Principal Components Analysis (PCA)

Mathematical Tools for Neuroscience (NEU 314)
Fall, 2021

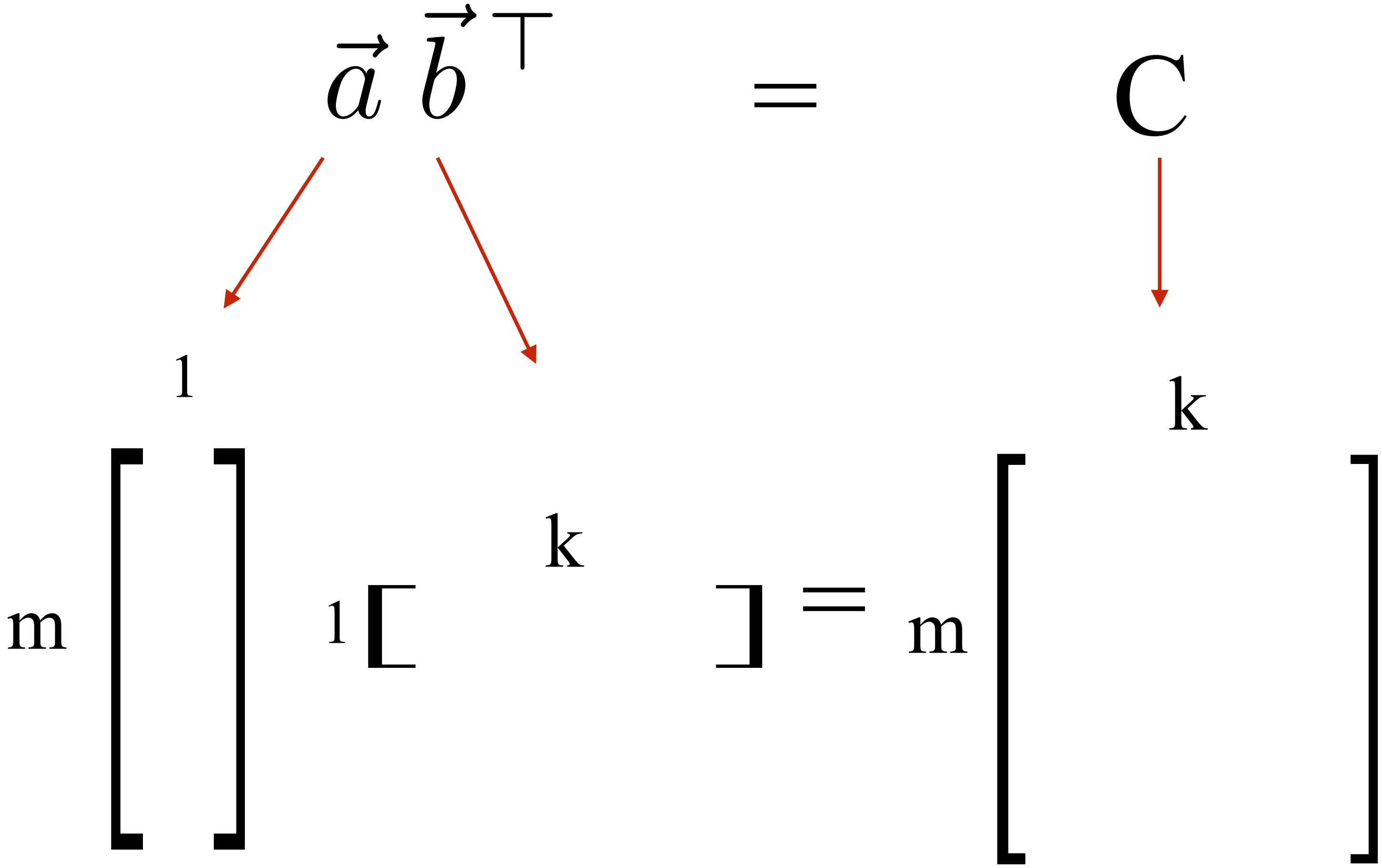
lecture 11

Jonathan Pillow

Summary of prev (online-only) lecture

- outer product (review)
- SVD as a sum of weighted outer products
- optimal low-rank matrix approximation using SVD
- Frobenius norm (Euclidean norm for matrices)

