EE16B
Designing Information Devices and Systems II

Lecture 9A
Geometry of SVD, PCA

Intro

- Last time:
  - Described the SVD in
    - Compact matrix form: $U_1 S V_1^T$
    - Full form: $U \Sigma V^T$
  - Showed a procedure to SVD via $A^T A$
- Today:
  - Show procedure via AA^T
  - Continue proofs (symmetric matrices)
  - PCA

Computing the SVD with AᵀA

$$\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$$

Proof concept: let $A^T A\vec{v}_i = \lambda_i \vec{v}_i \Rightarrow A^T A V_1 = \Lambda V_1$

$$\sigma_i^2 = \lambda_i \qquad S^2 = \Lambda$$

Show that $A\vec{v}_i = \sigma_i \vec{u}_i$, where

$$\vec{u}_i^T \vec{u}_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad \Longrightarrow \quad U_1^T U_1 = I_{r\times r}$$

Show that $A = U_1 S V_1^T$

Alternate Procedure using AA^T

Step 1: Find eigenvalues of AA^T and order s.t.

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 \cdots 0$$

Step 2: Find orthonormal eigenvectors of AA^T:

$$AA^T \vec{u}_i = \lambda_i \vec{u}_i \qquad i = 1, \cdots, r$$

Step 3: Set,

$$\sigma_i = \sqrt{\lambda_i} \qquad \vec{v}_i = \frac{1}{\sigma_i} A^T \vec{u}_i$$

Example

$$A = \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \qquad r = 2$$

$$A^T A = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix} \qquad AA^T = \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}$$

$$\lambda_1 = 32 \qquad \lambda_2 = 18$$

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\vec{v}_i = \frac{1}{\sigma_i} A^T \vec{u}_i \qquad \vec{v}_1 = \frac{1}{4\sqrt{2}} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Signs of u_1,v_1 (u_2,v_2) can be flipped!

Uniqueness of the SVD

Find SVD of A

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

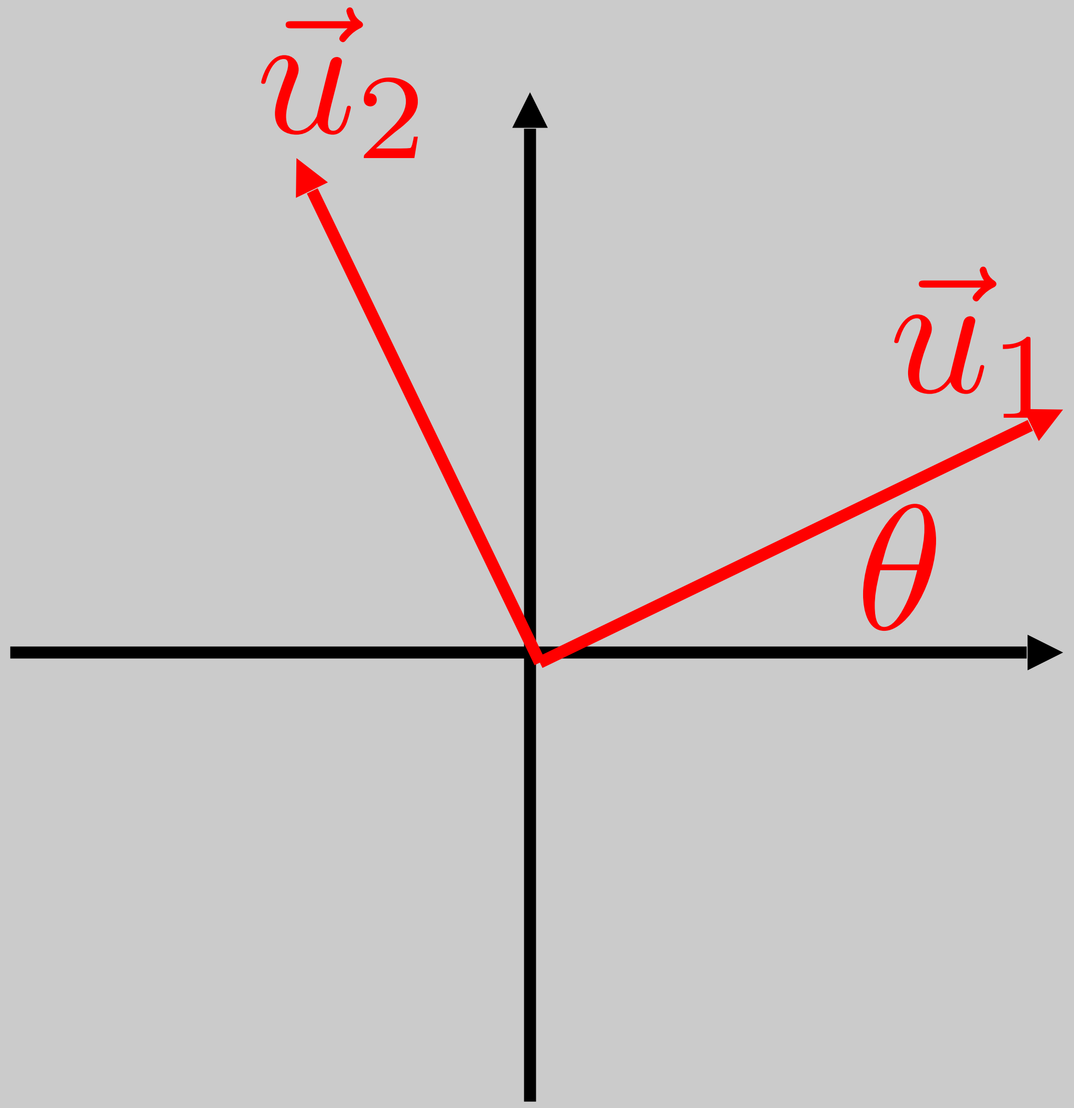
Uniqueness of the SVD

Find SVD of A

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

$$AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \lambda_1 = \lambda_2 = 1 \quad \Rightarrow \sigma_1 = \sigma_2 = 1$$

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



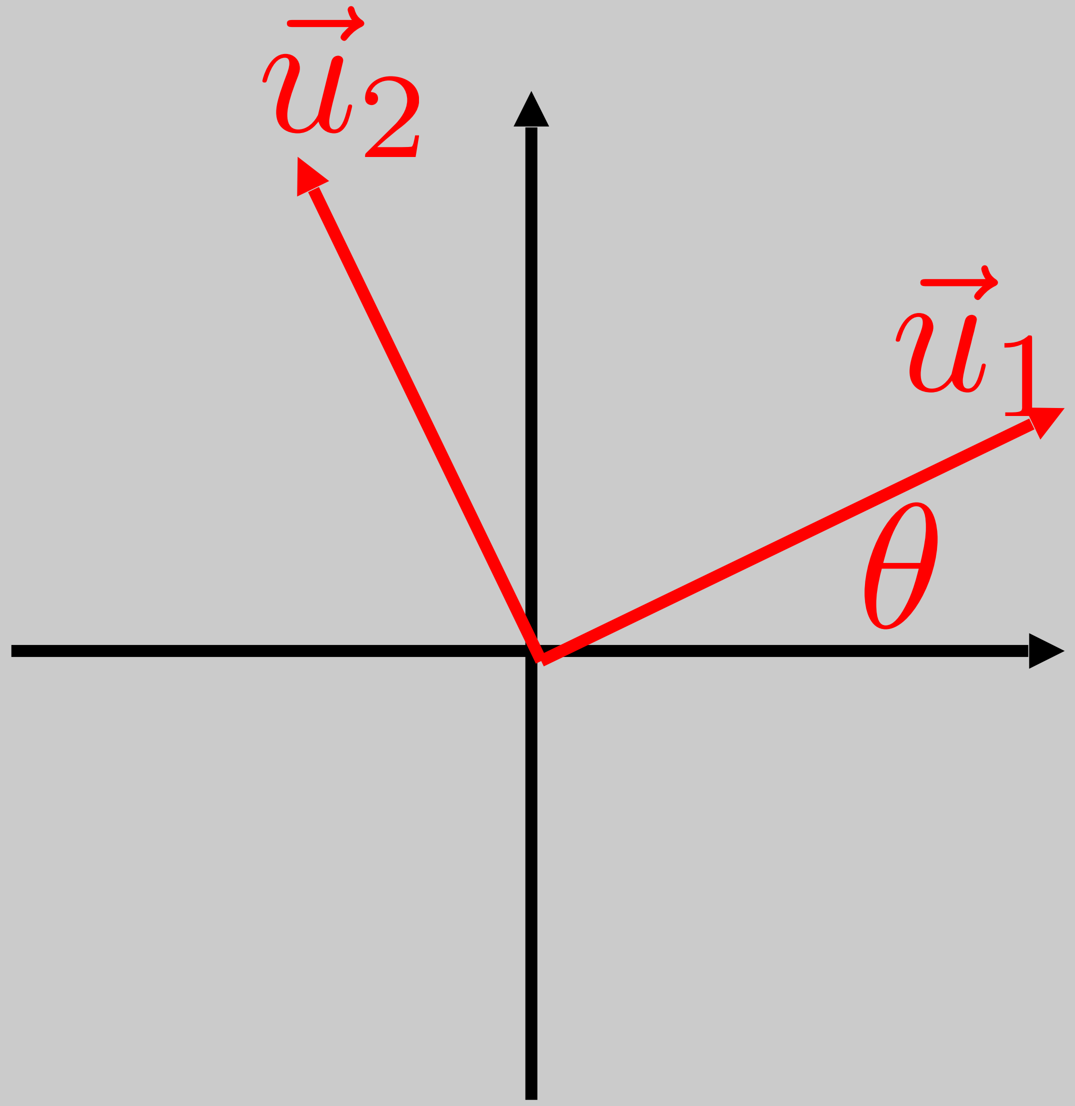
$$\vec{u}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \quad \vec{u}_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

Uniqueness of the SVD

Find SVD of A

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

$$AA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \lambda_1 = \lambda_2 = 1 \qquad \Rightarrow \sigma_1 = \sigma_2 = 1$$



$$\vec{u}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

$$\vec{v}_1 = \frac{1}{\sigma_1} A^T \vec{u}_1 = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} -\sin\theta \\ -\cos\theta \end{bmatrix}$$

Uniqueness of the SVD

Find SVD of A

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \Rightarrow \sigma_1 = \sigma_2 = 1$$

$$\vec{u}_1 = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \; \vec{u}_2 = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad \vec{v}_1 = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix} \vec{v}_2 = \begin{bmatrix} -\sin\theta \\ -\cos\theta \end{bmatrix}$$

