Linear Algebra

20. 12. 2023

Afonso Bandeira

* Please take a look at the typed lecture notes

* The full notes are now available!!

$A \in \mathbb{R}^{m \times n}$

Prop 6.3.16. $A \in \mathbb{R}^{m \times n}$

$n \times n$ A^TA are both symmetric.

$m \times m$ AA^T both PSD

have the same non-zero eigenvalues.

Proof: $(A^TA)^T = A^T A^{T^T} = A^TA$. Same for AA^T ✓

$\forall_x$: $x^T A^T A x = (Ax)^T (Ax) = \|Ax\|^2 \geq 0$ ✓

let r be the rank of A. $\text{rank}(A) = \text{rank}(A^T) =$
$= \text{rank}(A^TA) = \text{rank}(AA^T)$.

A^TA is sym, and has rank r.

so it has r non-zero e-values $\lambda_1, \dots, \lambda_r$

let $v_1, \dots, v_r$ be the corresponding e-vectors.

$1 \leq k \leq r$ $\quad A^TA v_k = \lambda_k v_k$

$\Downarrow$

$A\underbrace{A^TA v_k} = \lambda_k \underbrace{A v_k}$ so Av_k is an e-vector of AA^T with e-value λ_k.

since $Av_k \neq 0$. (if $Av_k = 0$ then $A^TAv_k = 0$ ↯)

$k \neq j$.

$$(Av_k)^T(Av_j) = v_k^T A^T A v_j$$
$$= \lambda_j v_k^T v_j = 0.$$

and so $Av_1, \dots, Av_r$ is an orthogonal set of e-vectors of AA^T with e-values $\lambda_1, \dots, \lambda_r$.

Prop. 6.3.17. Cholesky decomp.

Every sym PSD matrix M is the gram matrix of an upper triangular matrix C

$$M = C^T C \quad \text{for } C \text{ upper triangular}$$

<u>Proof:</u> by the spectral Thm.

$$M = V \Lambda V^T$$

V is orthogonal
Λ is diag.
$\Lambda_{ii} \geq 0$.
$\Lambda^{\frac{1}{2}}$ diag with the sq. root of the e-values in the diagonal.

$$= (V\Lambda^{\frac{1}{2}})(V\Lambda^{\frac{1}{2}})^T$$

$$(V\Lambda^{\frac{1}{2}})^T = QR \leftarrow \text{Q-R decomp.}$$

$$M = (QR)^T(QR)$$
$$= R^T Q^T Q R$$
$$= R^T R \quad \text{and } R \text{ is upper triangular.}$$

Singular Value Decomposition
SVD

$A \in \mathbb{R}^{m \times n}$ can be written as

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

U is $m \times m$ orthogonal, columns are o.n.b ($u_1, \ldots, u_m \in \mathbb{R}^m$) and are called left singular vectors

V is $n \times n$ orthogonal, " " " " ($v_1, \ldots, v_n \in \mathbb{R}^n$) " " " right singular vectors.

Σ is $m \times n$ diagonal

$\Sigma_{ij} = 0$ if $i \neq j$.

$\Sigma_{ii} = \sigma_i \in \mathbb{R}$ are the singular values of A

$\sigma_1 \geq \cdots \geq \sigma_{\min\{n,m\}} \geq 0$

$$[\; A \;] = [\; U \;]\left[\begin{array}{c|c} \ddots & \Sigma \; 0 \end{array}\right][\; V^T \;]$$

$$= [\; U_r \;][\; \Sigma_r \;][\; V_r^T \;]$$

Compact form of the SVD

$r = \text{rank}(A)$ $\qquad A$ $m \times n$

then we can write

$$A = U_r \Sigma_r V_r^T$$

U_r is $m \times r$ $\qquad U_r^T U_r = I$

Σ_r is $r \times r$ diag $\qquad \Sigma_{ii} = \sigma_i$

V_r is $n \times r$ $\qquad V_r^T V_r = I$

* needs saving $r \times m + r + r \times n$ numbers.

||

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

$$AA^T = U \Sigma V^T V \Sigma^T U^T = U \underbrace{\left(\Sigma \Sigma^T \right)}_{\text{diag.}} U^T$$

$u_1, \dots, u_m$ are the e.vectors of AA^T

$$(\Sigma \Sigma^T)_{ii} = \sigma_i^2 .$$

if $m \leq n$ then $(\Sigma\Sigma^T)_{ii} = \sigma_i^2$

if $m > n$ $(\Sigma\Sigma^T)_{ii} = \sigma_i^2$ if $i \leq n$

$= 0$ o.w.