quick review: **outer product**



$$\vec{a}\,\vec{b}^{\top} = C$$

- produces a rank-1 matrix

SVD as a sum of outer-products

$$A = USV^\top$$

$$\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_n \end{bmatrix} \begin{bmatrix} - & v_1 & - \\ - & v_2 & - \\ & \vdots & \\ - & v_n & - \end{bmatrix}$$

$$= s_1 \underbrace{\boxed{u_1 v_1^T}}_{\begin{matrix} | \\ u_1 \\ | \end{matrix} \begin{matrix} - & v_1 & - \end{matrix}} + s_2 \boxed{u_2 v_2^T} + \cdots + s_n \boxed{u_n v_n^T}$$

(rank 1 matrix)

matrix approximation

- the best rank-K approximation to A (in terms of squared error) is given by truncating the SVD after K terms.

$$= s_1 \boxed{u_1 v_1^T} + \cdots + s_K \boxed{u_K v_K^T}$$

$$= \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix} \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_k \end{bmatrix} \begin{bmatrix} — & v_1 & — \\ & \vdots & \\ — & v_k & — \end{bmatrix}$$

(this is an easier way to compute it)

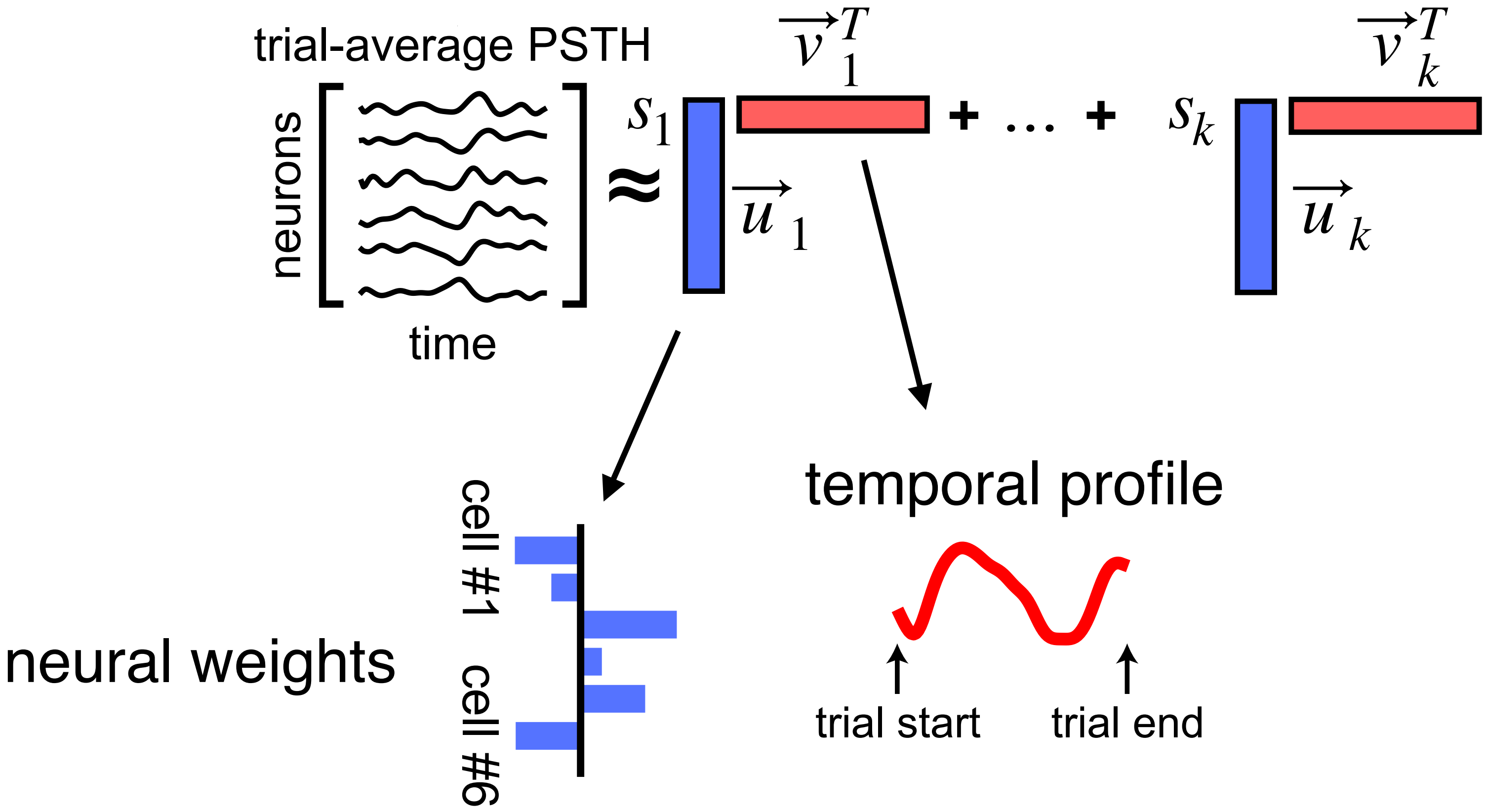
Fraction of variance accounted for
(by the rank-K approximation):