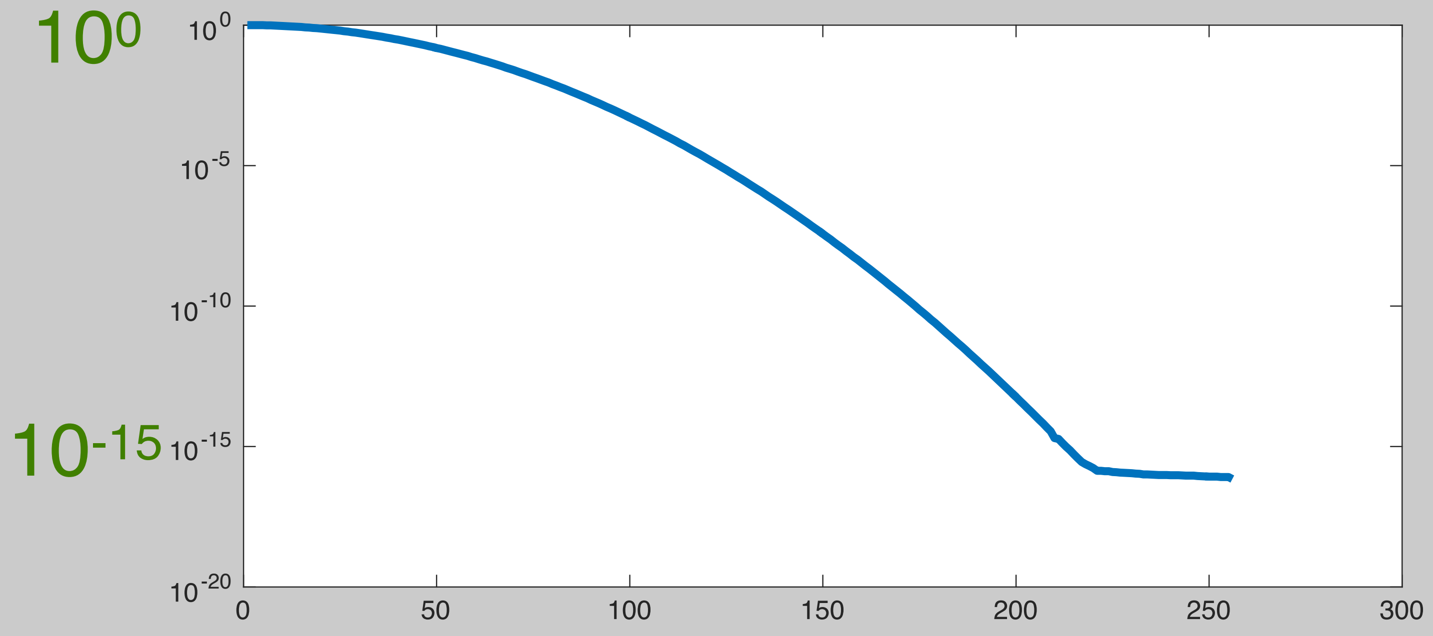
$$A = \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T$$

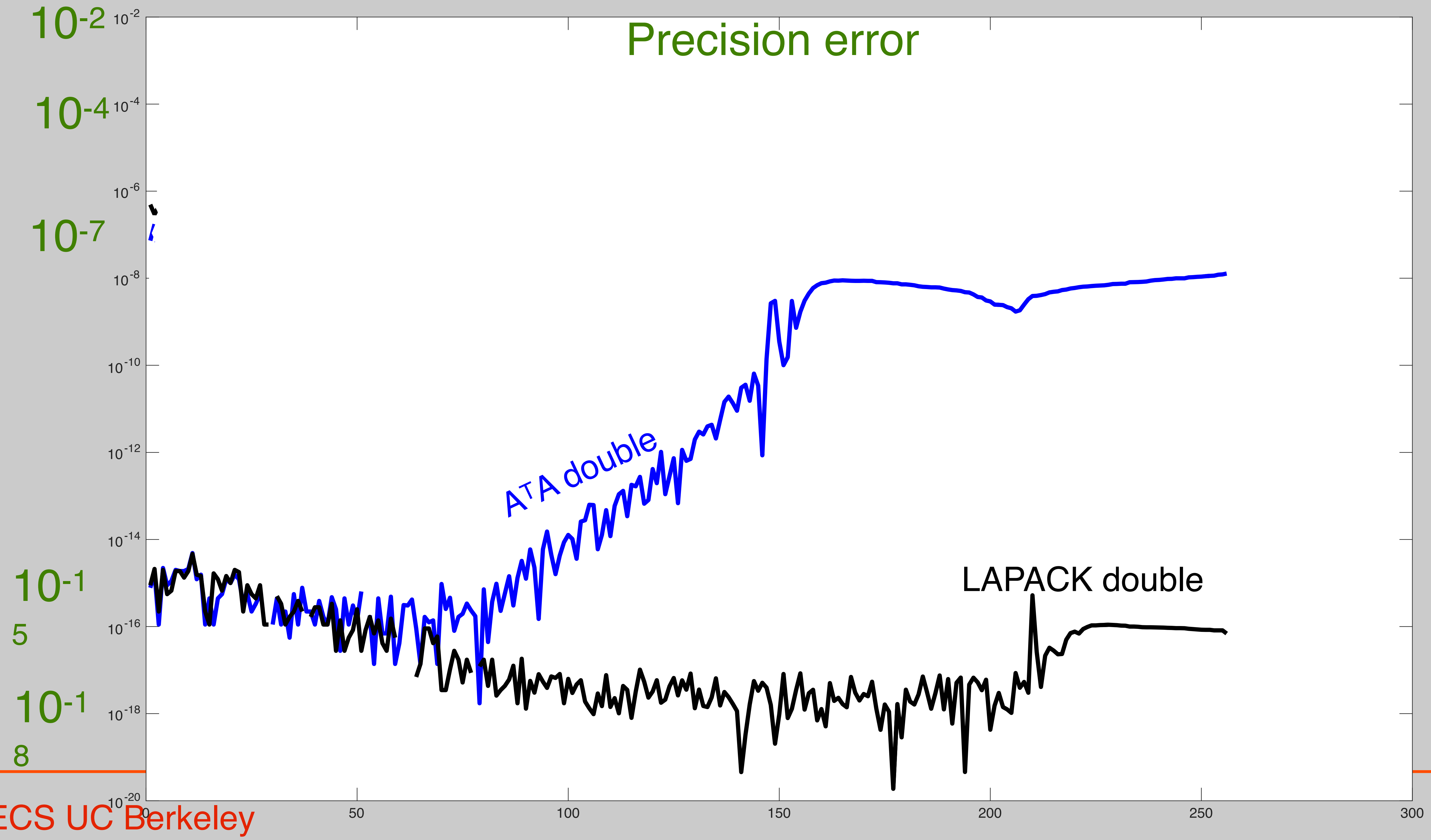
$$= \begin{bmatrix} \cos^2\theta & -\sin\theta\cos\theta \\ \sin\theta\cos\theta & -\sin^2\theta \end{bmatrix} + \begin{bmatrix} \sin^2\theta & \sin\theta\cos\theta \\ -\sin\theta\cos\theta & -\cos^2\theta \end{bmatrix}$$

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

Accuracy with Finite Precision

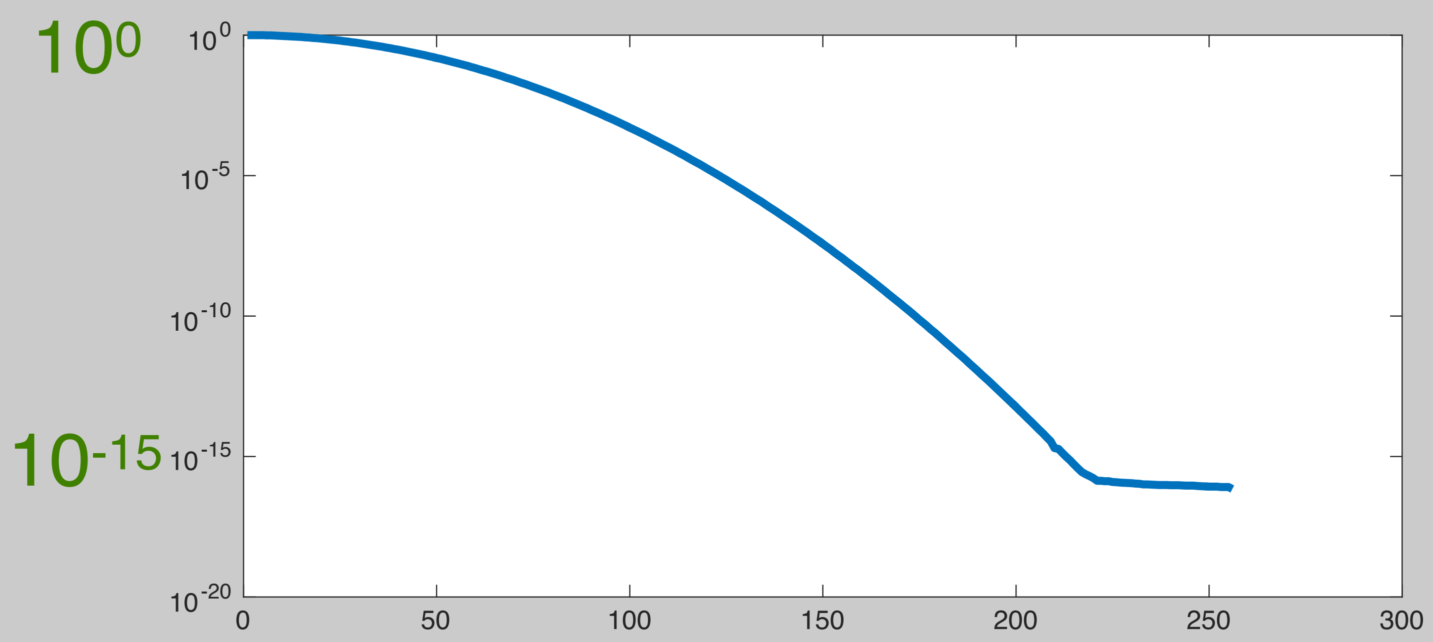
Consider matrix $A \in R^{512 \times 256}$ with the following singular values:

