samples?
 - How to best present what is “interesting”?
 - Which “sub-set” does this new sample rightfully belong?

PCA: *General*

From p original variables: x_1, x_2, \dots, x_p :

Produce k new variables: x'_1, x'_2, \dots, x'_p :

$$x'_1 = v_{11}x_1 + v_{12}x_2 + \dots + v_{1p}x_p$$

$$x'_2 = v_{21}x_1 + v_{22}x_2 + \dots + v_{2p}x_p$$

...

$$x'_p = v_{p1}x_1 + v_{p2}x_2 + \dots + v_{pp}x_p$$

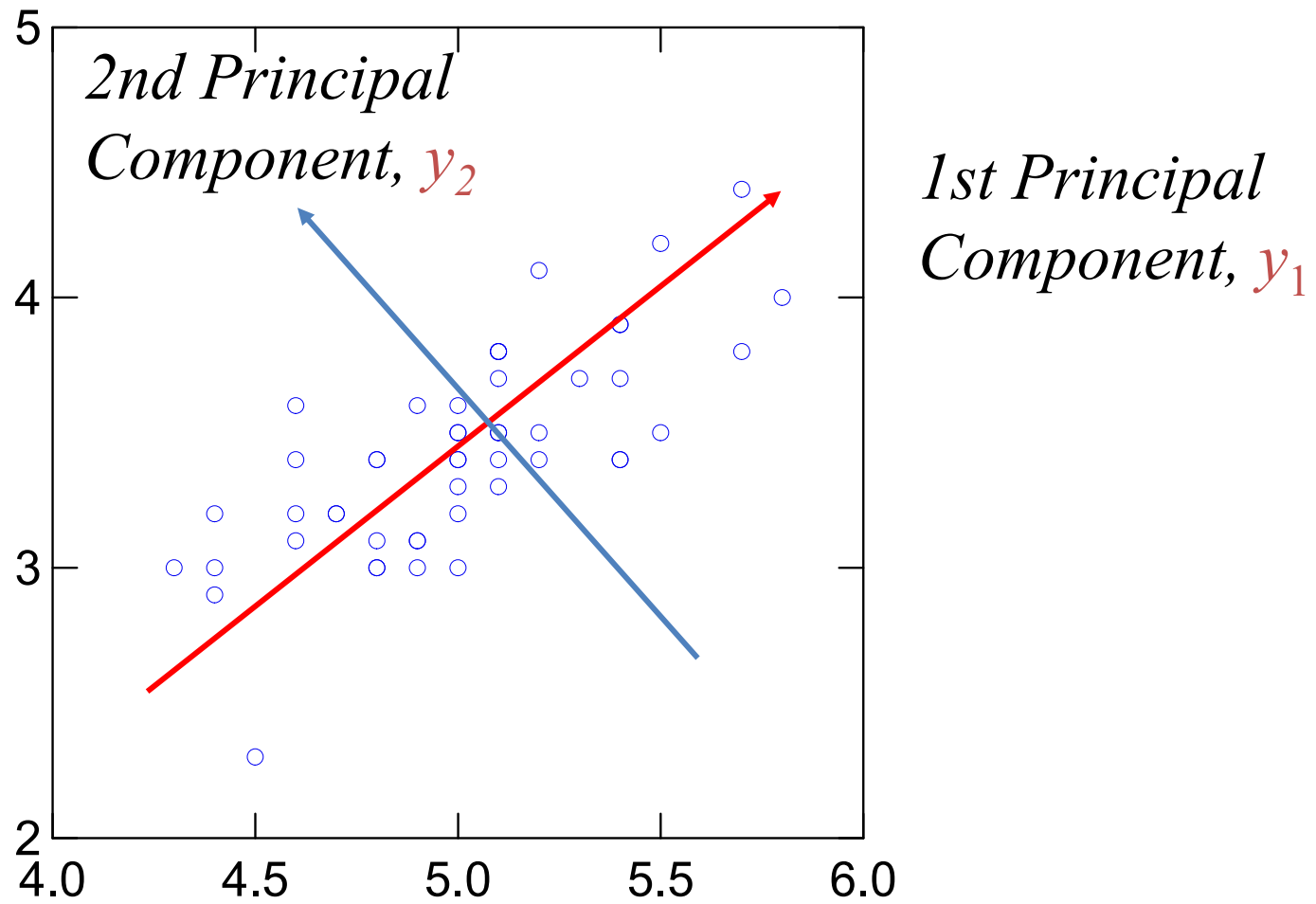
such that:

x'_i 's are uncorrelated (orthogonal)

