SVD Applications 2: low-rank matrix approximation

Mathematical Tools for Neuroscience (NEU 314)
Fall, 2021

lecture 10 (online only)

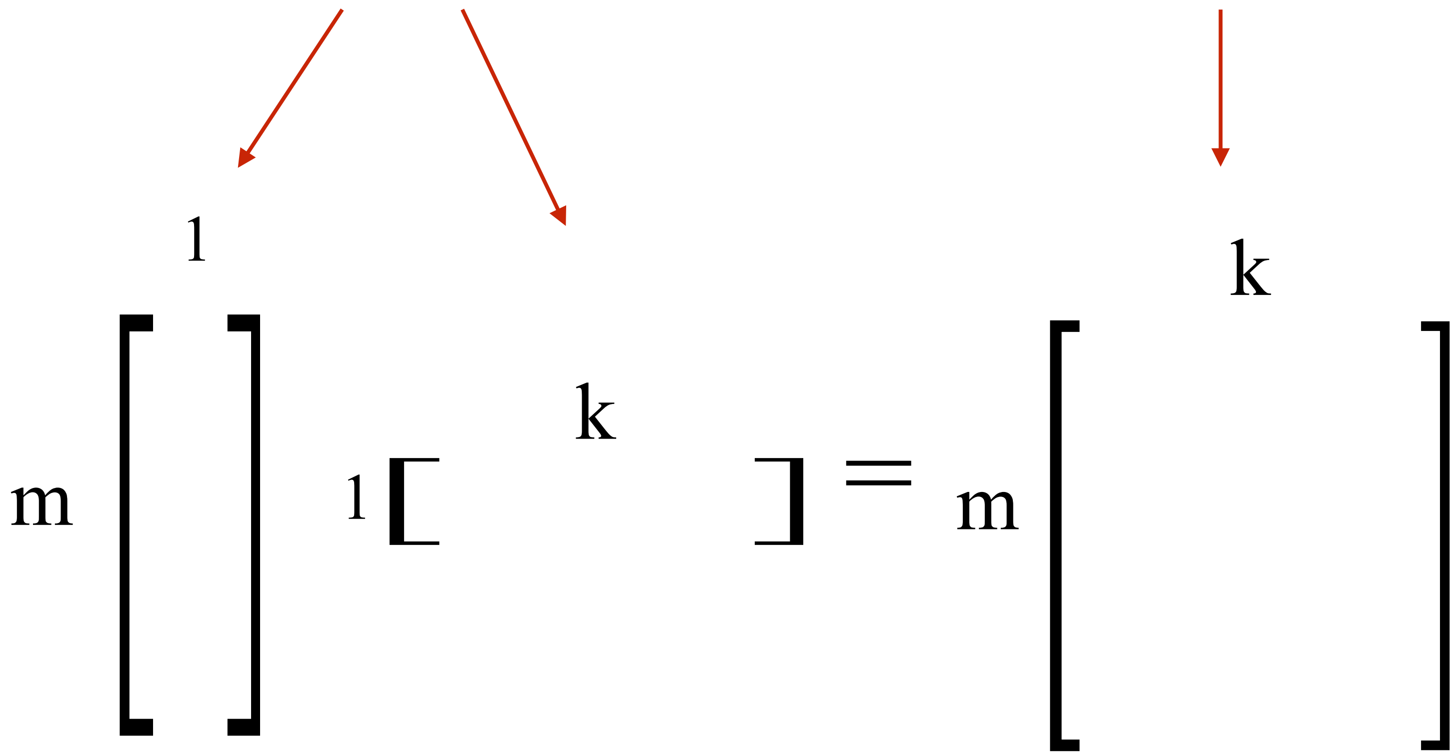
Jonathan Pillow

quick review: **outer product**

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

quick review: **outer product**

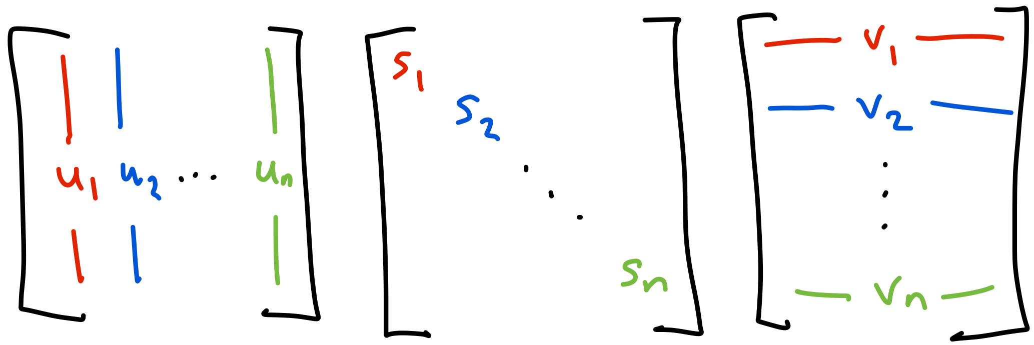
$$\vec{a}\,\vec{b}^{\top} = C$$



- produces a rank-1 matrix

another view of SVD: a sum of outer-products

$$A = USV^\top$$



SVD as a sum of outer-products

$$A = USV^\top$$

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

$$= s_1 \boxed{u_1 v_1^\top} + s_2 \boxed{u_2 v_2^\top} + \cdots + s_n \boxed{u_n v_n^\top}$$

SVD as a sum of outer-products

$$A = USV^\top$$

