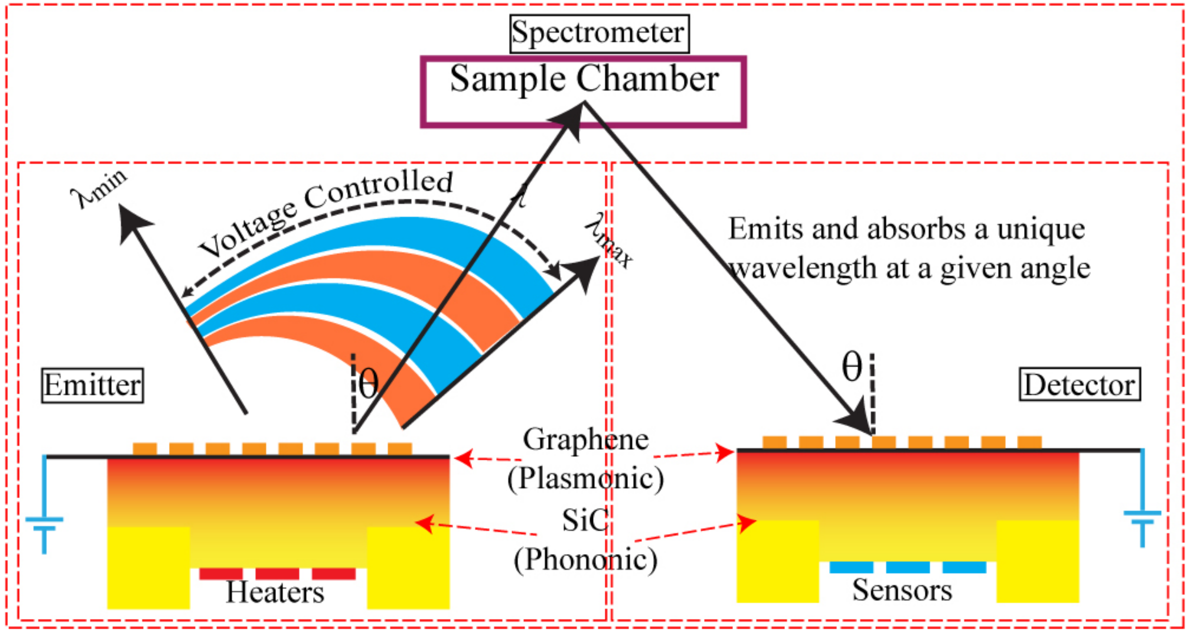
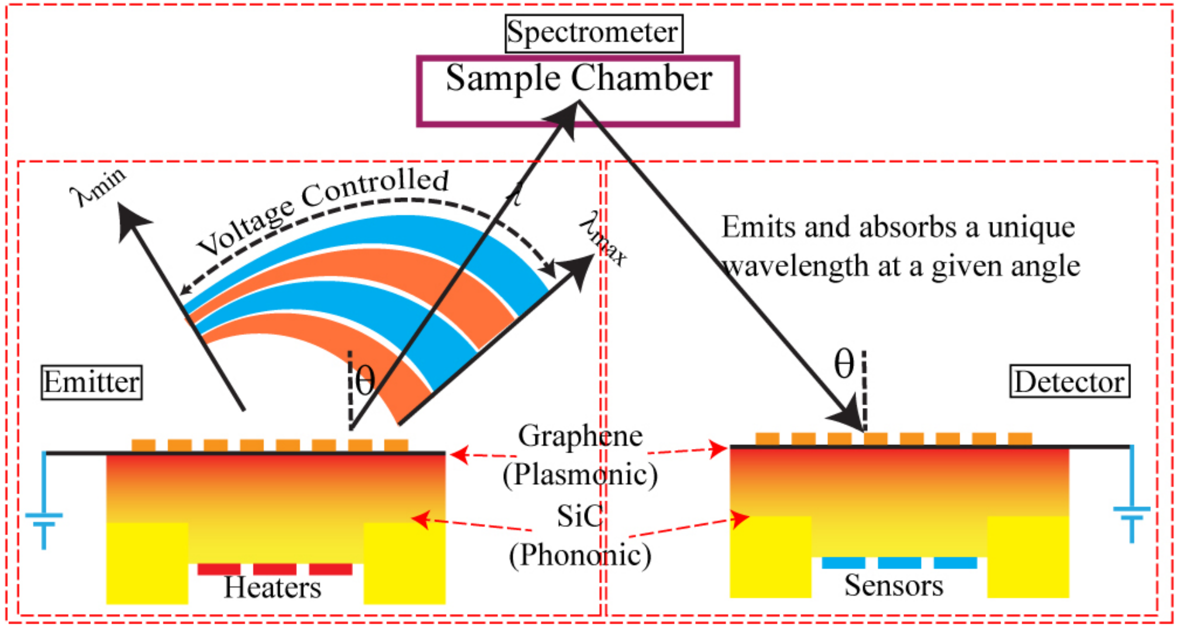
DS 5020
Introduction to Linear Algebra, Statistics, and Probability

