



Lecture 2: The Singular Value Decomposition



Nando de Freitas

www.cs.ubc.ca/~nando/340-2009/

September 2009

Outline

The Singular Value Decomposition (SVD) is a matrix factorization that has many applications, including:

- Information retrieval,
- Least-squares problems,
- Image processing,
- Dimensionality reduction.

Eigenvalue decomposition

Let $\mathbf{A} \in \mathbb{R}^{m \times m}$. If we put the eigenvalues of \mathbf{A} into a diagonal matrix $\mathbf{\Lambda}$ and gather the eigenvectors into a matrix \mathbf{X} , then the eigenvalue decomposition of \mathbf{A} is given by

$$\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}.$$

But what if \mathbf{A} is not a square matrix? Then the SVD comes to the rescue.

Formal definition of the SVD

Given $\mathbf{A} \in \mathbb{R}^{m \times n}$, the SVD of \mathbf{A} is a factorization of the form

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

where \mathbf{u} are the left **singular vectors**, σ are the **singular values** and \mathbf{v} are the right singular vectors.

$\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ is diagonal with positive entries (singular values in the diagonal).

$\mathbf{U} \in \mathbb{R}^{m \times n}$ with orthonormal columns.

$\mathbf{V} \in \mathbb{R}^{n \times n}$ with orthonormal columns.

($\Rightarrow \mathbf{V}$ is orthogonal so $\mathbf{V}^{-1} = \mathbf{V}^T$)

Formal definition of the SVD

The equations relating the right singular values $\{\mathbf{v}_j\}$ and the left singular vectors $\{\mathbf{u}_j\}$ are

$$\mathbf{A}\mathbf{v}_j = \sigma_j\mathbf{u}_j \quad j = 1, 2, \dots, n$$

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}$$

$$\mathbf{A} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \dots & \\ & & & \sigma_n \end{bmatrix}$$

CPSC 340

5

SVD properties

1. There is no assumption that $m \geq n$ or that \mathbf{A} has full rank.
2. All diagonal elements of $\mathbf{\Sigma}$ are non-negative and in non-increasing order:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$$

where $p = \min(m, n)$

CPSC 340

6

SVD properties

Theorem 3 Every matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ has singular value decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$

Furthermore, the singular values $\{\sigma_j\}$ are uniquely determined.

If \mathbf{A} is square and $\sigma_i \neq \sigma_j$ for all $i \neq j$, the left singular vectors $\{\mathbf{u}_j\}$ and the right singular vectors $\{\mathbf{v}_j\}$ are uniquely determined to within a factor of ± 1 .

SVD in terms of eigenvalues

Theorem 4 The nonzero singular values of \mathbf{A} are the (positive) square roots of the nonzero eigenvalues of $\mathbf{A}^T\mathbf{A}$ or $\mathbf{A}\mathbf{A}^T$ (these matrices have the same nonzero eigenvalues).





Matrix norms

Theorem 5 $\|\mathbf{A}\|_2 = \sigma_1$, where $\|\mathbf{A}\|_2 = \max_{\mathbf{x} \neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|\mathbf{Ax}\|$.





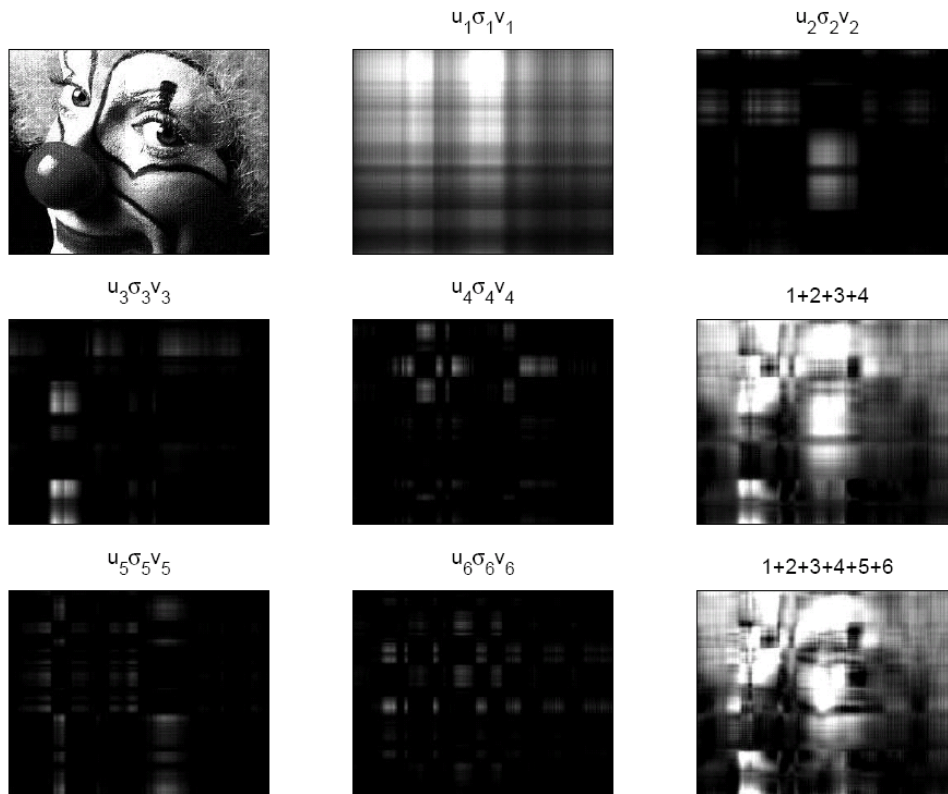
SVD in terms of rank 1 matrices

Theorem 6

$$\mathbf{A} = \sum_{j=1}^r \sigma_j \mathbf{u}_j \mathbf{v}_j^T$$

where r is the rank of \mathbf{A} .





Energy interpretation and approximation error

What is so useful about this expansion is that the ν^{th} partial sum captures as much of the “energy” of \mathbf{A} as possible by a matrix of at most rank- ν . In this case, “energy” is defined by the 2-norm.

Theorem 7 For any ν with $0 \leq \nu \leq r$ define

$$\mathbf{A}_\nu = \sum_{j=1}^{\nu} \sigma_j \mathbf{u}_j \mathbf{v}_j^T$$

If $\nu = p = \min(m, n)$, define $\sigma_{\nu+1} = 0$.

Then,

$$\|\mathbf{A} - \mathbf{A}_\nu\|_2 = \sigma_{\nu+1}$$