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

$$= s_1 \boxed{u_1 v_1^T} + s_2 \boxed{u_2 v_2^T} + \cdots + s_n \boxed{u_n v_n^T}$$

$$\underbrace{}_{\begin{bmatrix} | \\ u_1 \\ | \end{bmatrix} \begin{bmatrix} — & v_1 & — \end{bmatrix}} \quad \text{(rank 1 matrix)}$$

matrix approximation

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

$$= s_1 \, u_1 v_1^T + \cdots + s_K \, u_K v_K^T$$

matrix approximation

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

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

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

(this is an easier way to compute it)

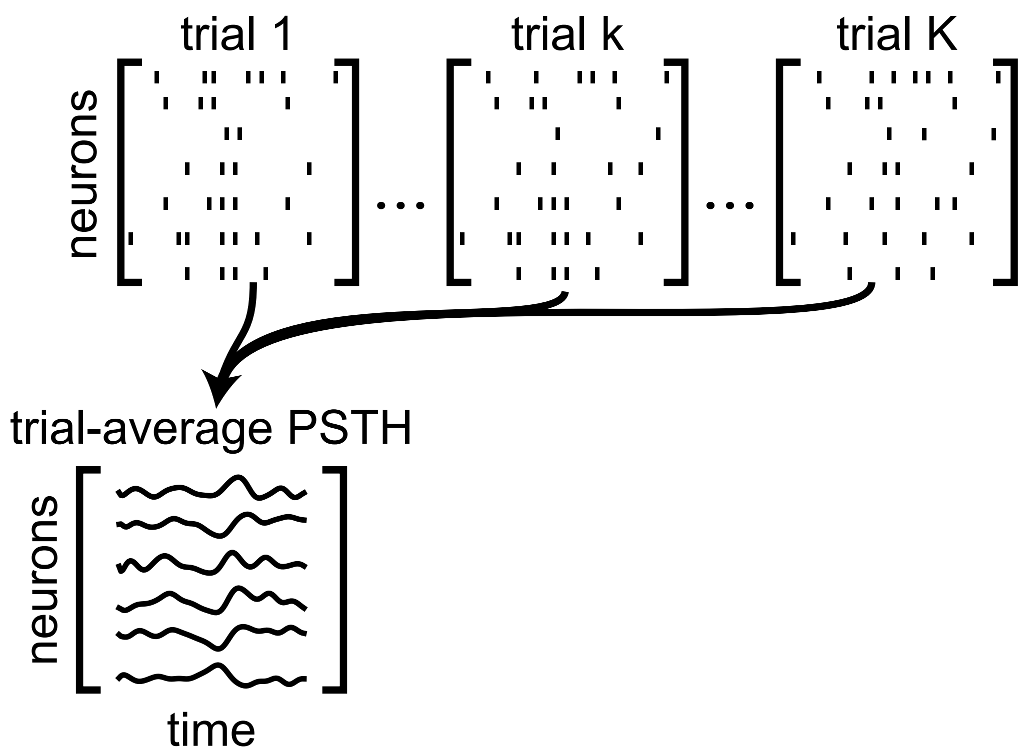
Fraction of variance accounted for (by the rank-K approximation):