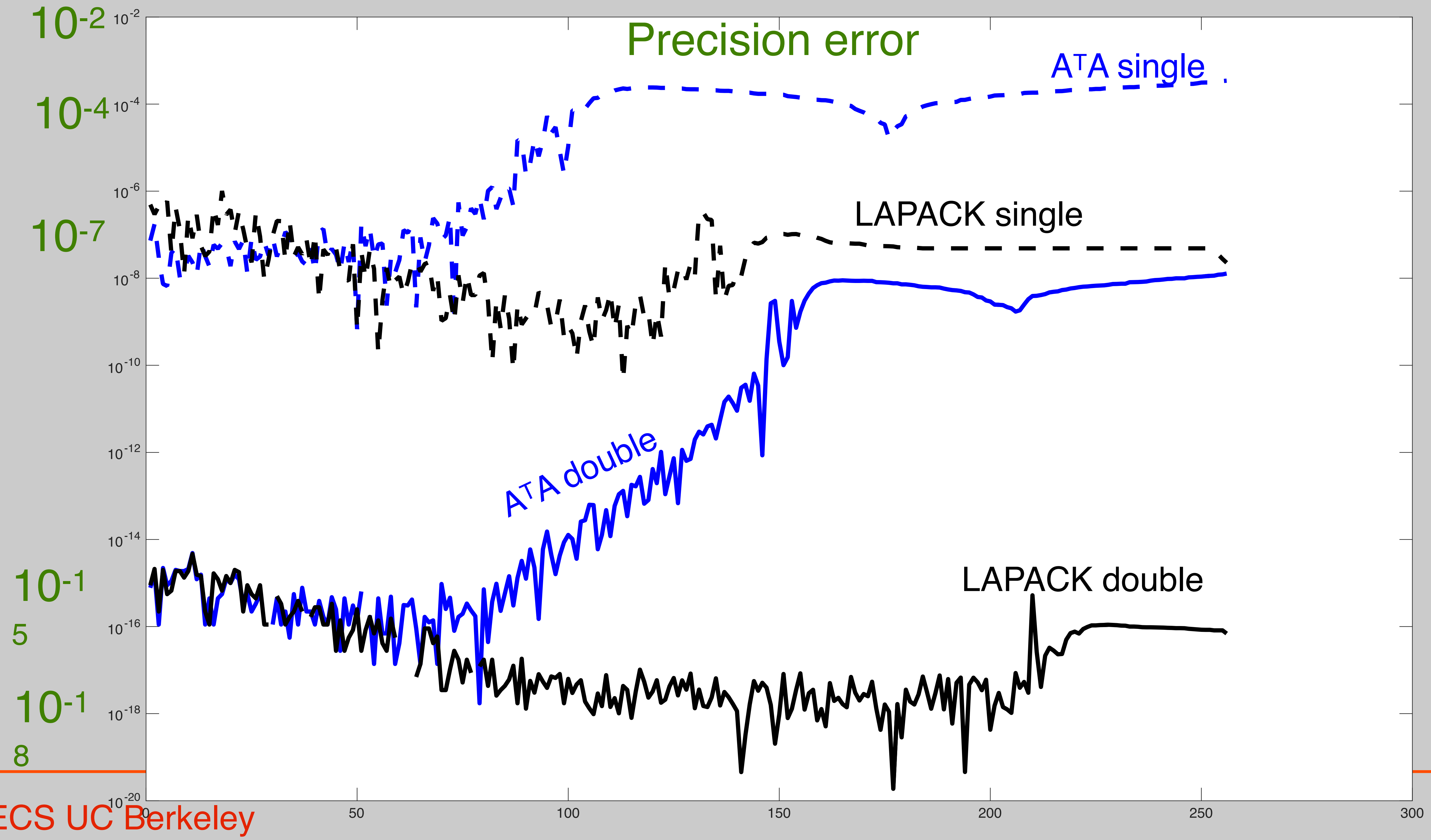




Accuracy with Finite Precision

Consider matrix $A \in R^{512x256}$ with the following singular values:





Full Matrix Form of SVD

$$U_1 = \underbrace{\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_r \end{bmatrix}}_{m \times r} \qquad S = \underbrace{\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}}_{r \times r} \qquad V_1 = \underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_r \end{bmatrix}}_{n \times r}$$

$$U = \underbrace{\begin{bmatrix} U_1 & U_2 \end{bmatrix}}_{m \times m} \qquad \Sigma = \underbrace{\left[\begin{array}{c|c} S & 0 \\ \hline 0 & 0 \end{array}\right]}_{m \times n} \qquad V = \underbrace{\begin{bmatrix} V_1 & V_2 \end{bmatrix}}_{n \times n}$$

$$A = U\Sigma V^T \qquad \begin{array}{l} U^T U = I_{m\times m} \\ V^T V = I_{n\times n} \end{array}$$

Unitary Matrices

Multiplying with unitary matrices does not change the length

$$||U\vec{x}|| = \sqrt{(U\vec{x})^T(U\vec{x})} = \sqrt{\vec{x}^T U^T U \vec{x}} = \sqrt{\vec{x}^T \vec{x}} = ||\vec{x}||$$

Example: Rotation, or reflection matrices

$$U = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Geometric Interpretation

$$A = U\Sigma V^T$$

$$A\vec{x} = U\Sigma V^T\vec{x}$$

1) $V^T\vec{x}$ re-orients $\vec{x}$ without changing length.

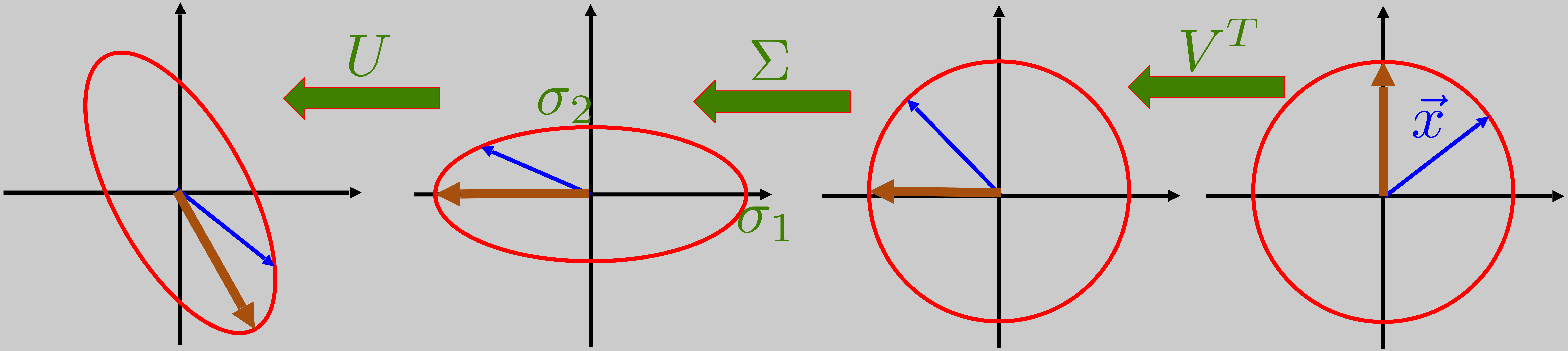
2) $\Sigma(V^T\vec{x})$ Stretches along the axis with singlular values

$$\begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 x_1 \\ \sigma_2 x_2 \end{bmatrix}$$

3) $U(\Sigma V^T\vec{x})$ re-orients again without changing length

Geometric Interpretation

$$A = U\Sigma V^T \qquad\qquad A\vec{x}$$



$$||A\vec{x}|| \leq \sigma_1 ||\vec{x}||$$

Q: What vector would amplify the most?

Symmetric Matrices

We assumed before that,

A^TA has only real eigenvalues, r of them are positive and the rest are zero
A^TA has orthonormal eigenvectors (to be proven next time)

For symmetric matrices: $$Q^T = Q$$

$$(AB)^T = B^TA^T$$

$$(A^TA)^T = A^TA$$

$$(AA^T)^T = AA^T$$

Properties of Symmetric Matrices

1) A real-valued symmetric matrix has real eigenvalues and eigenvectors

$$Qx = \lambda x \qquad \lambda = a + ib \qquad \overline{\lambda} = a - ib$$

Somehow we need to use the symmetric and real-ness property of Q to show that b==0

$$Q\overline{x} = \overline{\lambda}\overline{x}$$

$$\overline{x}^T Q = \overline{\lambda}\overline{x}^T$$

$$\overline{x}^T Q x = \overline{\lambda}\overline{x}^T x \qquad \overline{x}^T Q x = \lambda \overline{x}^T x$$

$$\overline{\lambda}\overline{x}^T x = \lambda \overline{x}^T x \quad \Rightarrow \lambda = \overline{\lambda} \Rightarrow \lambda \in \mathrm{R}$$

Properties of Symmetric Matrices

$$Qx = \lambda x$$

$$\underbrace{(Q - \lambda I)}_{\text{real}}x = 0$$

So x is real as well

✅A real-valued symmetric matrix has real eigenvalues and eigenvectors

Properties of Symmetric Matrices

2) Eigenvectors of a symmetrix matrix can be chosen to be orthonormal

Choose two distinct eigenvalues and vectors $\lambda_1 \neq \lambda_2$

$$Qx_1 = \lambda_1 x_1 \qquad Qx_2 = \lambda_2 x_2$$

$$x_2^T Q x_1 = \lambda_1 x_2^T x_1 \qquad x_1^T Q x_2 = \lambda_2 x_1^T x_2$$

$$(\lambda_1 - \lambda_2) x_2^T x_1 = 0$$

$$\lambda_1 \neq \lambda 2 \Rightarrow x_2^T x_1 = 0$$

✅ Eigenvectors of a symmetrix matrix can be chosen to be orthonormal

Positiveness of Eigenvalues

3) If Q can be written as Q = R^TR for real R, then Q is positive semidefinite – eigenvalues greater of equal to zero

$$Qx = \lambda x$$

$$R^T R x = \lambda x$$

$$x^T R^T R x = \lambda x^T x$$

$$(Rx)^T (Rx) = \lambda x^T x$$

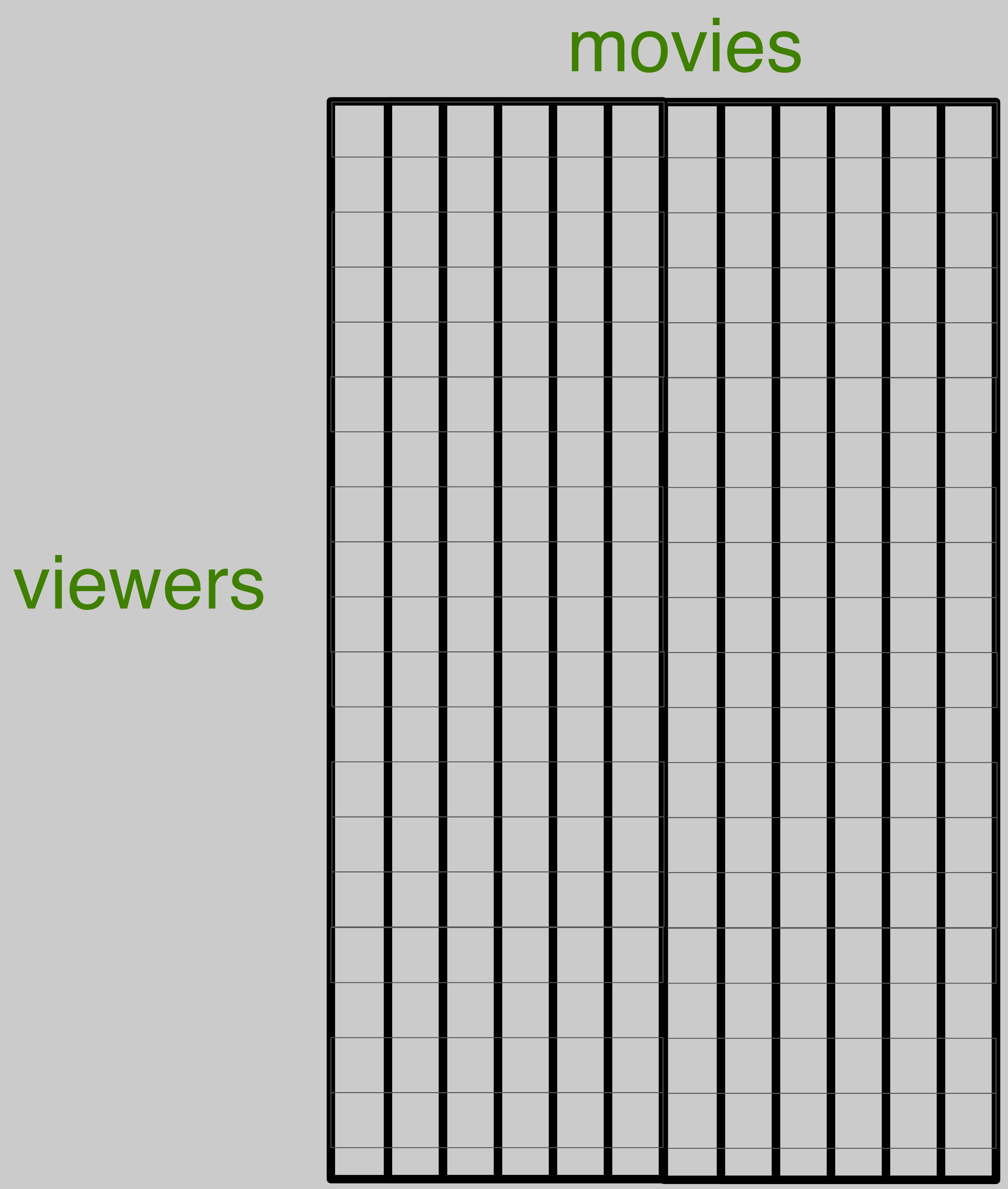
$$||Rx||^2 = \lambda ||x||^2 \quad \Rightarrow \lambda \geq 0$$