$$\frac{\sum_{i=1}^{K} s_i^2}{\sum_{j=1}^{N} s_j^2}$$

sum of squared first K singular values

sum of squares of all singular values

applications to neural data

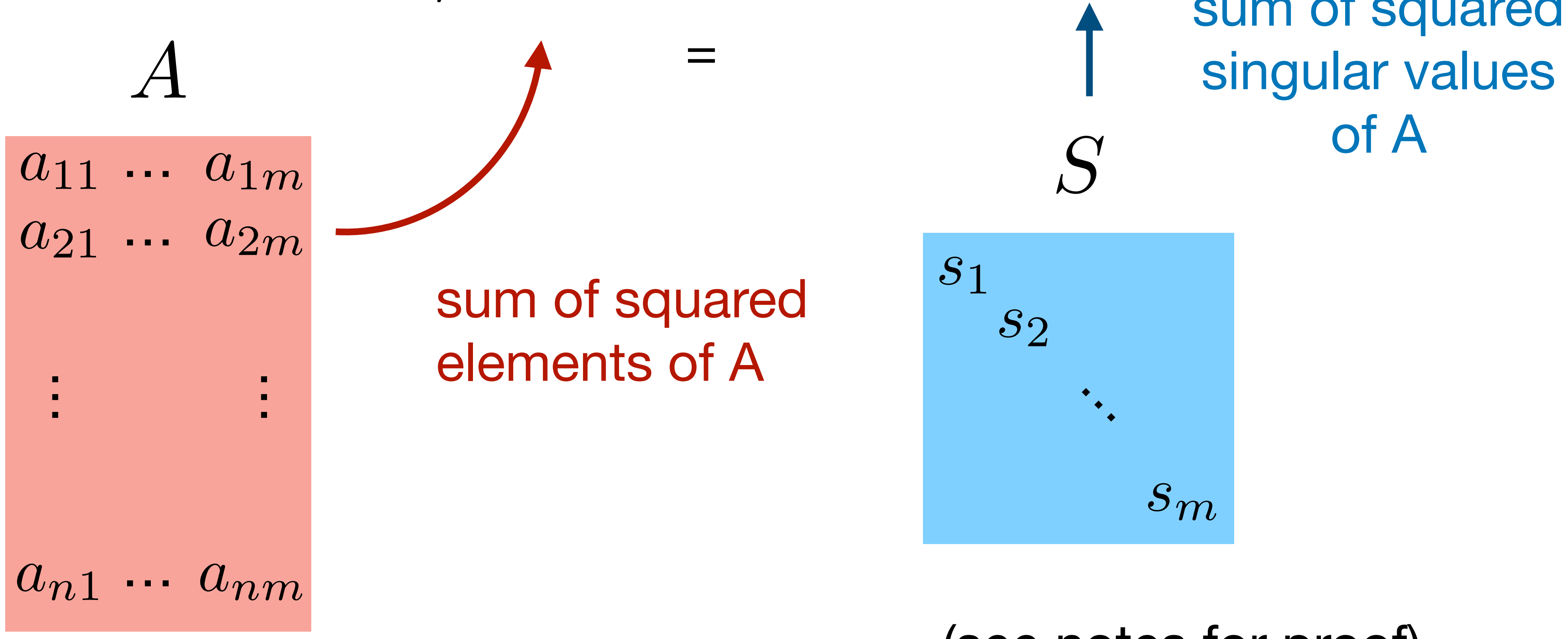


(adapted from Williams *et al*, Neuron 2018)

Frobenius norm

(the Euclidean norm for matrices)

$$||A||_F = \sqrt{\sum_{ij} a_{ij}^2} = \sqrt{\sum_i s_i^2}$$



$$A = \begin{matrix} a_{11} & \cdots & a_{1m} \\ a_{21} & \cdots & a_{2m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{matrix}$$

sum of squared elements of A

=

$$S = \begin{matrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_m \end{matrix}$$

sum of squared singular values of A

(see notes for proof)

PCA

PCA summary

the data

$$X = \overbrace{\begin{bmatrix} — & \vec{x}_1 & — \\ — & \vec{x}_2 & — \\ & \vdots & \\ — & \vec{x}_N & — \end{bmatrix}}^{d} \left.\vphantom{\begin{bmatrix} a \\ a \\ a \\ a \end{bmatrix}}\right\} N$$

2nd moment matrix

$$C = X^\top X$$

SVD

$$C = USU^\top$$

first k PCs: $\{u_1, \ldots u_k\}$

sum of squares of data within subspace: $s_1 + \cdots + s_k$

PCA summary

the data

$$X = \overbrace{\left.\begin{bmatrix} — \vec{x}_1 — \\ — \vec{x}_2 — \\ \vdots \\ — \vec{x}_N — \end{bmatrix}\right\} N}^{d}$$