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

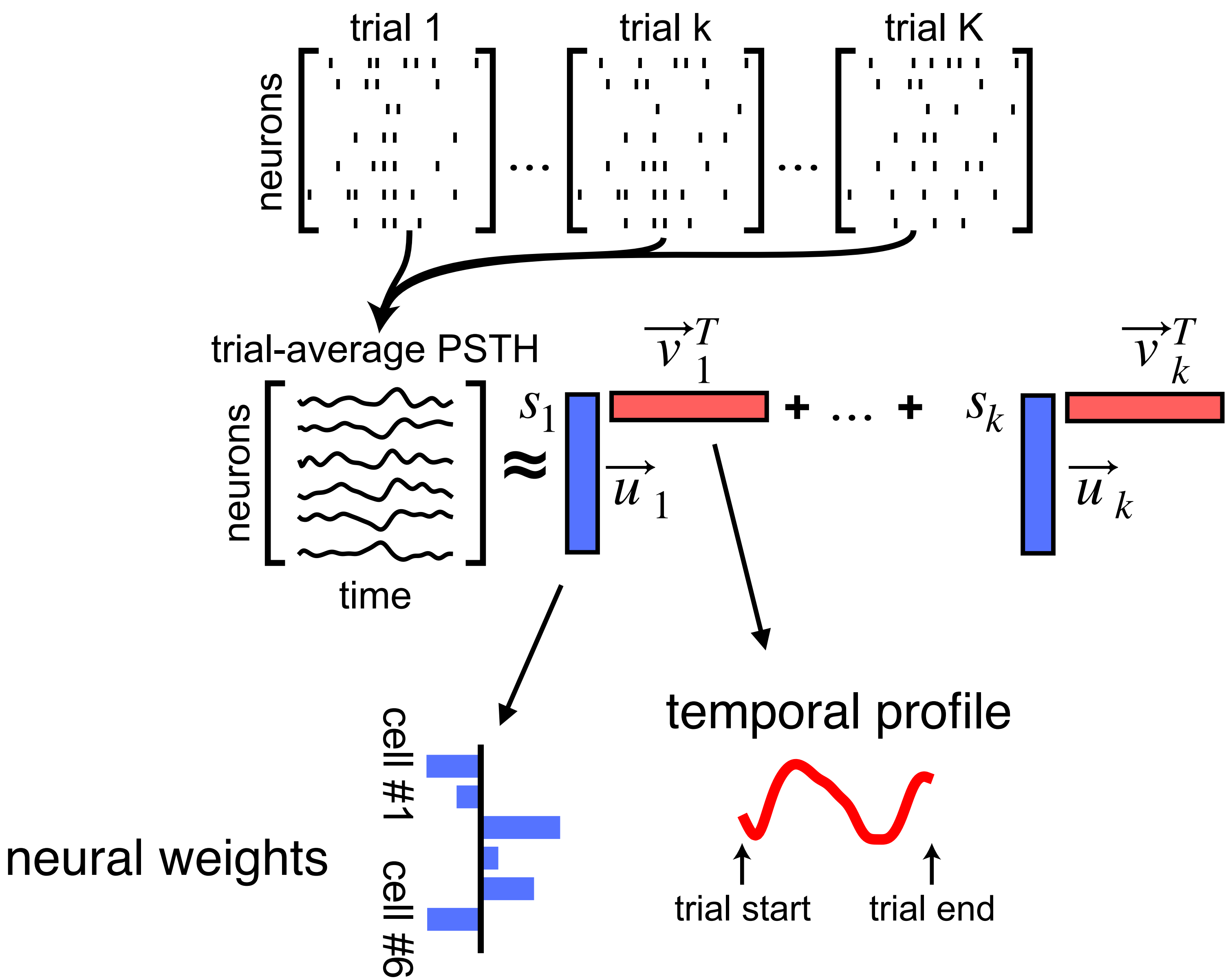
sum of squared first K singular values

sum of squares of all singular values

applications to neural data

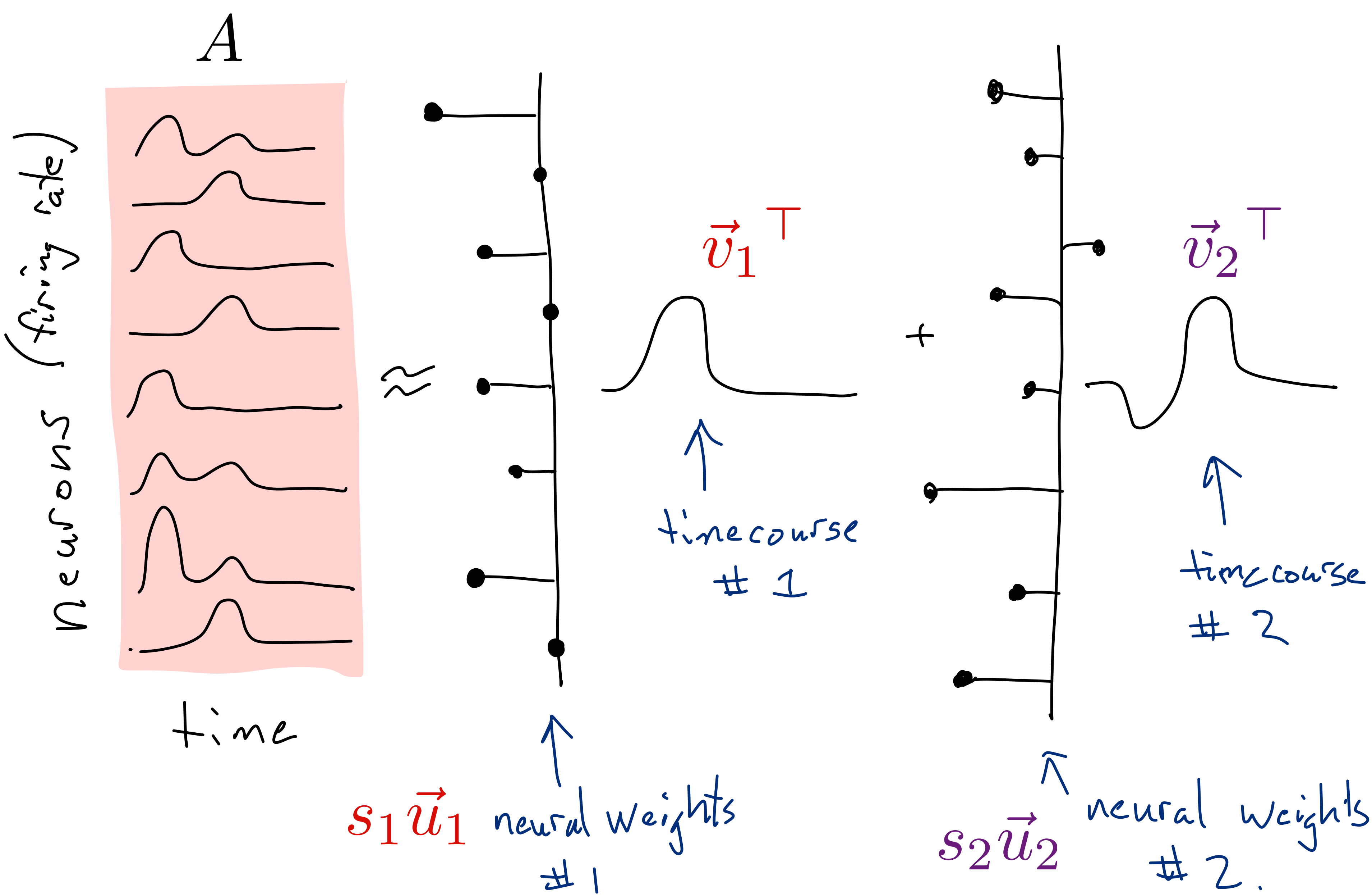


applications to neural data



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

my (admittedly poor) attempt: rank 2 approximation



Frobenius Norm

Frobenius norm

(the Euclidean norm for matrices)

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

$$A$$

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

sum of squared
elements of A

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

=

sum of squared singular values of A

S

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

(see notes for proof)

Thus we can also write the "fraction of variance" accounted for by the rank-K approximation as:

$$\frac{\sum_{i=1}^{K} s_i^2}{||A||_F^2}$$

sum of squared first K singular values

sum of squares of all singular values

Summary

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