Singular Value Decomposition

Shusen Wang

Orthonormal Basis

Definition (Orthonormal Basis).

A set of vector $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\} \subset \mathbb{R}^n$ $(k \leq n)$ forms an orthonormal basis (of some space or subspace) if:

- unit ℓ_2-norm: $||\mathbf{v}_i||_2 = 1$ for all i,
- orthogonal to each other: $\mathbf{v}_i^T \mathbf{v}_j = 0$ for all $i \neq j$.

Singular Value Decomposition (SVD)

- Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be any matrix.
- Rank: $r = \mathrm{rank}(\mathbf{A})$. $(r \leq m, n)$
- SVD: $\mathbf{A} = \sum_{i=1}^{r} \boxed{\sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^T}$

A rank-1 matrix

Singular Value Decomposition (SVD)

- Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be any matrix.
- Rank: $r = \mathrm{rank}(\mathbf{A})$. $(r \leq m, n)$
- SVD: $\mathbf{A} = \sum_{i=1}^{r} \sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^T$
    - Singular values: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.
    - Left singular vectors: $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_r\} \subset \mathbb{R}^m$ forms an orthonormal basis.
    - Right singular vectors: $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_r\} \subset \mathbb{R}^n$ forms an orthonormal basis.

Truncated SVD

- Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be any matrix.
- Rank: $r = \text{rank}(\mathbf{A})$. $(r \leq m, n)$
- SVD: $\mathbf{A} = \sum_{i=1}^{r} \sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^T$
- **Truncated SVD**: for $0 < k < r$, $\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^T$
  - $\mathbf{A}_k$ is the best rank-k approximation to $\mathbf{A}$.
  - $\mathbf{A}_k = \text{argmin}_{\mathbf{B}} \, ||\mathbf{A} - \mathbf{B}||_F^2 \, ; \quad s.t. \quad \text{rank}(\mathbf{B}) \leq k.$

Matrix Frobenius Norm

- Let $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times n}$ be any matrix.
- Frobenius norm: $||\mathbf{A}||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}$.
  - A generalization of the ℓ_2-vector norm to matrix.
- Property: $||\mathbf{A}||_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$.
  - σ_i is the i-th singular value of $\mathbf{A}$.

Error of Truncated SVD

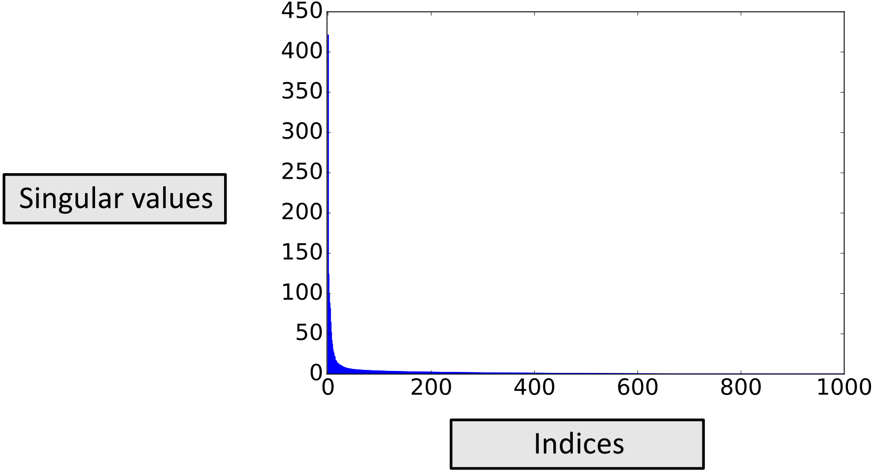
- SVD: $\mathbf{A} = \sum_{i=1}^{r} \sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^T$
- Truncated SVD: for $0 < k < r$, $\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i \, \mathbf{u}_i \, \mathbf{v}_i^T$.
- Error of the truncated SVD:

$$\left|\left|\mathbf{A} - \mathbf{A}_k\right|\right|_F^2 = \underbrace{\sum_{i=k+1}^{r} \sigma_i^2}_{\text{Bottom (the smallest) singular values}}.$$

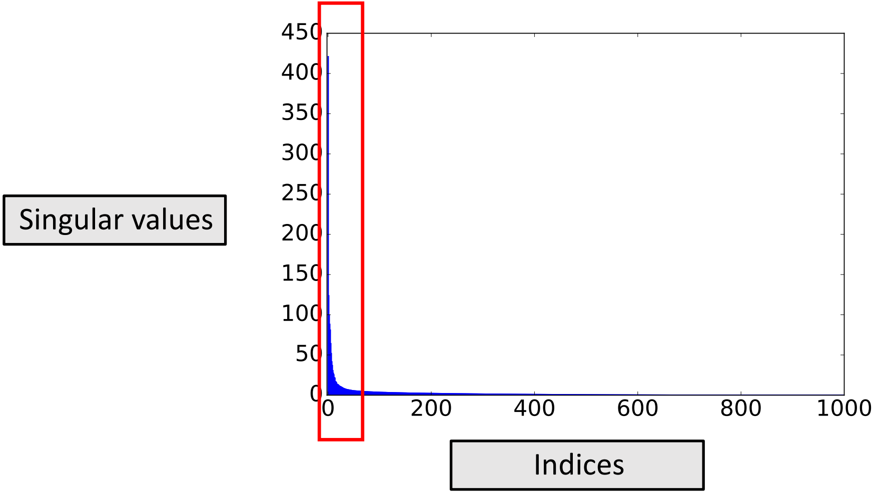
SVD: Example

Original image (1000×1500)