Lecture 7: Singular Value Decomposition





Sort the information in the order of importance

Singular Value Decomposition

Assume A as an $\mathrm{m} \times n$ matrix

$\mathrm{A}A^T$ is and $m \times m$ matrix

AA^T is Symmetric

Let u_i be the eigenvectors of $AA^T$with eigenvalues $\lambda_i (written\ as\ \sigma_i^2) \Rightarrow AA^T u_i = \sigma_i^2 u_i$

Chose $\mathrm{v}_i = \frac{A^T u_i}{\sigma_i}$

$$
\begin{aligned}
A^T A v_i &= A^T A \frac{A^T u_i}{\sigma_i} \\
&= A^T \frac{AA^T u_i}{\sigma_i} \\
&= A^T \frac{\sigma_i^2 u_i}{\sigma_i} \\
&= \sigma_i^2 \left(\frac{A^T u_i}{\sigma_i} \right) \\
&= \sigma_i^2 v_i
\end{aligned}
$$

$$A^T A v_i = \sigma_i^2 v_i$$

So v_i will be the eigenvectors of $A^T A$with eigenvalues $\lambda_i (written\ as\ \sigma_i^2)$

$$\mathrm{v}_i = \frac{A^T u_i}{\sigma_i} \quad \Rightarrow \quad A\mathrm{v}_i = \frac{AA^T u_i}{\sigma_i} \quad \Rightarrow \quad A\mathrm{v}_i = \frac{\sigma_i^2 u_i}{\sigma_i} = \sigma_i u_i$$

$$A\mathrm{v}_i = \sigma_i u_i$$
$$A^T u_i = \sigma_i v_i$$

$$A\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_1 & v_2 & v_3 & \cdots & \cdots & v_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_1 & u_2 & u_3 & \cdots & \cdots & u_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \sigma_3 & & \\ & & & & \\ & & & & \sigma_n \end{bmatrix}$$

$$AV = U\Sigma$$

$$A = U\Sigma V^{-1}$$
$$A = U\Sigma V^T \longrightarrow$$ Singular value decomposition (not exactly)

$Av_i = \sigma_i u_i \Rightarrow$

$$\underset{m \times n}{A} \underset{n \times r}{\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_r \\ | & | & & | \end{bmatrix}} = \underbrace{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_r \\ | & | & & | \end{bmatrix}}_{m \times r} \underset{r \times r}{\begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \ \sigma_r \end{bmatrix}}$$

$AV_r = U_r \Sigma_r$. Add orthogonal $n-r$ vectors fror null space of A $(v_{r+1}, \dots, v_n)$, and $m-r$ vectors from null space of A^T $(u_{r+1}, \dots, u_m)$;.

$$\underset{m \times n}{A} \underset{n \times n}{\begin{bmatrix} | & | & | & & | \\ v_1 & v_2 & v_3 & \cdots & v_n \\ | & | & | & & | \end{bmatrix}} = \underset{m \times m}{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{bmatrix}} \underset{m \times n}{\begin{bmatrix} \sigma_1 & & & \overbrace{}^{n-r \text{ zero rows.}} \\ & \sigma_2 & & 0 \\ & & \ddots \ \sigma_r & \\ & 0 & & 0 \end{bmatrix}} \}m-r \text{ zero rows}$$

$$AV = U\Sigma \quad \Rightarrow \quad A = U\Sigma V^T$$

$$= \underbrace{u_1 \sigma_1 v_1^T + \cdots + u_r \sigma_r v_r^T}_{r \text{ matrices of rank } 1.}$$

Rank 1: $A = uv^T$

Rank 2: $A = u_1 v_1^T + u_2 v_2^2$ where u_1, u_2 are LI.