2nd moment matrix

$$C = X^\top X$$

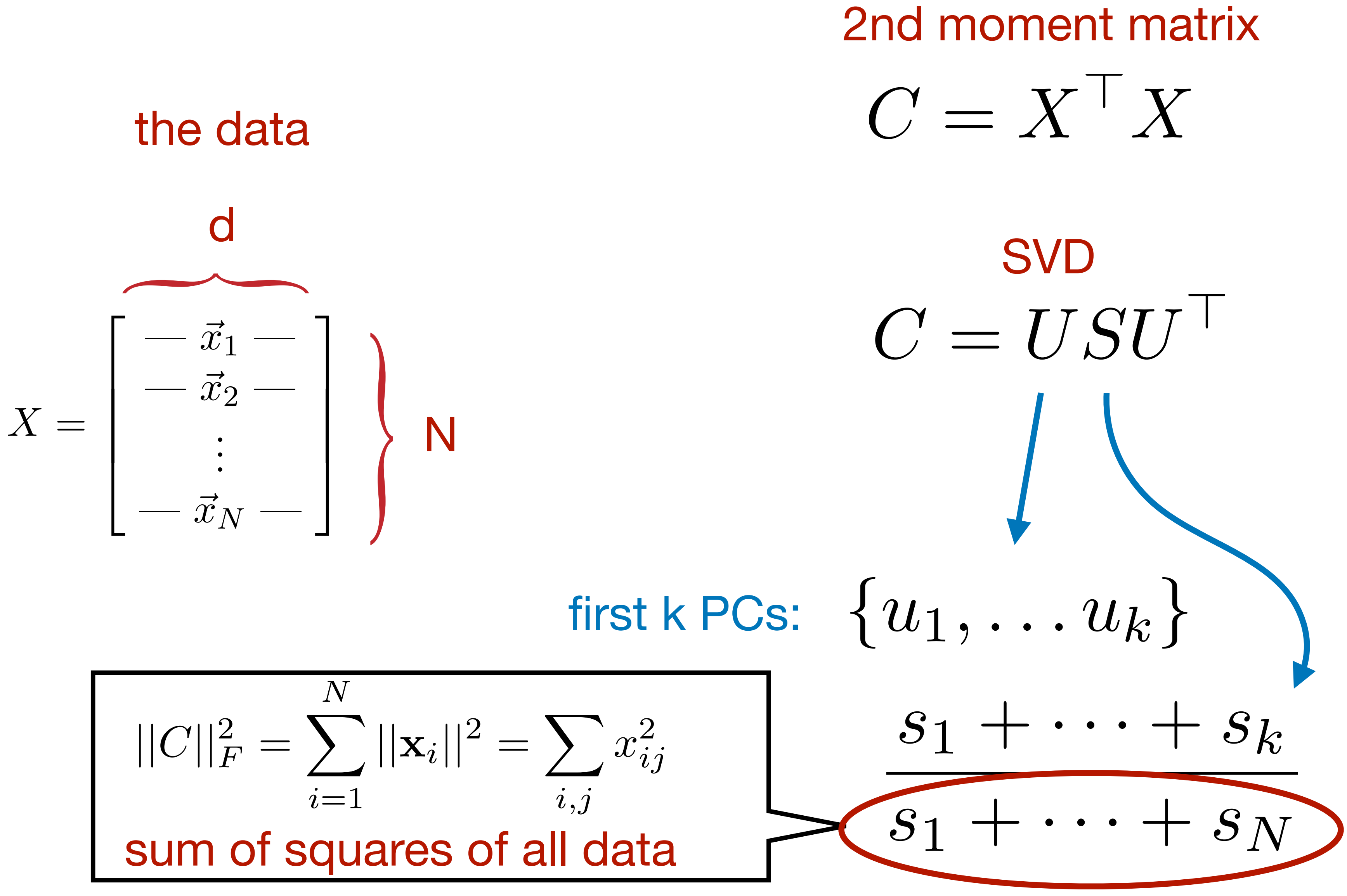
SVD

$$C = USU^\top$$

first k PCs: $\{u_1, \ldots u_k\}$

fraction of sum of squares: $\dfrac{s_1 + \cdots + s_k}{s_1 + \cdots + s_N}$

PCA summary



the data

$$X = \left.\overbrace{\begin{bmatrix} — \vec{x}_1 — \\ — \vec{x}_2 — \\ \vdots \\ — \vec{x}_N — \end{bmatrix}}^{d}\right\} N$$

2nd moment matrix

$$C = X^\top X$$

SVD

$$C = USU^\top$$

first k PCs: $\{u_1, \ldots u_k\}$

$$\frac{s_1 + \cdots + s_k}{s_1 + \cdots + s_N}$$

$$||C||_F^2 = \sum_{i=1}^{N} ||\mathbf{x}_i||^2 = \sum_{i,j} x_{ij}^2$$

sum of squares of all data

Full derivation of PCA: see notes

Two equivalent formulations:

1. $$\hat{B}_{pca} = \arg\max_B ||XB||_F^2$$

$$\text{such that } B^\top B = I$$

find subspace that preserves *maximal* sum-of-squares

2. $$\hat{B}_{pca} = \arg\min_B ||X - \underbrace{XBB^\top}||_F^2$$

$$\text{such that } B^\top B = I$$

minimize sum-of-squares of orthogonal component

reconstruction of X in subspace spanned by B