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

$$= V(\Sigma^T\Sigma)V^T \longleftarrow \text{eigendecomposition of } A^TA.$$

left s.v. of A are e.vectors of AA^T
right s.v. of A are e-vector of A^TA
sing val. of A are square root of e-values of AA^T or A^TA taking the $\min\{n, m\}$ of them.

rank (A) = # non-zero sing. values of A.

Proof (of existence of SVD)

let $r = \text{rank}(A)$ we will build the compact SVD

we want $A = U_r \Sigma_r V_r^T$

we know $AA^T = U \Lambda U^T \longleftarrow$ eigendecomposition of AA^T

since $\text{rank}(AA^T) = r$ then AA^T has r non-zero e-values.

$$AA^T = \begin{bmatrix} U \end{bmatrix}\begin{bmatrix} \Lambda \end{bmatrix}\begin{bmatrix} U^T \end{bmatrix}$$

$$AA^T = [U_r][\Lambda_r][U_r^T]$$ where Λ_r is diag with r non-zero e-val. and U_r has as col. corresponding e-vectors.

$$U_r^T U_r = I \quad (r \times r)$$

Σ_r is diag and $\Sigma_r = \Lambda_r^{1/2}$ (recall that Λ_r is diag. $(\Lambda_r)_{ii} > 0$)

we need to build a matrix V_r $n \times r$ s.t.

$$A = U_r \Sigma_r V_r^T$$

"Candidate V_r".

if $A = U_r \Sigma_r V_r^T$ then we would have

$$\Sigma_r^{-1} U_r^T A = \Sigma_r^{-1} U_r^T U_r \Sigma_r V_r^T = V_r^T$$

Candidate: $V_r = (\Sigma_r^{-1} U_r^T A)^T$.

Now we check that the candidate is good:

① we need $V_r^T V_r = I$.

$$V_r^T V_r = \Sigma_r^{-1} U_r^T \underbrace{A A^T}_{= U_r \Lambda_r U_r^T} U_r \Sigma_r^{-1}$$

$$= \Sigma_r^{-1} \underbrace{U_r^T U_r}_{I} \Lambda_r \underbrace{U_r^T U_r}_{I} \Sigma_r^{-1} =$$

$$= \Sigma_r^{-1} \Lambda_r \Sigma_r^{-1} = I_{r\times r}$$ ✓

② We need

$$A = U_r \Sigma_r V_r^T.$$

$$U_r \Sigma_r V_r^T = U_r \Sigma_r \left(\Sigma_r^{-1} U_r^T A \right) =$$

$$\begin{array}{cc} & r \\ m & \begin{bmatrix} U_r \end{bmatrix} \end{array} \quad U_r^T U_r = I$$

$$= \underbrace{U_r U_r^T}_{\text{Proj } C(U_r).} A.$$

$$= \text{Proj } C(AA^T)$$

$$C(AA^T) = C(A)$$

and so $U_r U_r^T = \text{Proj } C(A)$

and so $U_r U_r^T A = A.$ $\square$

Ultimate fact in L.A.: Any matrix A can be written as

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

"any matrix is diagonal when viewed in the orthonormal basis of the sing. vectors".

HW: Write the A^+ in terms of SVD of A.