Rank r: $A = u_1 v_1^T + \ldots + u_r v_r^T$ where $u_1, u_2, \ldots, u_r$ are LI.

SVD idea: Write A as a sum of rank 1 matrices. And use eigenvectors of AA^T and A^TA to define those.

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

$$AA^T = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad A^TA = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\det(AA^T - \lambda I) = 0 \Rightarrow \lambda_1 = \frac{3+\sqrt{5}}{2} \qquad \lambda_2 = \frac{3-\sqrt{5}}{2}$$

$$\sigma_1 = \sqrt{\lambda_1} = \frac{\sqrt{5}+1}{2} \qquad \sigma_2 = \sqrt{\lambda_2} = \frac{\sqrt{5}-1}{2} \qquad \sigma_1 \sigma_2 = 1$$

$$u_1 = \begin{bmatrix} 1 \\ \sigma_1 \end{bmatrix} \quad u_2 = \begin{bmatrix} \sigma_1 \\ -1 \end{bmatrix} \qquad v_1 = \begin{bmatrix} \sigma_1 \\ 1 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1 \\ -\sigma_1 \end{bmatrix}$$ all divided by $\sqrt{1+\sigma_1^2}$

Then

$$A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} \quad \text{or} \quad A \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} \sigma_1 u_1 & \sigma_2 u_2 \end{bmatrix}$$

Example:

$$A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix} \qquad A^TA = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$$

$\sigma_1^2 = 45 \quad \sigma_2^2 = 5$, $\quad \lambda_1 = \sqrt{45} \quad \lambda_2 = \sqrt{5}$ and $\lambda_1 \lambda_2 = \det A = 15.$

$$\begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} v_1 = \sigma_1 v_1 \Rightarrow \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} v_1 = \sqrt{45} \begin{bmatrix} v_1 \end{bmatrix} \Rightarrow v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} v_2 = \sigma_2 v_2 \Rightarrow v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} / \sqrt{2} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$Av_1 = \frac{3}{\sqrt{2}} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \sqrt{45} \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \sigma_1 u_1.$$

$$Av_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \sqrt{5} \frac{1}{\sqrt{10}} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \sigma_2 u_2$$

Note that u_1 and u_2 are orthonormal.

$$A = U \Sigma V^T \quad \text{with}$$

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sqrt{45} & \\ & \sqrt{5} \end{bmatrix} \qquad V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Note: u_i's are called left singular vectors.
v_i's are called right singular vectors.
σ_i^2's are called singular values.

Repeat: SVD: Decompose $\underset{m\times n}{A} = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V}^T$ such that

- U and V are orthonormal ($U^TU = I$, $V^TV = I$)
- Σ is diagonal with r non-zero singular-values.

Steps: ⊗ v_i's (1 to r) are orthonormal eigenvectors of A^TA.
Complete V by selecting $n-r$ orthonormal vectors from null space of A $N(A)$.

⊗ $Av_i = \sigma_i u_i$ gives unit vectors $u_1, \dots, u_r$.

Note: $u_i^T u_j = \left(\frac{Av_i}{\sigma_i}\right)^T \left(\frac{Av_j}{\sigma_j}\right) = \frac{v_i^T A^T A v_j}{\sigma_i \sigma_j} = \frac{\sigma_j^2 v_i^T v_j}{\sigma_i \sigma_j} = 0$

⊗ Complete U by selecting any orthonormal basis for null space of A^T $N(A^T)$.