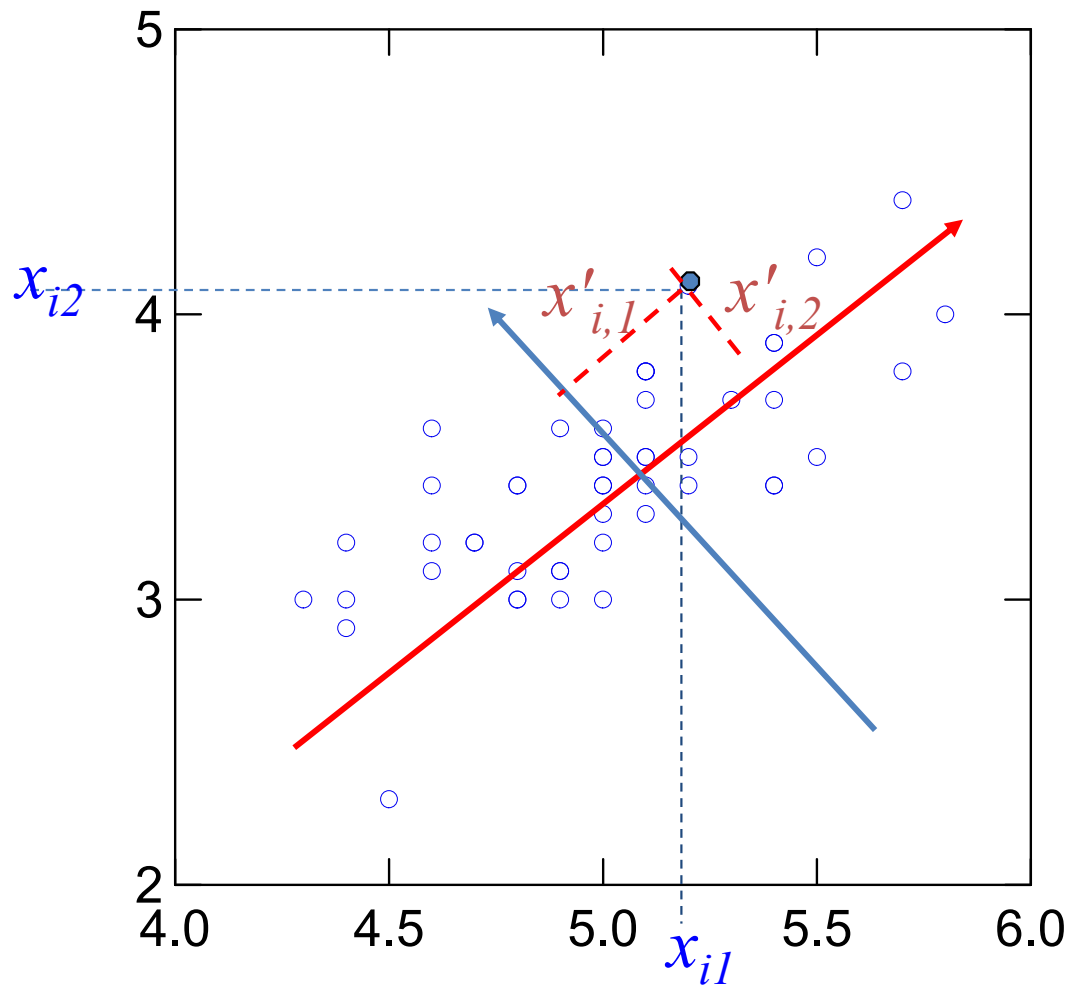
x'_1 explains as much as possible of original variance in data set

x'_2 explains as much as possible of remaining variance

etc.