One can write

$$A = \sum_{k=1}^{r} \sigma_k u_k v_k^T$$

CS lens
Principal Component Analysis

people

10000

bases pairs

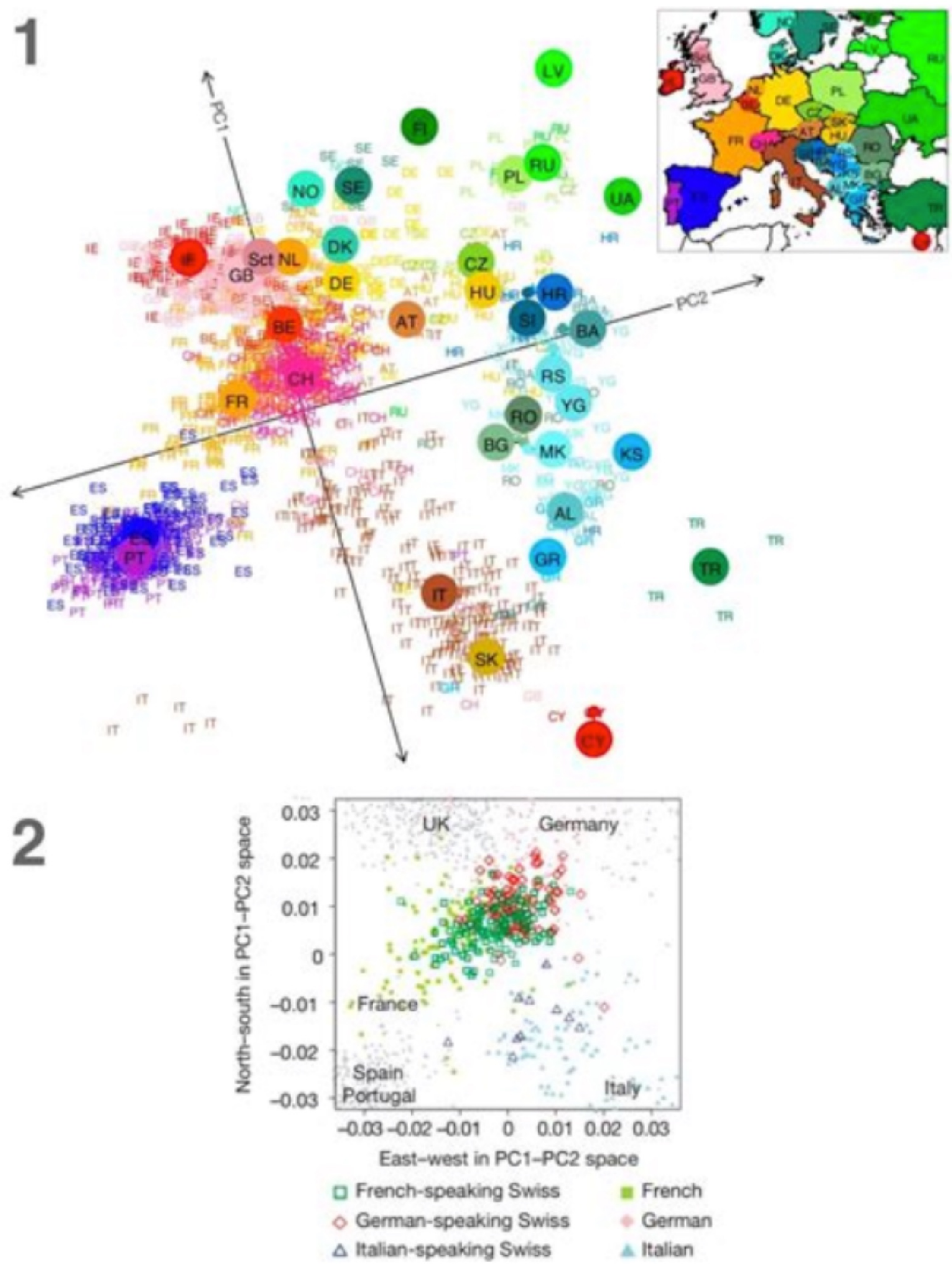
2000

$$[\;] = U \Sigma V^T$$

but keep only the largest 2 sing. val.

$$= U_2 \Sigma_2 V_2^T$$

$$= \underbrace{\Big[\; U_2^T \;\Big]}_{2000 \times 2} \Big[\; \cdot \;\; \cdot \;\Big] \underbrace{\Big[\; V_2^T \;\Big]}_{2 \times 10000}$$



courtesy: John Novembre, UCLA

See references in typed lecture notes.