✅ If Q can be written as Q = R^TR for real R, then Q is positive semidefinite – eigenvalues greater of equal to zero

Principal Component Analysis

Application of the SVD to datasets to learn features

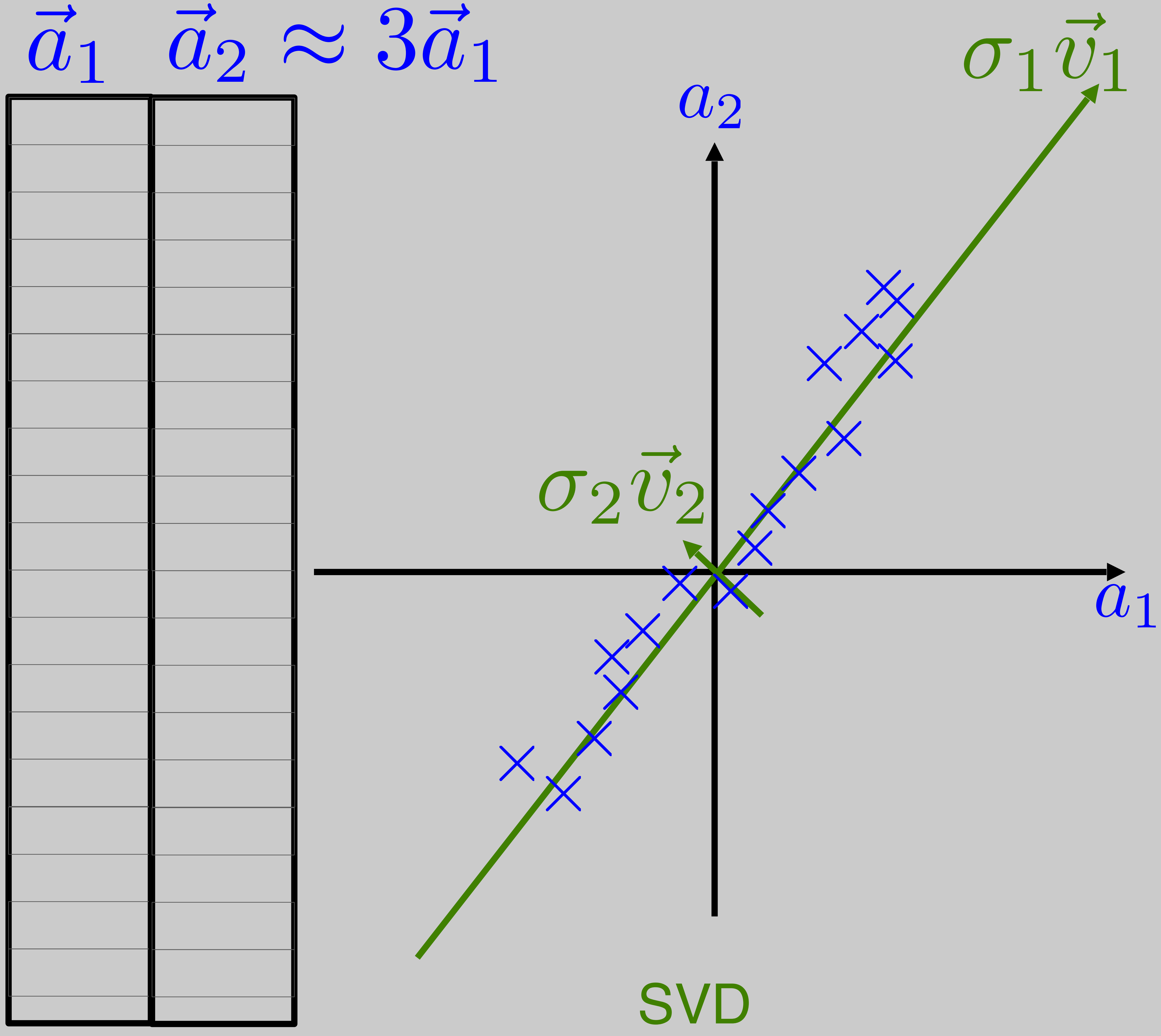
PCA is a tool in statistics and machine learning, which can be computed using SVD



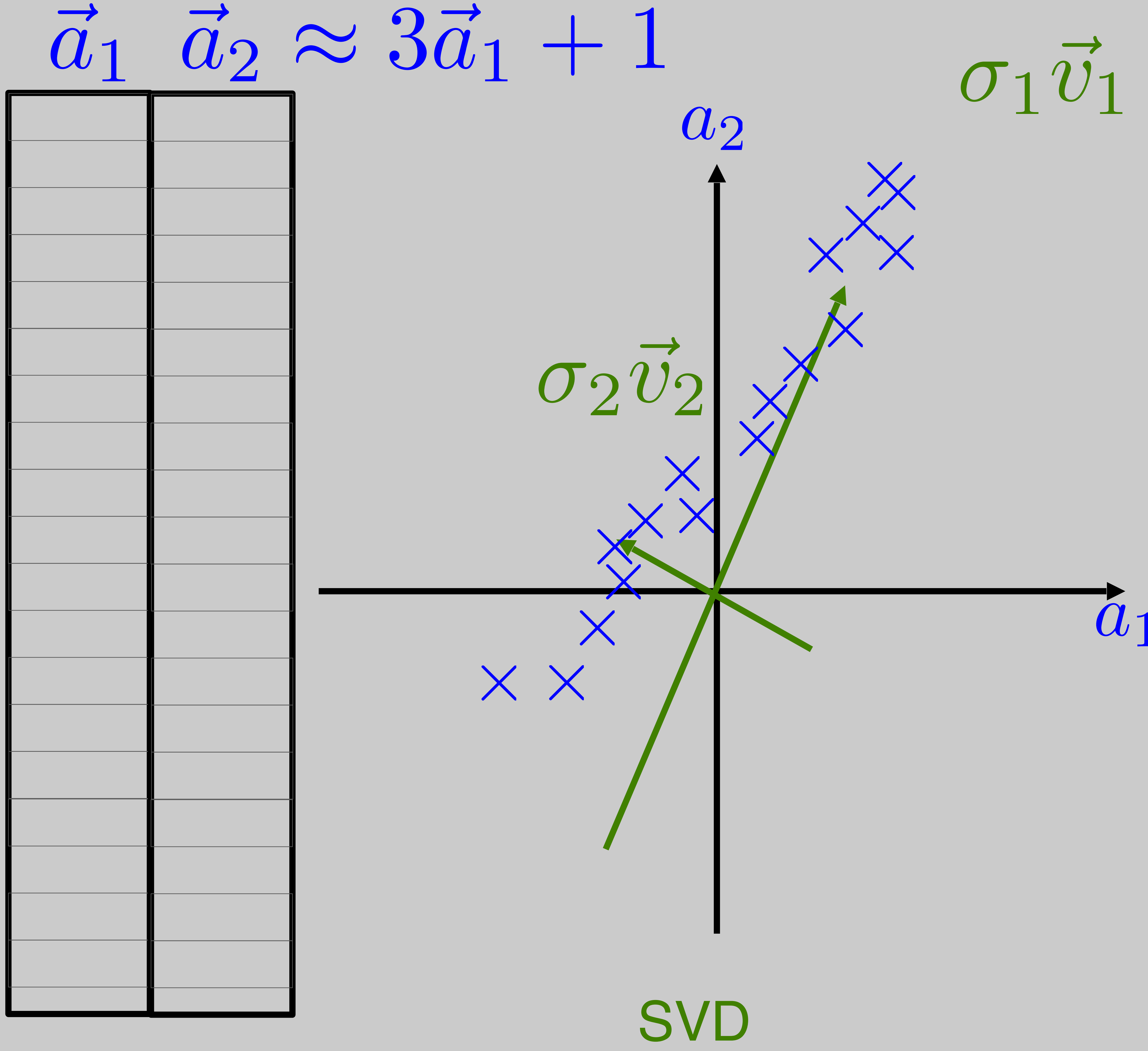
Example -- PCA

Consider data s.t.