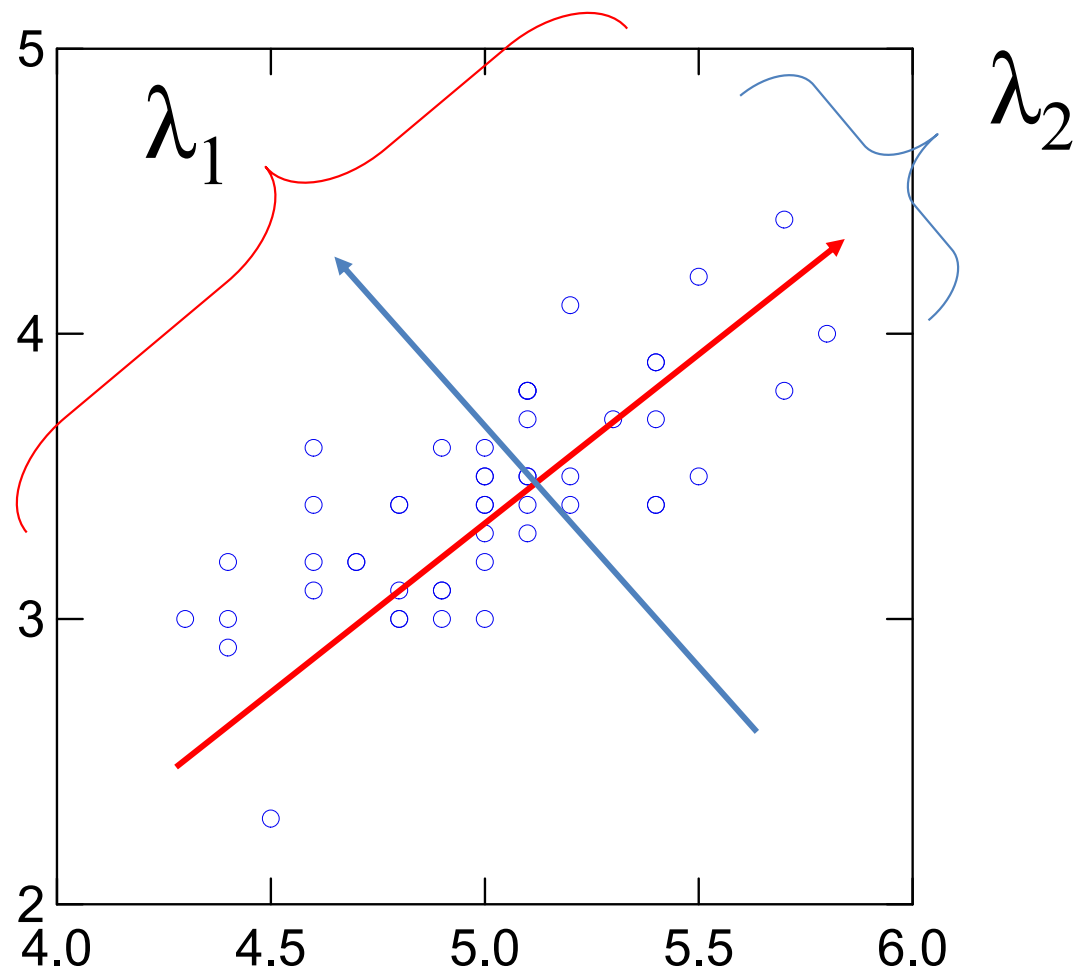


PCA Scores



Adapted from slides by Prof. S. Narasimhan, "Computer Vision" course at CMU

PCA Eigenvalues



Multivariable statistics and Principal Component Analysis (PCA)

- A table of **n observations** in which **p variables** were measured

Variables, components, coordinates

	x_{11}	x_{12}	...	x_{1j}	...	x_{1p}
	x_{21}	x_{22}	...	x_{2j}	...	x_{2p}
	\vdots	\vdots		\vdots		\vdots
Objects, observations	x_{i1}	x_{i2}	...	x_{ij}	...	x_{ip}
	\vdots	\vdots		\vdots		\vdots
	x_{n1}	x_{n2}	...	x_{nj}	...	x_{np}

i^{th} object

j^{th} variable

$p \times p$ symmetric
matrix R of
corr. coefficients

$$r_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$

PCA: Diagonalize
matrix R

Principle Component Analysis (PCA)

