# Recitation #7

Irina Espejo (iem244@nyu.edu)

Center for Data Science

DS-GA 1014: Optimization and Computational Linear Algebra for Data Science

# Announcement

**Next recitation session on Wed 28th October will be a midterm review.**

**The office hours on Thr 29th October is moved to Wed 28th 9 AM EST after the recitation**

# recall: how to find an eigenvalue

**Exercise 0**

Let $A \in \mathbb{R}^{n \times n}$ have an eigenvalues . How would you find its eigenvector $v \in \mathbb{R}^n$ that is $Av = \lambda v$

# recall: spectral theorem

## Spectral theorem

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, Then there is a orthonormal basis of $\mathbb{R}^n$ composed of eigenvectors of $A$.

## Let's get things clear

- If $A$ is symmetric then $A^T = A$ and it diagonalizes $A = PDP^{-1}$
- $P$ is an orthonormal matrix therefore $P^T = P^{-1}$

# practice: spectral theorem

**Exercise 1**

Prove that if $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix then $Tr(A) = \sum_{i=1}^{n} \lambda_i$ where $\lambda_i, i = 1..n$ are the eigenvalues of $A$.

# practice: spectral theorem

## Exercise 2

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. Show that $A$ is invertible if and only if $\lambda_i \neq 0$ for all eigenvalues of $A$.

# practice: spectral theorem

**Exercise 3**

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix and let $\lambda_1, ..., \lambda_n \in \mathbb{R}$ be its eigenvalues. Show that
$A$ is positive semi-definite if and only if $\lambda_i \geq 0$ for all $i = 1..n$

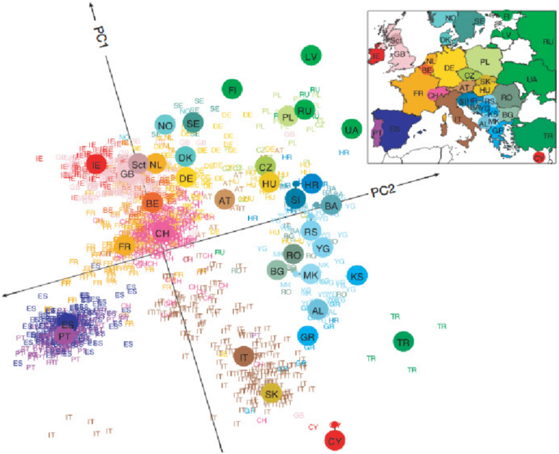
# practice: spectral theorem

**Exercise 3**

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix and let $\lambda_1, ..., \lambda_n \in \mathbb{R}$ be its eigenvalues. Show that
$A$ is positive semi-definite if and only if $\lambda_i \geq 0$ for all $i = 1..n$

## PCA

Problem: we have data $a_1, .., a_n \in \mathbb{R}^d$ with $d$ very large. We wish to represent the dataset in a lower dimension.



# recall: PCA

## Method

To perform PCA on a dataset we have to:

1. Calculate th sample covariance matrix $S$
2. Find the eigenvector $v_1$ of $S$ with largest eigenvalue. $v_1$ is the first singular vector
3. Continue finding as many eigenvectors $v_k$ of $S$ as necessary by decreasing order with respect to the respective eigenvalues $\lambda_k$.
4. The dimensionally reduced data set will be composed by the projection of the data into the $v_1, ..v_k$ eigenbasis.

# practice: PCA

### Exercise 4

Suppose there are two eigenvectors of the covariance matrix that correspond to large eigenvalues and the rest of eigenvalues are small. How do we interpret this geometrically?

# recall: Singular Value decomposition

**SVD**

Let $A \in \mathbb{R}^{n \times m}$. Then there exists two orthogonal matrices $U \in \mathbb{R}^{n \times n}, V \in \mathbb{R}^{m \times m}$ and a matrix $\Sigma \in \mathbb{R}^{n \times m}$ such that

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

with $\Sigma_{11}, \Sigma_{22}, ..., \Sigma_{11} \geq 0$ and $\Sigma_{ij} \neq 0$

# practice: SVD

**Exercise 5**

Let $A \in \mathbb{R}^{n \times m}$ give a method to compute numerically $rank(A)$ using SVD.

**Exercise 6**

Let $A \in \mathbb{R}^{n\times m}$ explain why the set $\{Ax | \|x\| = 1\}$ is an ellipsoid.

# Announcement

**Next recitation session on Wed 28th October will be a midterm review.**

**The office hours on Thr 29th October is moved to Wed 28th 9 AM EST after the recitation**