

# Singular Value Decomposition

L. Eldén

LiU

September 2020

# Singular Value Decomposition

## Theorem (SVD)

Any  $m \times n$  matrix  $A$ , with  $m \geq n$ , can be factorized

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T, \tag{1}\text{span}$$

where  $U \in \mathbb{R}^{m \times m}$  and  $V \in \mathbb{R}^{n \times n}$  are orthogonal, and  $\Sigma \in \mathbb{R}^{n \times n}$  is diagonal,

$$\begin{aligned} \Sigma &= \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n), \\ \sigma_1 &\geq \sigma_2 \geq \dots \geq \sigma_n \geq 0. \end{aligned}$$

# SVD

$$A = U \begin{pmatrix} 0 & \\ & 0 \end{pmatrix} V^T$$

Diagram illustrating the Singular Value Decomposition (SVD) of a matrix  $A$ . The matrix  $A$  is shown as a tall rectangle labeled  $A$ , with dimensions  $m \times n$  below it. An equals sign ( $=$ ) is positioned between  $A$  and the decomposition components. The matrix  $U$  is shown as a wide rectangle labeled  $U$ , with dimensions  $m \times m$  below it. The diagonal matrix is shown as a tall rectangle with two entries labeled  $0$ , with dimensions  $m \times n$  below it. The matrix  $V^T$  is shown as a square rectangle labeled  $V^T$ .

# Thin SVD

Partition  $U = (U_1 \ U_2)$ , where  $U_1 \in \mathbb{R}^{m \times n}$ ,

$$A = U_1 \Sigma V^T,$$

$$\begin{array}{c|c|c|c} A & = & U_1 & \Sigma \\ \hline m \times n & & m \times n & n \times n \end{array}$$

The diagram illustrates the thin Singular Value Decomposition (SVD) of matrix A. It shows four boxes: a tall rectangular box labeled 'A' representing the input matrix, followed by an equals sign, then a tall rectangular box labeled 'U1', then a square box containing a diagonal matrix labeled 'Σ' with zeros at the top-right and bottom-left, and finally a square box labeled 'V^T'.

# SVD Expansion

$$A = \sum_{i=1}^n \sigma_i u_i v_i^T = \left| \begin{array}{c} \text{---} \\ \text{---} \end{array} \right| + \left| \begin{array}{c} \text{---} \\ \text{---} \end{array} \right| + \dots$$

## SVD of a matrix with full column rank I

```
A = 1 1  
    1 2  
    1 3  
    1 4
```

```
>> [U,S,V]=svd(A)
```

```
U = 0.2195 -0.8073 0.0236 0.5472  
    0.3833 -0.3912 -0.4393 -0.7120  
    0.5472 0.0249 0.8079 -0.2176  
    0.7110 0.4410 -0.3921 0.3824
```

```
S = 5.7794 0  
    0 0.7738  
    0 0  
    0 0
```

# SVD of a matrix with full column rank II

$$V = \begin{pmatrix} 0.3220 & -0.9467 \\ 0.9467 & 0.3220 \end{pmatrix}$$

## Thin SVD

```
>> [U,S,V]=svd(A,0)
```

```
U = 0.2195 -0.8073  
     0.3833 -0.3912  
     0.5472  0.0249  
     0.7110  0.4410
```

```
S = 5.7794 0  
     0 0.7738
```

```
V = 0.3220 -0.9467  
     0.9467  0.3220
```

## Rank deficient matrix |

```
>> A(:,3)=A(:,1)+0.5*A(:,2)
```

```
A = 1.0000    1.0000    1.5000  
     1.0000    2.0000    2.0000  
     1.0000    3.0000    2.5000  
     1.0000    4.0000    3.0000
```

```
>> [U,S,V]=svd(A,0)
```

```
U = 0.2612    -0.7948    -0.5000  
     0.4032    -0.3708    0.8333  
     0.5451     0.0533   -0.1667  
     0.6871     0.4774   -0.1667
```