- $p \times p$ symmetric matrix R of corr. coefficients $r_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$
- $R = n^{-1} Z' * Z$ is a “square” of the matrix Z of standardized r.v.:
 $Z_{\alpha i} = \frac{x_{\alpha i} - \mu_i}{\sigma_i} \rightarrow$ all eigenvalues of R are non-negative
- Diagonal elements=1 $\rightarrow tr(R)=p$
- Can be diagonalized:
 $R = V * D * V'$ where D is the diagonal matrix
- $d(1,1)$ – largest eig. value, $d(p,p)$ – the smallest one
- The meaning of $V(i,k)$ – contribution of the data type i to the k -th eigenvector
- $tr(D)=p$, the largest eigenvalue $d(1,1)$ absorbs a fraction $=d(1,1)/p$ of all correlations can be $\sim 100\%$
- **Scores**: $X' = Z * V$: $n \times p$ matrix. Meaning of $X'(\alpha, k)$ – participation of the sample # α in the k -th eigenvector

Credit: XKCD
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

QUESTIONS

FOUND IN GOOGLE AUTOCOMLETE



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

WHY AREN'T ECONOMISTS RICH

WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS THERE PHLEGM

WHY DO AMERICANS CALL IT SOCCER

WHY IS PSYCHIC WEAK TO BUG

WHY ARE MY EARS RINGING

WHY DO CHILDREN GET CANCER

WHY ARE THERE SO MANY AVENGERS

WHY IS POSEIDON ANGRY WITH ODYSSEUS

WHY ARE THE AVENGERS FIGHTING THE X MEN

WHY IS THERE ICE IN SPACE

WHY ARE THERE ANTS IN MY LAPTOP

WHY IS EARTH TILTED

WHY ARE THERE GHOSTS



WHY IS THERE AN OWL IN MY BACKYARD

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY DO OWLS ATTACK PEOPLE

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE GODS

WHY ARE THERE TWO SPOCKS

WHY ARE CIGARETTES LEGAL

WHY ARE THERE DUCKS IN MY POOL

WHY IS JESUS WHITE

WHY IS THERE LIQUID IN MY EAR

WHY DO Q TIPS FEEL GOOD

WHY DO GOOD PEOPLE DIE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY DO SPIDERS COME INSIDE

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY DO SPIDER BITES ITCH

WHY IS DYING SO SCARY

WHY IS THERE NO GPS IN LAPTOPS

WHY IS SEX SO IMPORTANT



WHY DO KNEES CLICK

WHY AREN'T THERE E GRADES

WHY DO WHALES JUMP

WHY ARE WITCHES GREEN

WHY ARE THERE MIRRORS ABOVE BEDS

WHY DO I SAY UH

WHY IS SEA SALT BETTER

WHY ARE THERE TREES IN THE MIDDLE OF FIELDS

WHY IS THERE NOT A POKEMON MMO

WHY IS THERE LAUGHING IN TV SHOWS

WHY ARE THERE DOORS ON THE FREEWAY

WHY ARE THERE SO MANY SVCHOST.EXE RUNNING

WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA

WHY ARE THERE SCARY SOUNDS IN MINECRAFT

WHY IS THERE KICKING IN MY STOMACH

WHY ARE THERE TWO SLASHES AFTER HTTP

WHY ARE THERE CELEBRITIES

WHY DO SNAKES EXIST

WHY DO OYSTERS HAVE PEARLS

WHY ARE DUCKS CALLED DUCKS

WHY DO THEY CALL IT THE CLAP

WHY ARE KYLE AND CARTMAN FRIENDS

WHY IS THERE AN ARROW ON AANG'S HEAD

WHY ARE TEXT MESSAGES BLUE

WHY ARE THERE MUSTACHES ON CLOTHES

WHY ARE THERE MUSTACHES ON CARS

WHY ARE THERE MUSTACHES EVERYWHERE

WHY ARE THERE SO MANY BIRDS IN OHIO

WHY IS THERE SO MUCH RAIN IN OHIO

WHY IS OHIO WEATHER SO WEIRD

WHY ARE THERE SQUIRRELS



WHY IS PROGRAMMING SO HARD

WHY IS THERE A 0 OHM RESISTOR

WHY DO AMERICANS HATE SOCCER

WHY DO RHYMES SOUND GOOD

WHY DO TREES DIE

WHY IS THERE NO SOUND ON CNN

WHY AREN'T POKEMON REAL

WHY AREN'T BULLETS SHARP

WHY DO DREAMS SEEM SO REAL

WHY DO IGUANAS DIE

WHY AREN'T THERE DINOSAUR GHOSTS