$\vec{a}_1 \quad \vec{a}_2 \approx 3\vec{a}_1$

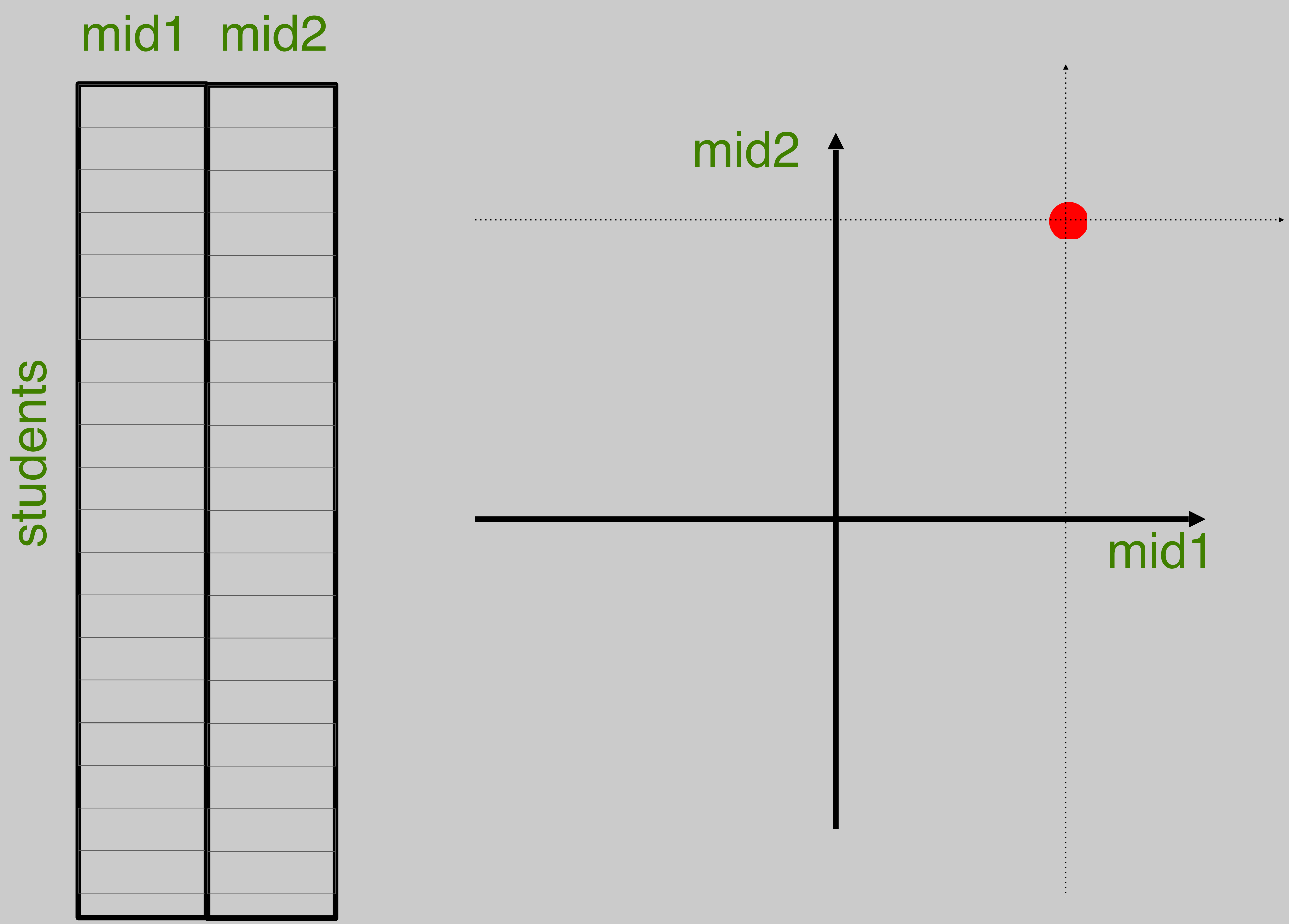


$\vec{a}_1 \quad \vec{a}_2 \approx 3\vec{a}_1 + 1$



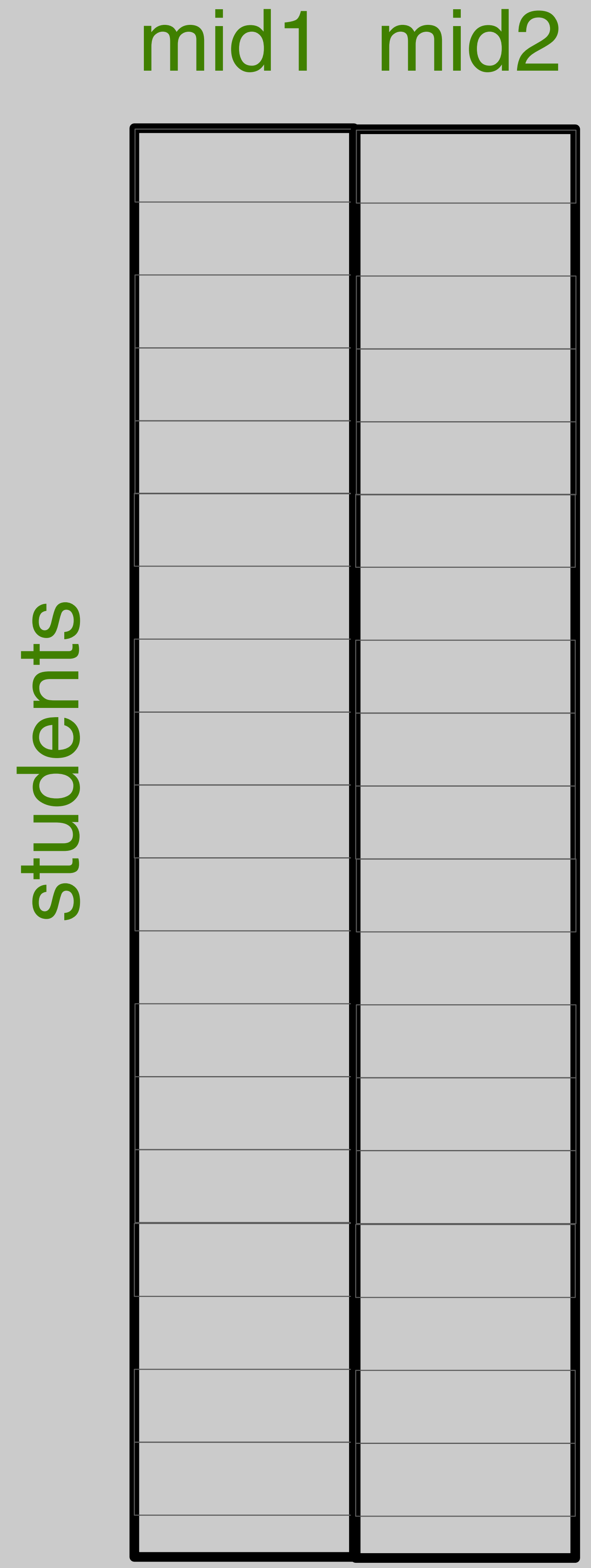
Example -- PCA

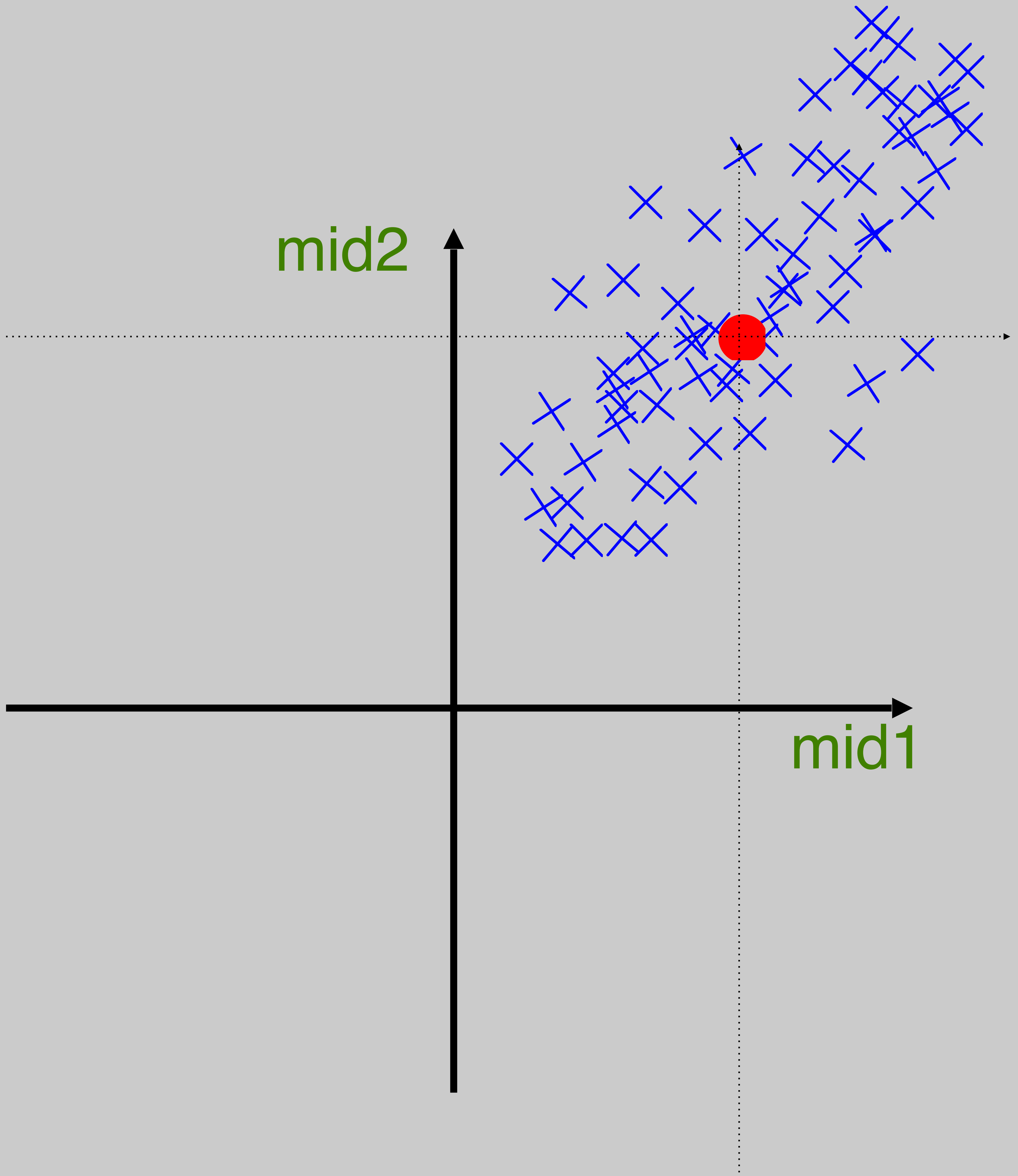
Consider miterm data



Example -- PCA

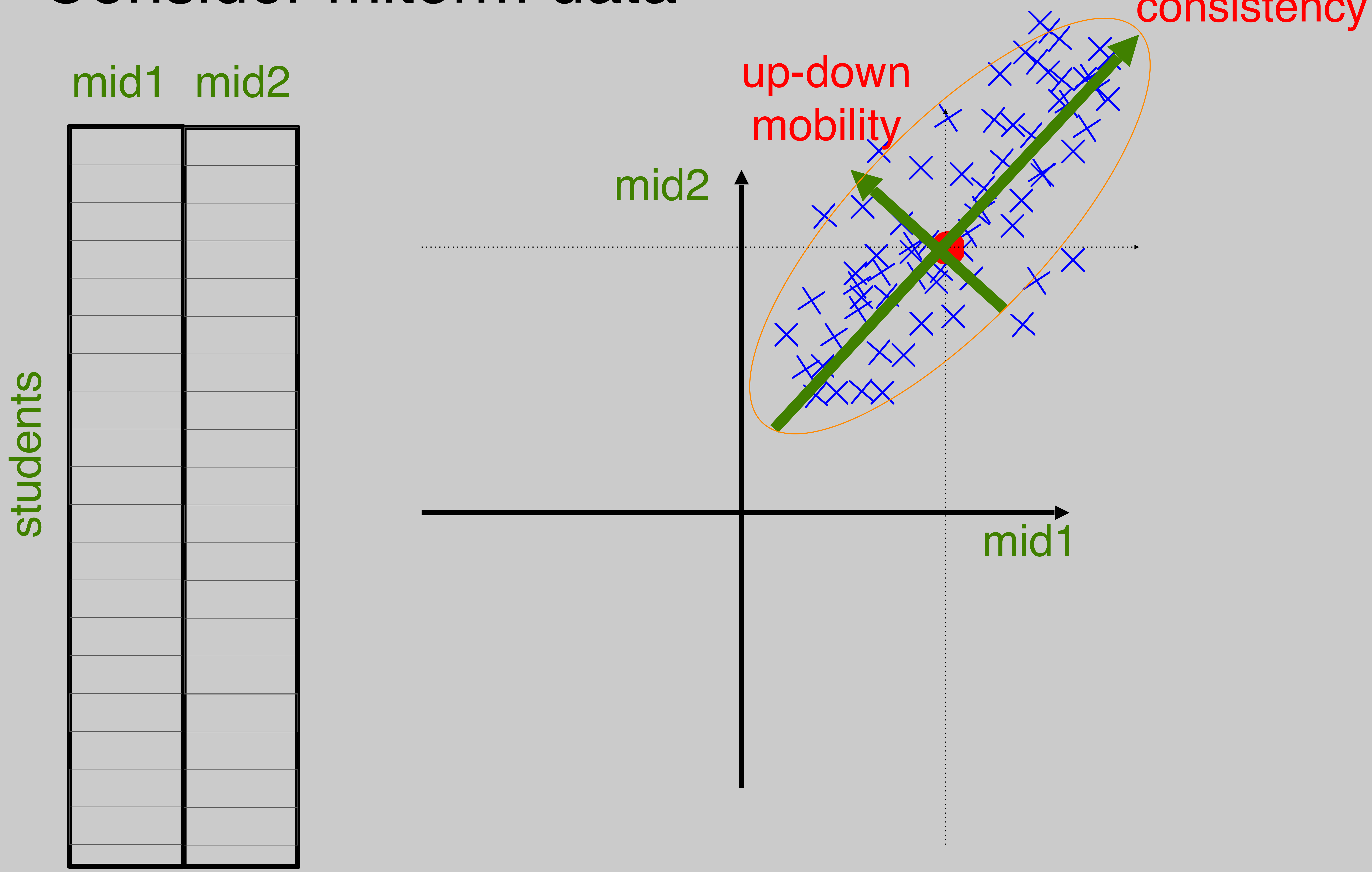