SVD: Example


Singular values
450
400
350
300
250
200
150
100
50
0
0
200
400
600
800
1000
Indices

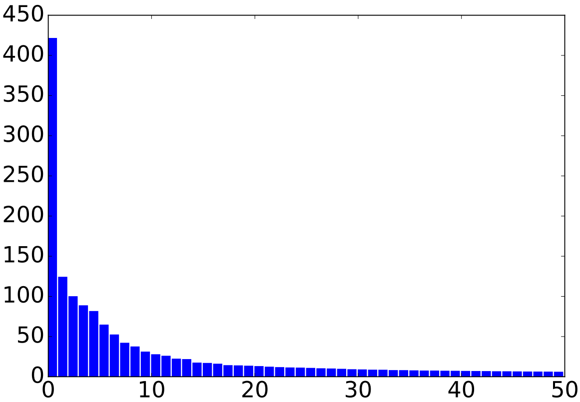
SVD: Example



450
400
350
300
250
200
150
100
50
0
0
200
400
600
800
1000
Singular values
Indices

SVD: Example

Singular values



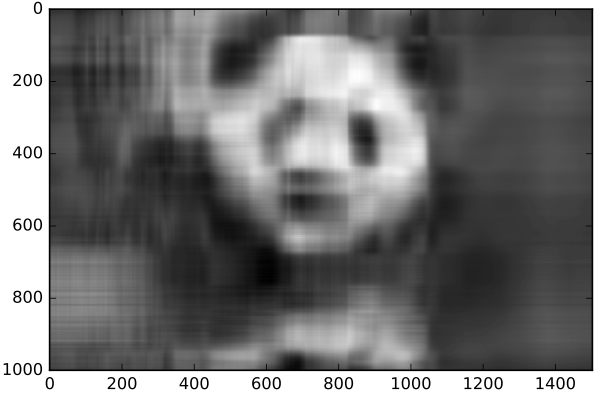
Indices (first 50)

SVD: Example

Original image (1000×1500)

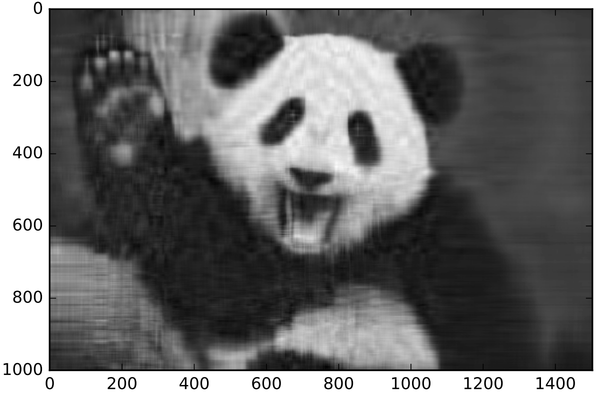
SVD: Example

Rank-5 truncated SVD



SVD: Example

Rank-20 truncated SVD



SVD: Example

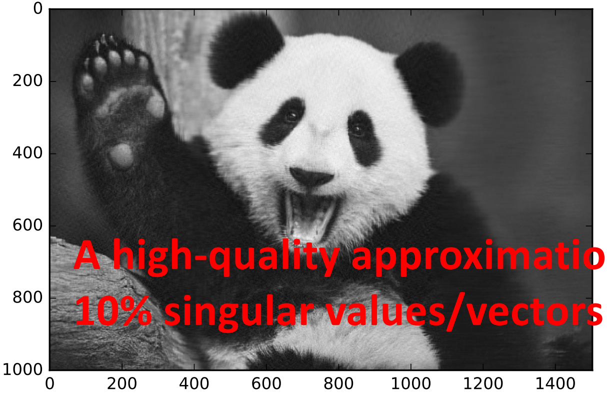
Rank-50 truncated SVD

SVD: Example

Rank-100 truncated SVD

SVD: Example

Rank-100 truncated SVD



A high-quality approximation using 10% singular values/vectors!

SVD: Example

- The original matrix
  - Shape: 1000×1500
  - #Entries: $1.5\mathrm{M}$
- The rank-100 truncated SVD
  - Shape: 100×1, 100×1500, and 1000×100
  - #Entries: $0.25\mathrm{M}$
- Truncated SVD saves 83% storage