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

$$\left.\begin{array}{ll} u_1, \dots, u_r: & \text{orthonormal basis for } C(A) \\ u_{r+1}, \dots, u_m: & \text{orthonormal basis for } N(A^T) \end{array}\right\} \text{both} \in \mathbb{R}^m$$

$$\left.\begin{array}{ll} v_1, \dots, v_r: & \text{orthonormal basis for } C(A^T) \\ v_{r+1}, \dots, v_n: & \text{orthonormal basis for } N(A) \end{array}\right\} \text{both} \in \mathbb{R}^n$$

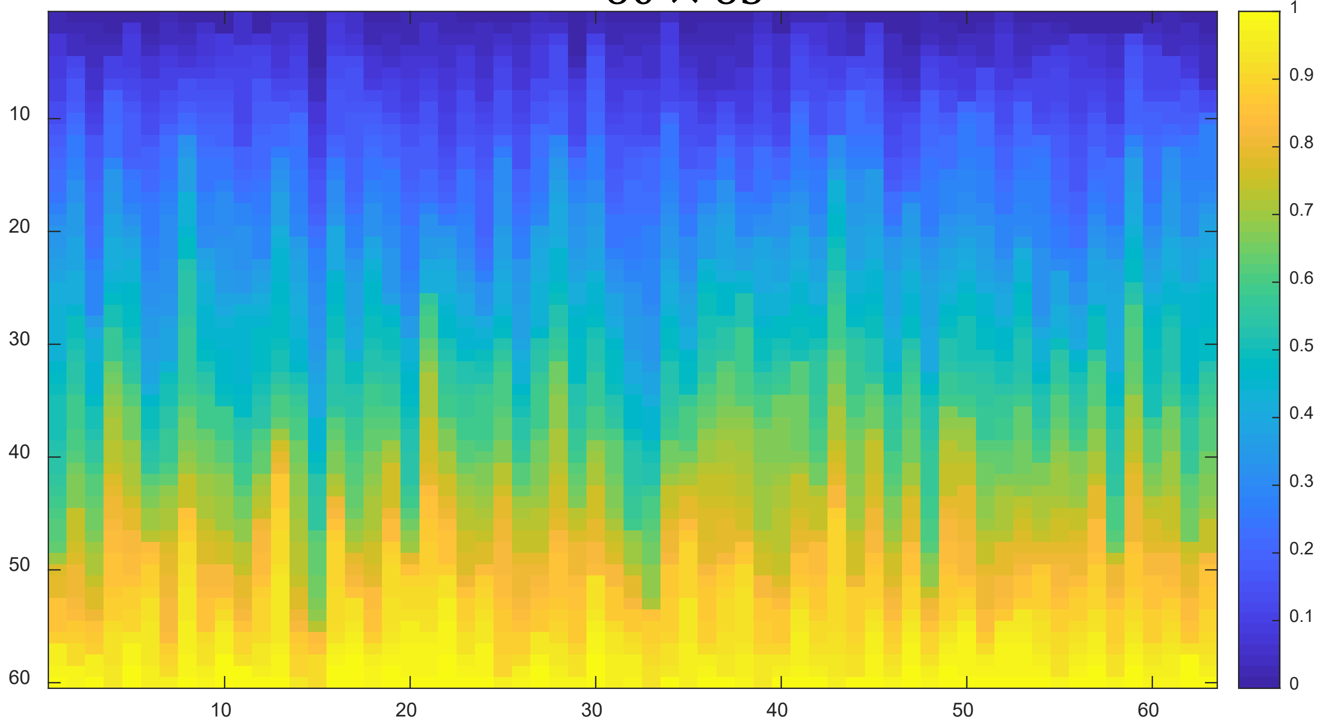
<u>Comparison:</u> (Symmetric S vs any A)