Consider miterm data





Example -- PCA

Consider miterm data

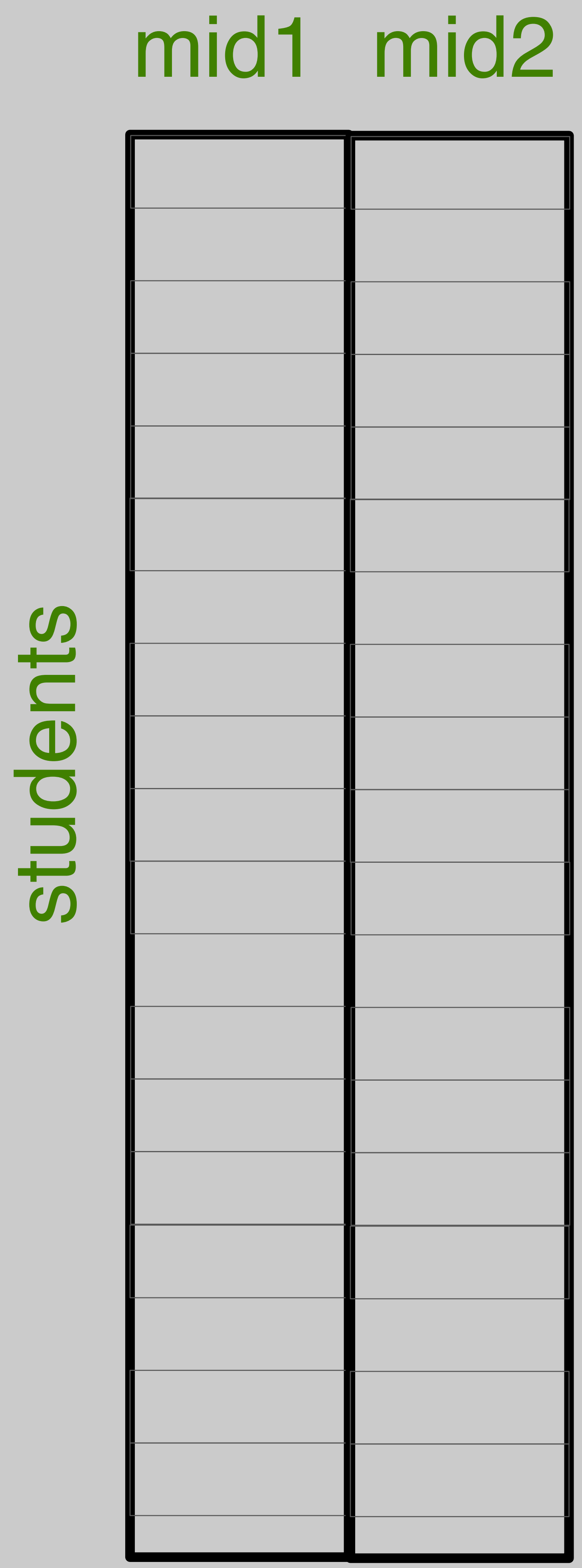


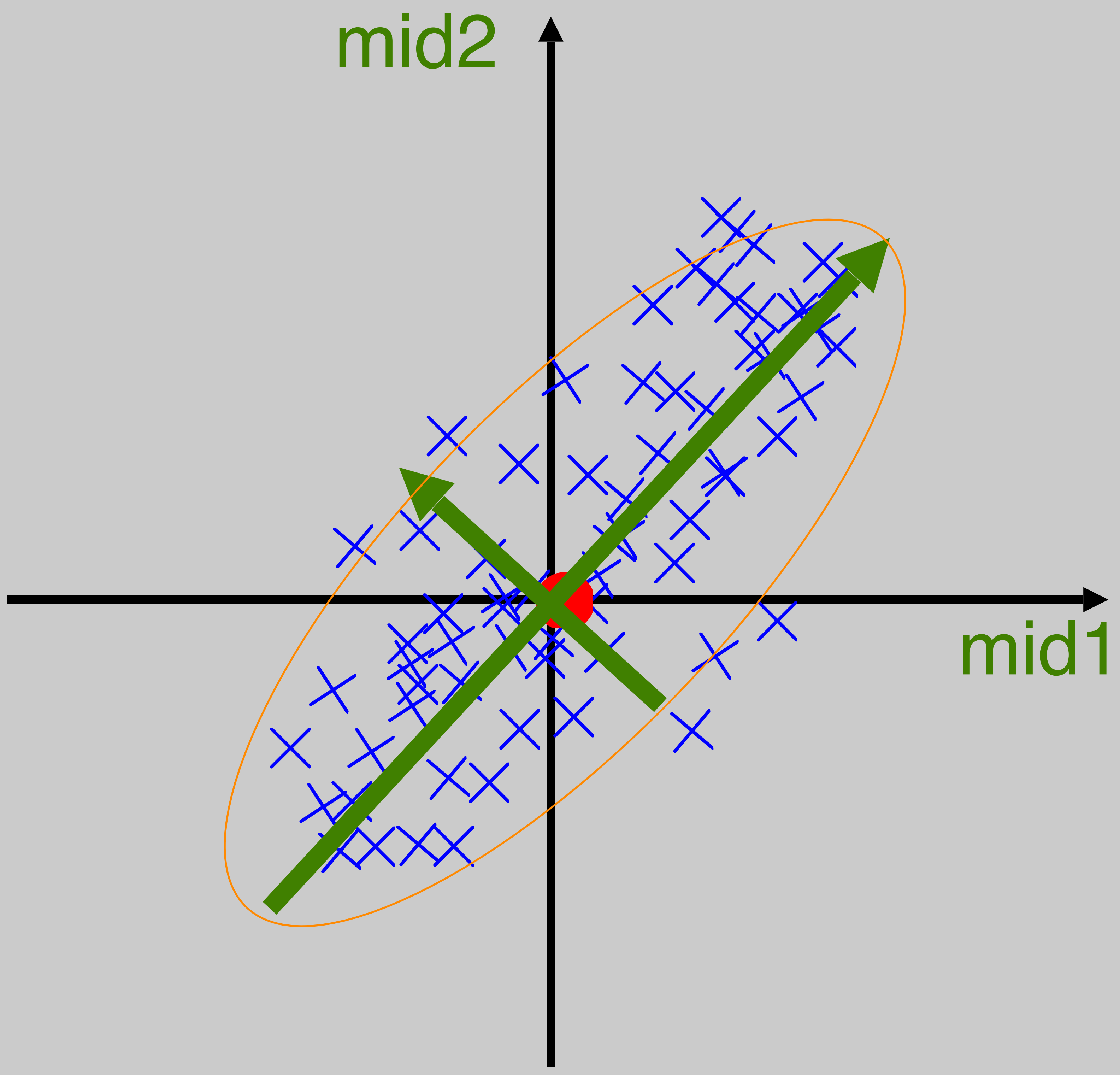
PCA Procedure

Remove averages from column of A

From A^TA, find $\sigma_i, \quad \vec{v_i}$

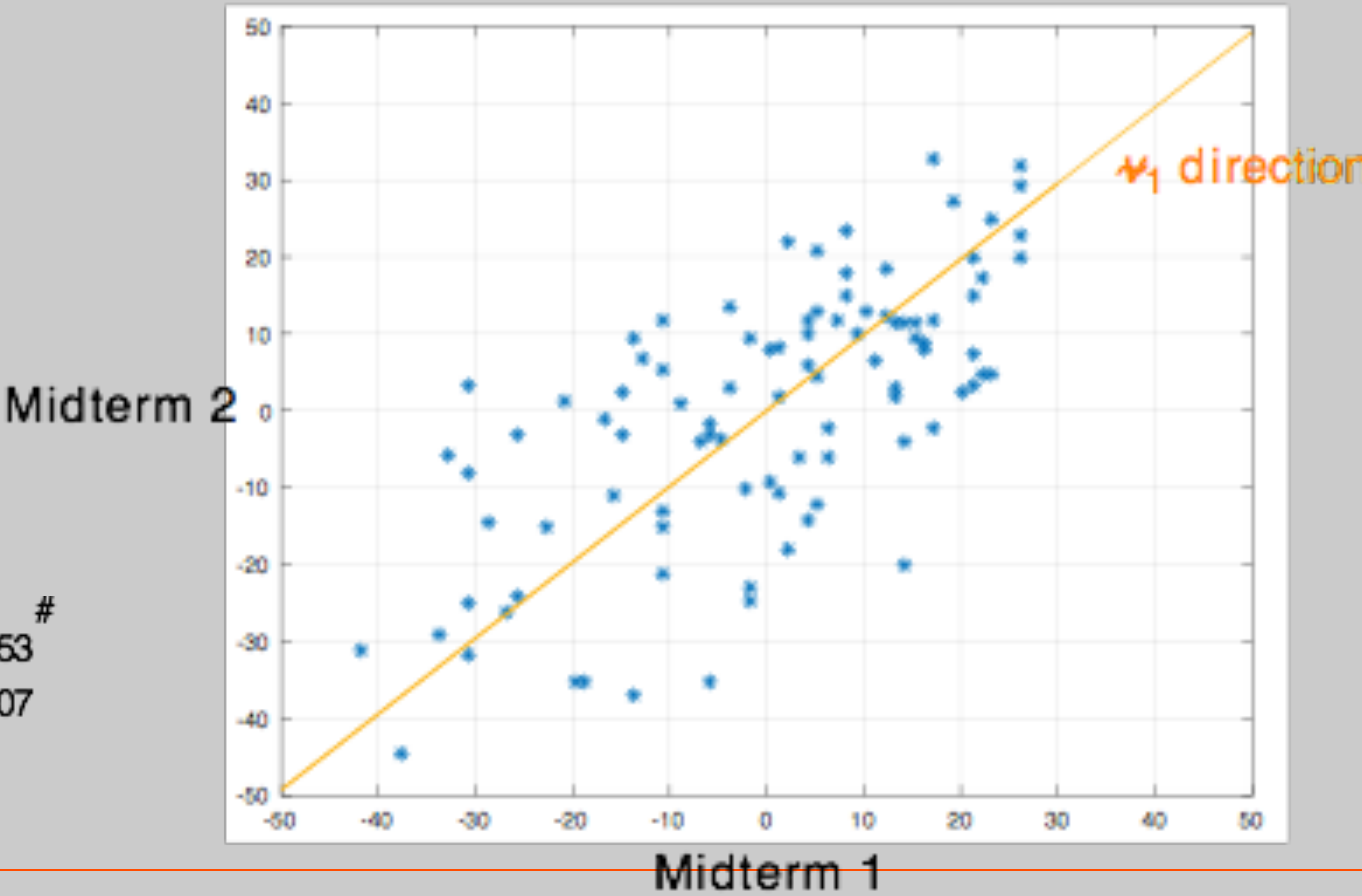
$\vec{v_i}$ are principal components!