## Rank deficient matrix II

$$S = \begin{matrix} 7.3944 & 0 & 0 \\ 0 & 0.9072 & 0 \\ 0 & 0 & 0 \end{matrix}$$
$$V = \begin{matrix} 0.2565 & -0.6998 & 0.6667 \\ 0.7372 & 0.5877 & 0.3333 \\ 0.6251 & -0.4060 & -0.6667 \end{matrix}$$

# Fundamental Subspaces I

The *range of the matrix A*:

$$\mathcal{R}(A) = \{y \mid y = Ax, \text{ for arbitrary } x\}.$$

Assume that  $A$  has rank  $r$ :

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

Outer product form:

$$y = Ax = \sum_{i=1}^r \sigma_i u_i v_i^T x = \sum_{i=1}^r (\sigma_i v_i^T x) u_i = \sum_{i=1}^r \alpha_i u_i.$$

The *null-space of the matrix A*:

$$\mathcal{N}(A) = \{x \mid Ax = 0\}.$$

## Fundamental Subspaces II

$$Ax = \sum_{i=1}^r \sigma_i u_i v_i^T x$$

Any vector  $z = \sum_{i=r+1}^n \beta_i v_i$  is in the null-space:

$$Az = \left( \sum_{i=1}^r \sigma_i u_i v_i^T \right) \left( \sum_{i=r+1}^n \beta_i v_i \right) = 0.$$

# Fundamental Subspaces

## Theorem (Fundamental subspaces)

- ① *The singular vectors  $u_1, u_2, \dots, u_r$  are an orthonormal basis in  $\mathcal{R}(A)$  and*

$$\text{rank}(A) = \dim(\mathcal{R}(A)) = r.$$

- ② *The singular vectors  $v_{r+1}, v_{r+2}, \dots, v_n$  are an orthonormal basis in  $\mathcal{N}(A)$  and*

$$\dim(\mathcal{N}(A)) = n - r.$$

- ③ *The singular vectors  $v_1, v_2, \dots, v_r$  are an orthonormal basis in  $\mathcal{R}(A^T)$ .*
- ④ *The singular vectors  $u_{r+1}, u_{r+2}, \dots, u_m$  are an orthonormal basis in  $\mathcal{N}(A^T)$ .*

# Null Space

The third column of V is a basis vector in  $N(A)$ :

```
>> A*V(:,3)
```

```
ans =
1.0e-15 *
0
-0.2220
-0.2220
0
```

# Eckart-Young Property

## Frobenius norm

$$\|A\|_F = \left( \sum_{i,j} a_{ij}^2 \right)^{1/2}$$

## Theorem

Assume that the matrix  $A \in \mathbb{R}^{m \times n}$  has rank  $r > k$ . The Frobenius norm matrix approximation problem

$$\min_{\text{rank}(Z)=k} \|A - Z\|_F$$

has the solution

$$Z = A_k = U_k \Sigma_k V_k^T,$$

# Principal Component Analysis (PCA) I

Data matrix  $\mathbb{R}^{m \times n} \ni X = U\Sigma V^T$

Each column is an observation of a real-valued random vector with mean zero.

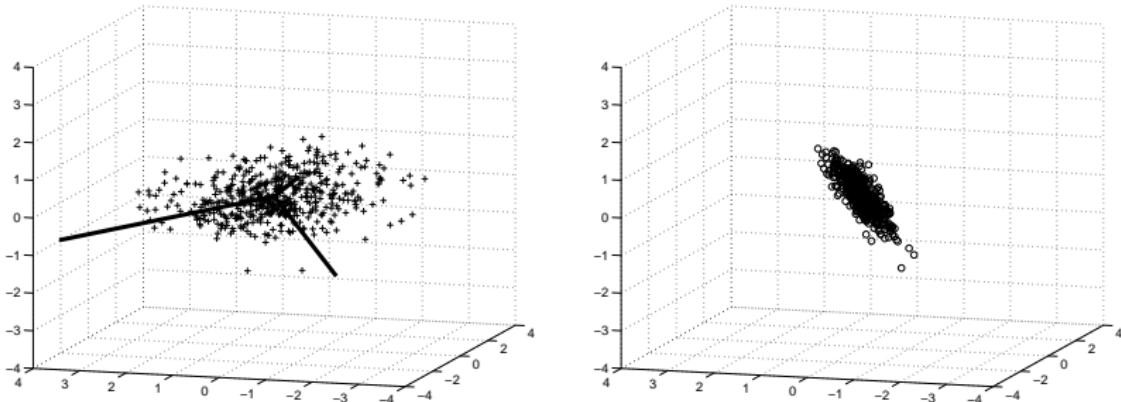
The right singular vectors  $v_i$  are called **principal components directions** of  $X$ . The vector