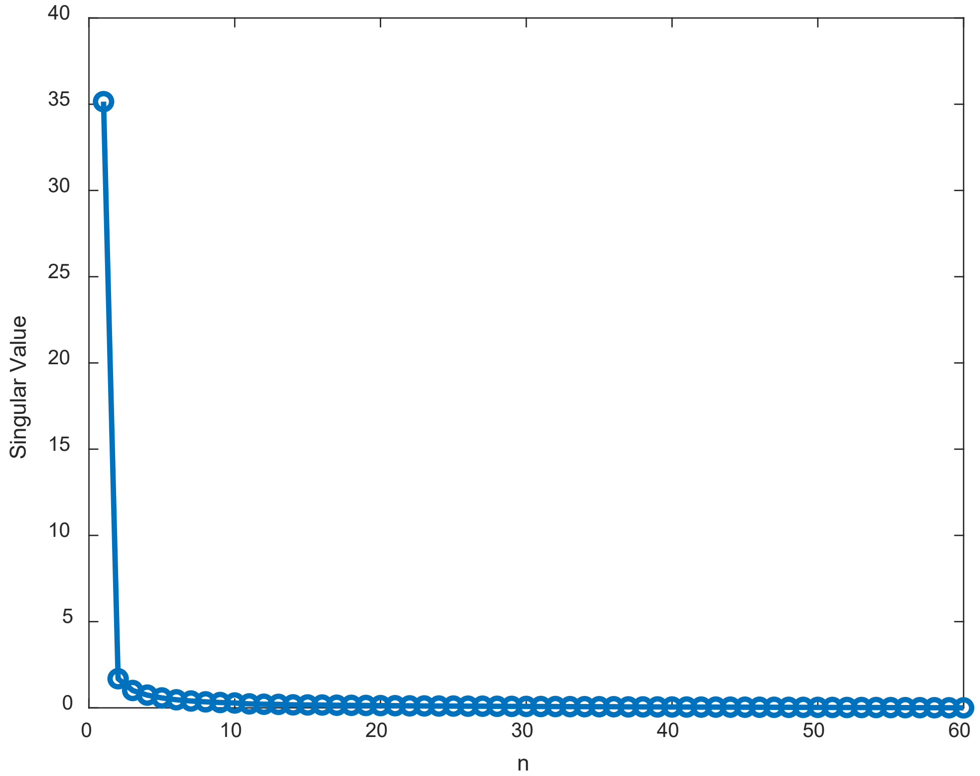
$$S = Q\Lambda Q^T = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \dots + \lambda_r q_r q_r^T$$