A^TA as sample covariance matrix

Example midterm



$$\frac{1}{93}A^TA = \begin{bmatrix} 297.69 & 202.53 \\ 202.53 & 292.07 \end{bmatrix}$$

PCA in Genetics Reveals Geography

Genes mirror geography within Europe
Nature **456**, 98-101 (6 November 2008)

Study:

Characterized genetic variatios in 3,000 Europeans from 36 Countries

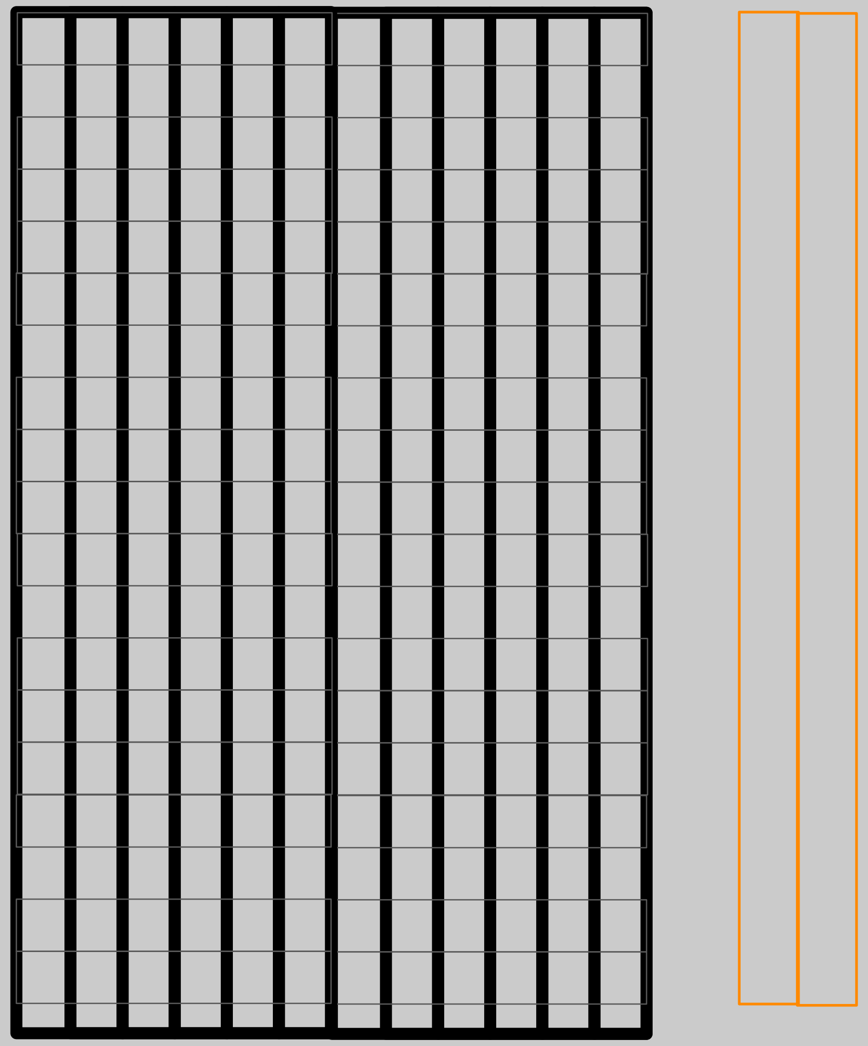
Built a matrix of 200K SNPs (single nucleotide polymorphisms)

Computed largest 2 principle components

Projected subjects on 2 dimentional data

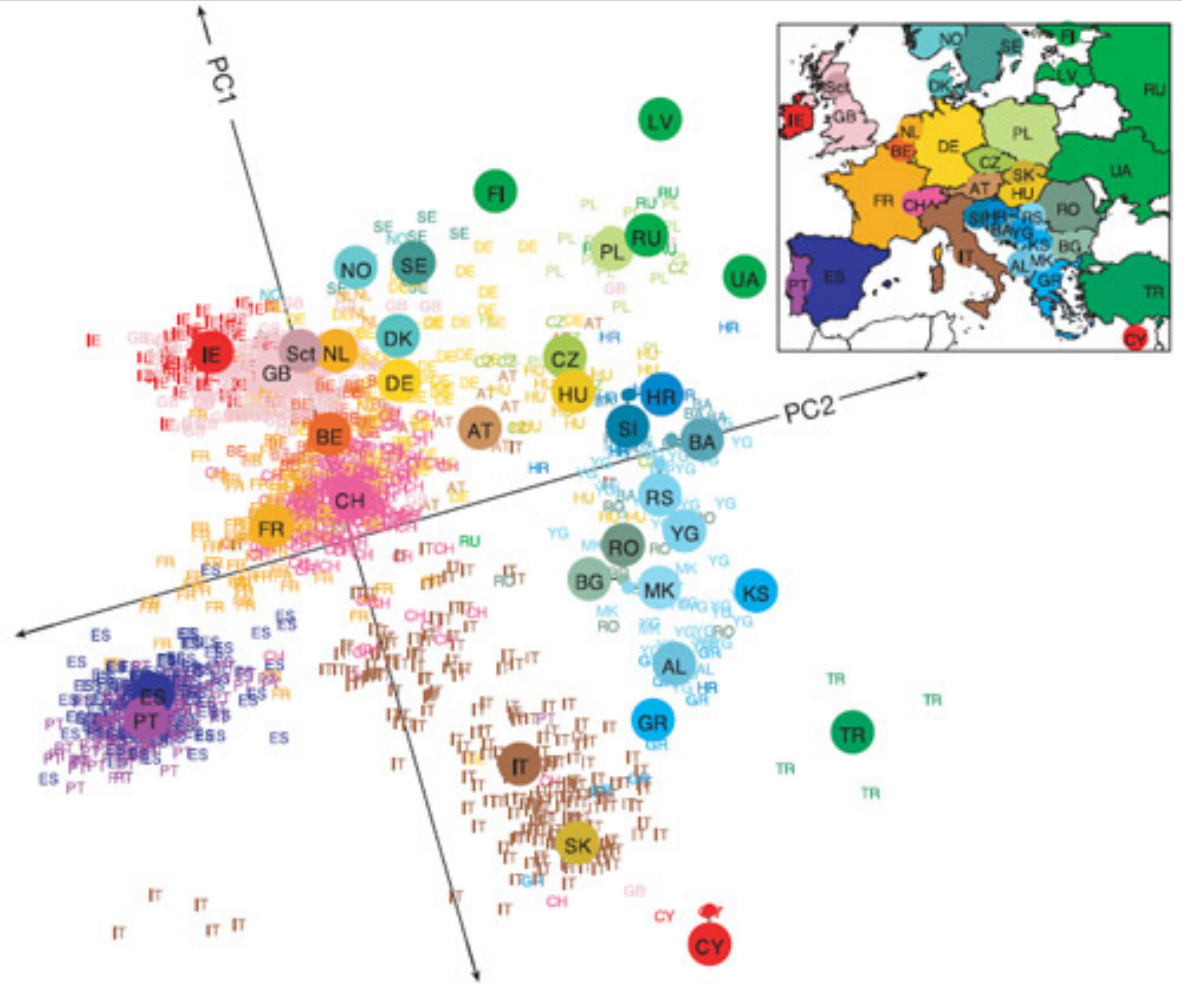
Overlayed the result on the map of Europe