$$z_1 = Xv_1 = \sigma_1 u_1$$

has the largest sample variance amongst all normalized linear combinations of the columns of  $X$ :

$$\text{Var}(z_1) = \text{Var}(Xv_1) = \frac{\sigma_1^2}{m}.$$

The normalized variable  $u_1$  is called the **normalized first principal component** of  $X$ . The second principal component is the vector of largest sample variance of the deflated data matrix  $X - \sigma_1 u_1 v_1^T$ , and so on.



**Figure:** Cluster of points in  $\mathbb{R}^3$  with (scaled) principal components (top). The same data with the contributions along the first principal component deflated (bottom).

# Least Squares by SVD I

A =	1	1	b = 7.9700
	1	2	10.2000
	1	3	14.2000
	1	4	16.0000
	1	5	21.2000

```
>> [U1,S,V]=svd(A,0)
```

U1 =	0.1600	-0.7579
	0.2853	-0.4675
	0.4106	-0.1772
	0.5359	0.1131
	0.6612	0.4035

## Least Squares by SVD II

```
S = 7.6912          0  
      0      0.9194
```

```
V = 0.2669      -0.9637  
    0.9637      0.2669
```

```
>> x=V*(S\((U1'*b)))
```

```
x = 4.2360  
    3.2260
```

# Least Squares by SVD, in R |

```
> A.svd<-svd(A)
> A.svd
$d
[1] 7.6912131 0.9193696

$u
      [,1]      [,2]
[1,] 0.1600071  0.7578903
[2,] 0.2853078  0.4675462
[3,] 0.4106086  0.1772020
[4,] 0.5359094 -0.1131421
[5,] 0.6612102 -0.4034862

$v
      [,1]      [,2]
```

## Least Squares by SVD, in R II

```
[1,] 0.2669336 0.9637149  
[2,] 0.9637149 -0.2669336
```

```
> x=A.svd$v %*% diag(1/A.svd$d) %*% t(A.svd$u) %*% b
```

```
> x
```

```
[,1]
```

```
[1,] 4.236  
[2,] 3.226
```

```
>
```

# Linear dependence – SVD

## Theorem

*Let the singular values of  $A$  satisfy*

$$\sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0.$$

*Then the rank of  $A$  is equal to  $r$ .*

Rank = the number of linearly independent columns of  $A$ .

# Linear dependence I

```
A=[1 1; 1 2; 1 3; 1 4]
singval=svd(A)

% Third col=linear combination of first two
A1=[A A(:,1)+0.5*A(:,2)]
singval1=svd(A1)
```

```
A = 1 1
     1 2
     1 3
     1 4
```

Result:

## Linear dependence II

```
singval = 5.7794  
          0.7738
```

```
A1 = 1.0000    1.0000    1.5000  
     1.0000    2.0000    2.0000  
     1.0000    3.0000    2.5000  
     1.0000    4.0000    3.0000
```

```
singval1 = 7.3944  
           0.9072  
           0
```

# Almost linear dependence I

```
A2=[A A(:,1)+0.5*A(:,2)+0.0001*randn(4,1)]  
singval2=svd(A2)
```

---

```
A2 = 1.0000    1.0000    1.4999  
      1.0000    2.0000    2.0001  
      1.0000    3.0000    2.5000  
      1.0000    4.0000    3.0001
```

```
singval2 = 7.3944  
          0.9072  
          0.0001
```