$$A = U\Sigma V^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T$$

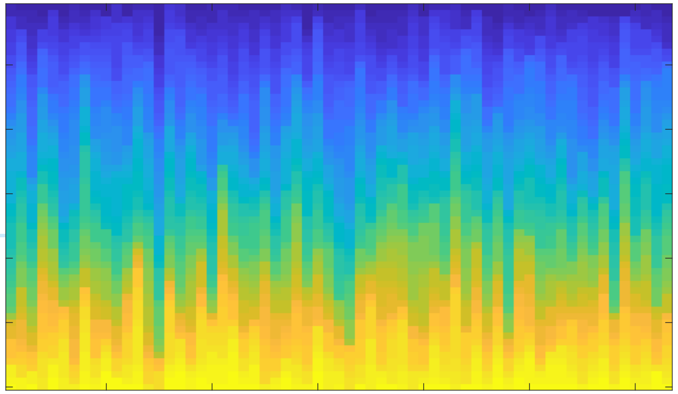
SVD-image compression-numerical example

$$A = \quad 60 \times 63$$

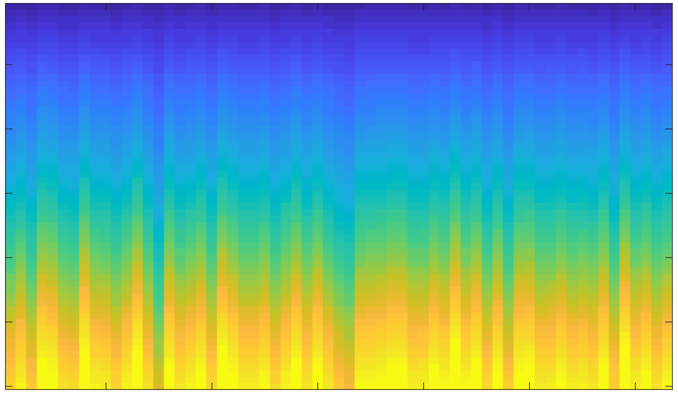




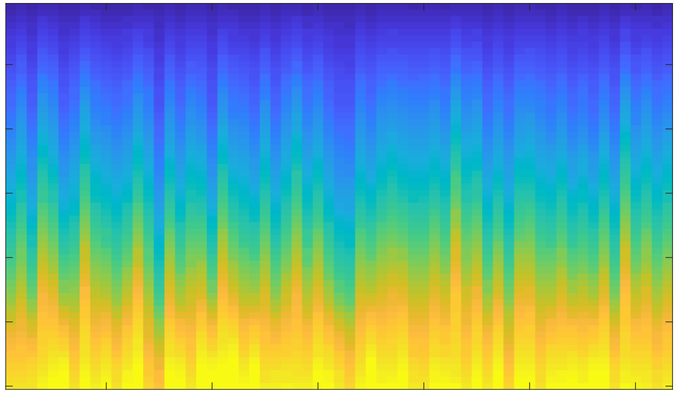
$$A = U \Sigma V^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$



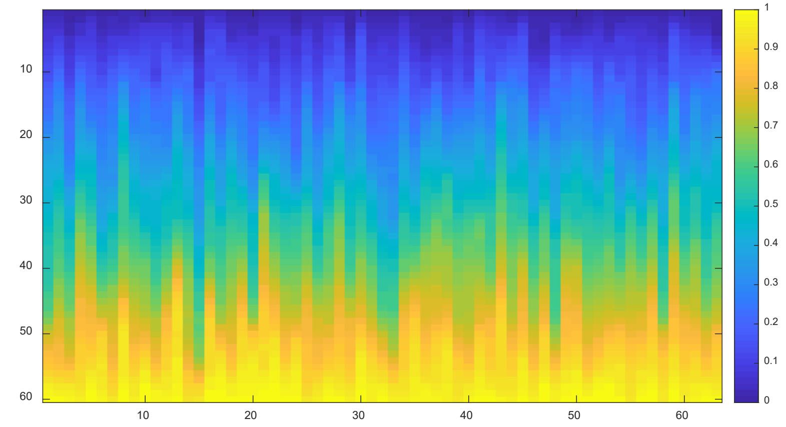
largest $\sigma_1 u_1 v_1^T$



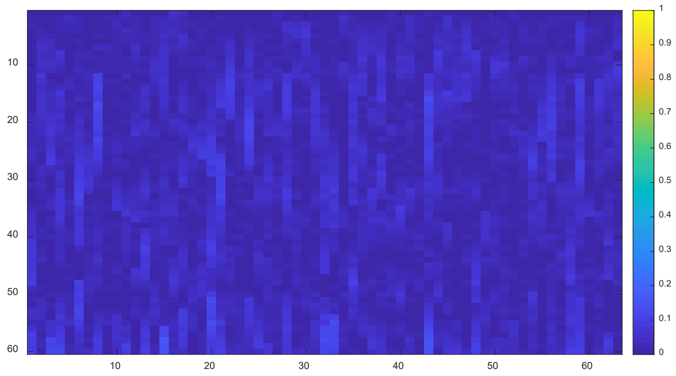
Largest+ second largest $\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$



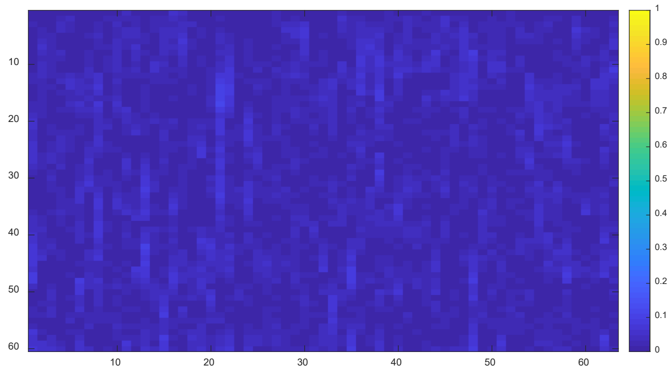
10 terms



$A - \sigma_1 u_1 v_1^T$



$A - (\sigma_1 u_1 v_1^T + +\sigma_2 u_2 v_2^T)$



$A - (10\ terms)$