WHY IS THERE HELL IF GOD FORGIVES

WHY IS GPS FREE

WHY ARE THERE SWARMS OF GNATS

WHY IS THERE LAW

WHY ARE THERE FEMALE MR NIMES

WHY IS LIFE SO BORING

WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND



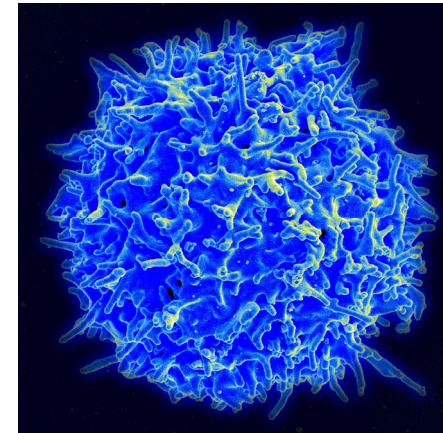
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

Human T cell expression data

- The matrix contains **47 expression samples** from Lukk et al, Nature Biotechnology 2010
- All samples are **from T cells in different individuals**
- Only the **top 3000 genes** with the largest variability **were used**
- The value is **log2 of gene's expression level** in a given sample as measured by the microarray technology

a T cell



A global map of human gene expression

Margus Lukk, Misha Kapushesky, Janne Nikkilä, Helen Parkinson, Angela Goncalves, Wolfgang Huber, Esko Ukkonen & Alvis Brazma

Affiliations | Corresponding author

Nature Biotechnology **28**, 322–324 (2010) | doi:10.1038/nbt0410-322

Although there is only one human genome sequence, different genes are expressed in many different cell types and tissues, as well as in different developmental stages or diseases. The structure of this 'expression space' is still largely unknown, as most transcriptomics experiments focus on sampling small regions. We have constructed a global gene expression map by integrating microarray data from 5,372 human samples representing 369 different cell and tissue types, disease states and cell lines. These have been compiled in an online resource (<http://www.ebi.ac.uk/gxa/array/U133A>) that allows the user to search for a gene of interest and

Matlab exercise on MLR

- Every group works with
g0=2907; g1=1527; g2=2629; g3=2881;
g4=1144; g5=1066;
- Compute **Multiple Linear Regression (MLR)**:
where
y=exp_t (g0); x1= exp_t (g1); x2= exp_t (g2);
- **How much better** the MLR did compared to the Single Linear Regression (SLR)?
- **Continue increasing** the number of genes in x until **R_adj** starts to decrease

How I did it

- `g0=2907; g1=1527; g2=2629; g3=2881;g4=1144; g5=1066;`
- `y=exp_t(g0,:)' ;`
- `%% first use one x to predict y`
- `x=exp_t(g1,:)' ;`
- `figure; plot(x,y,'ko')`
- `lm=fitlm(x,y)`
- `y_fit=lm.Fitted;`
- `hold on;`
- `plot(x,lm.Fitted,'r-');`
- `%% now use 2 x's to predict y`
- `x=[exp_t(g1,:)', exp_t(g2,:)]';`
- `lm2=fitlm(x,y)`
- `y_fit=lm2.Fitted;`
- `hold on; plot(x(:,1),y_fit,'gd');`
- `%% now use m x's to predict y`
- `corr_matrix=corr(exp_t');`
- `g0=2907;`
- `[u v]=sort(corr_matrix(g0,:), 'descend');`
- `x=[exp_t(v(2:m+1),:)]';`
- `lm3=fitlm(x,y)`
- `y_fit=lm3.Fitted;`
- `plot(x(:,1),y_fit,'s');`

Credit: XKCD
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

QUESTIONS FOUND IN GOOGLE AUTOCOMPLETE



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD

WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS

WHY ARE THERE SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE

WHY ARE THERE ANTS IN MY LAPTOP

WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KUNGONS DIFFERENT



WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY



WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS

WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE



WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL

WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD



WHY IS GPS FREE