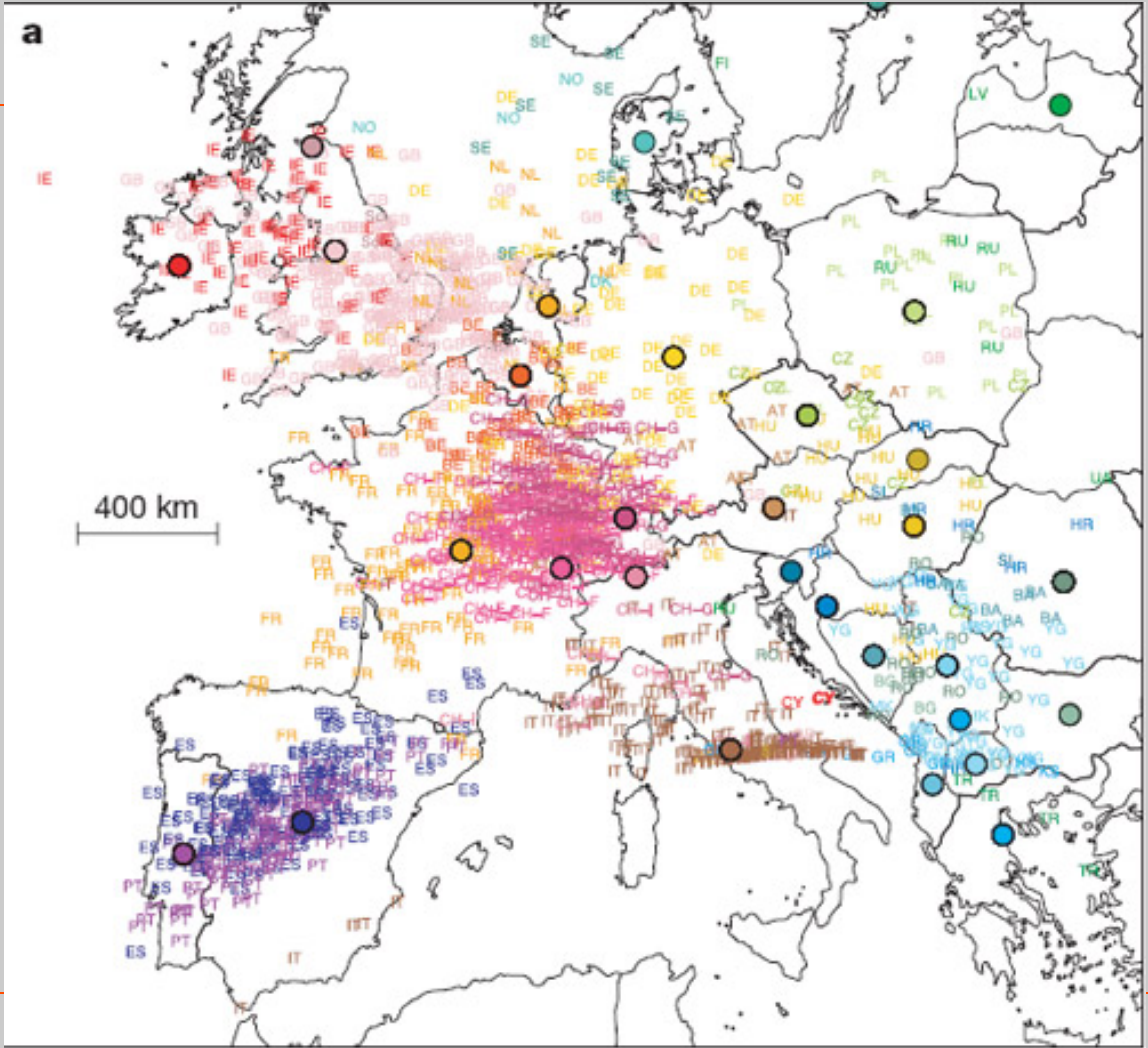
$A\vec{v}_1 \quad A\vec{v}_2$

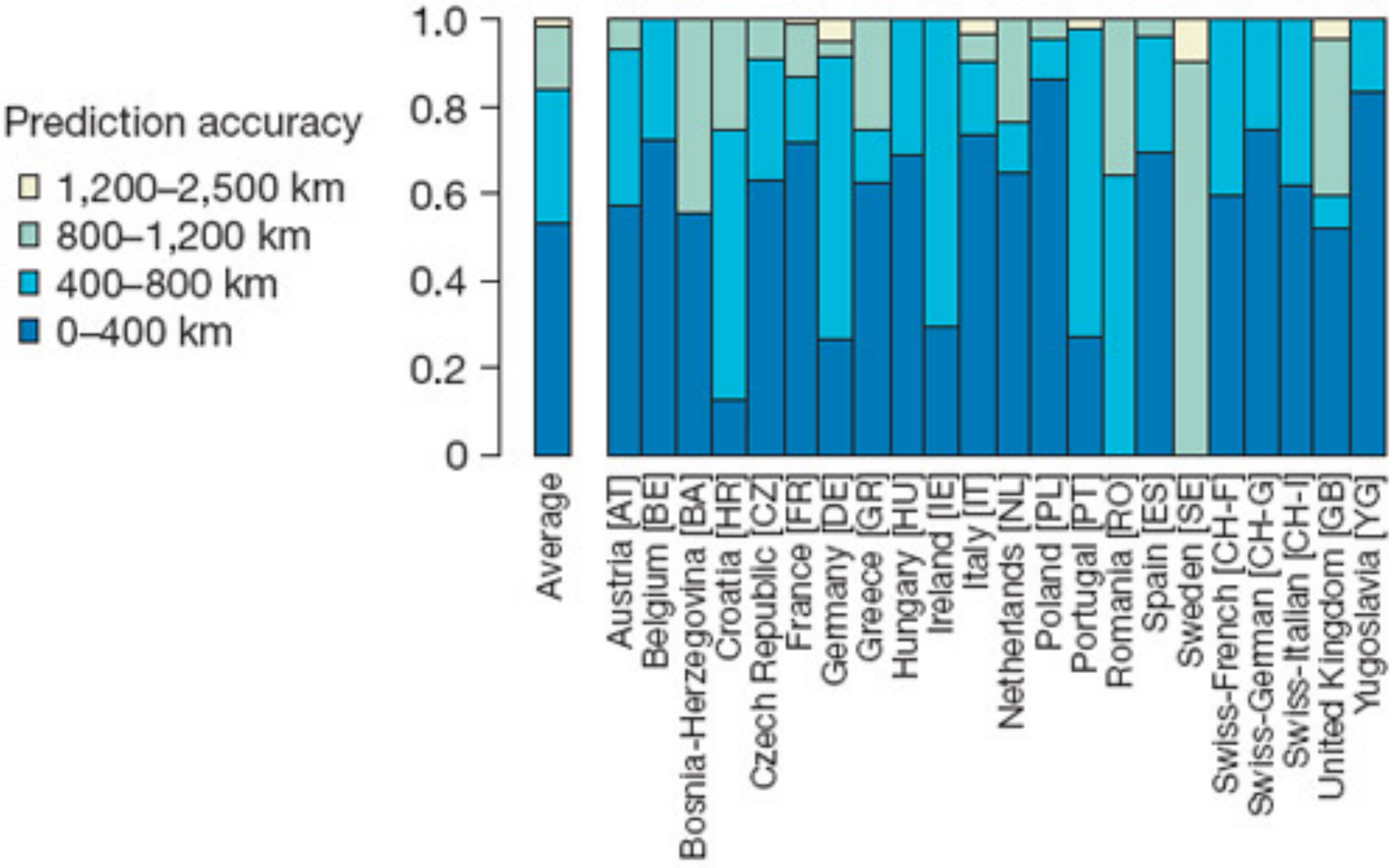


PC1 could be associated with food

PC2 associated with west migration




a
400 km

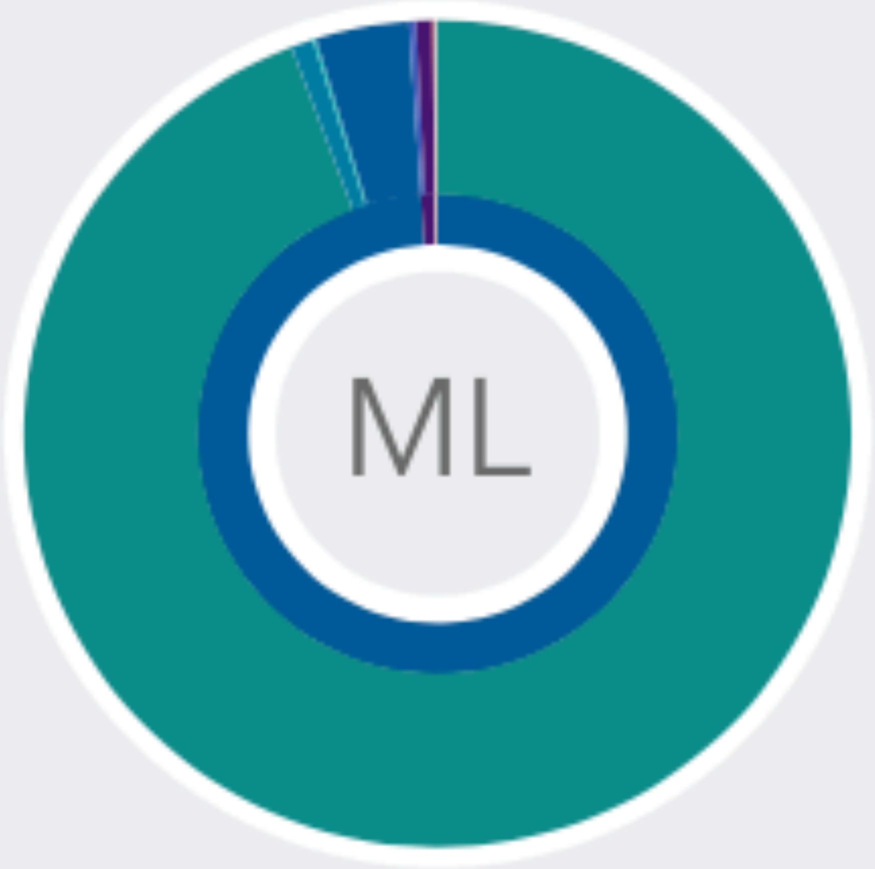

Prediction accuracy
1,200–2,500 km
800–1,200 km
400–800 km
0–400 km
1.0
0.8
0.6
0.4
0.2
0
Average
Austria [AT]
Belgium [BE]
Bosnia-Herzegovina [BA]
Croatia [HR]
Czech Republic [CZ]
France [FR]
Germany [DE]
Greece [GR]
Hungary [HU]
Ireland [IE]
Italy [IT]
Netherlands [NL]
Poland [PL]
Portugal [PT]
Romania [RO]
Spain [ES]
Sweden [SE]
Swiss-French [CH-F]
Swiss-German [CH-G]
Swiss-Italian [CH-I]
United Kingdom [GB]
Yugoslavia [YG]

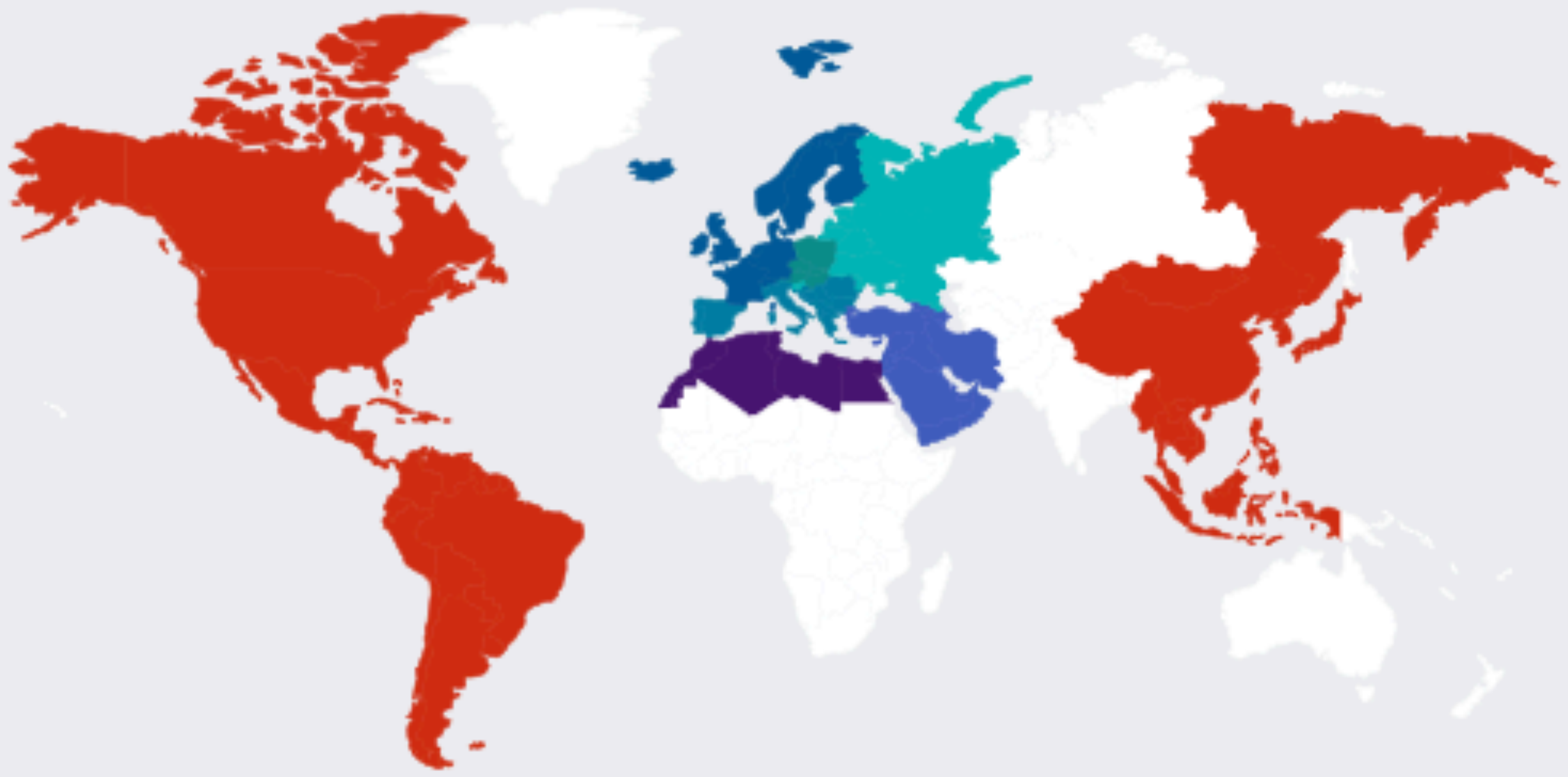
Interesting conclusions

"The results have implications for a lot of biomedical research. Many scientists are scanning entire genomes on a hunt for SNPs that affect a person's risk of diseases like cancer or their reaction to drugs. Novembre says that researchers who are running these "whole-genome studies" need to bear in mind where their sample has come from. Even if a study looks at a small and seemingly related parts of Europe, it would have to adjust for any geographical influences in the genetic variations it uncovers."

http://phenomena.nationalgeographic.com/2008/09/01/european-genes-mirror-european-geography/

23 and me





Michael Lustig	100%
European	98.9%
Middle Eastern & North African	0.9%
East Asian & Native American	< 0.1%
Unassigned	0.2%