- Review of matrix decompositions:
  - $\boldsymbol{L}\boldsymbol{U}=\boldsymbol{A}$ factorization (Gaussian elimination), used to solve linear systems (compute coordinates in new basis)
  - $\boldsymbol{Q}\boldsymbol{R}=\boldsymbol{A}$ factorization (Gram-Schmidt algorithm), used to solve least squares problems (compute best possible approximation)
  - $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{X}\boldsymbol{\Lambda}$, eigenproblem. If $\boldsymbol{X}$ nonsingular, eigendecomposition $\boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}=\boldsymbol{A}$ (reduction to diagonal form)
- Additional matrix decompositions:
  - $\boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^T=\boldsymbol{A}$, Schur decomposition (reduction to triangular form)
  - $\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}=\boldsymbol{A}$, Jordan decomposition (reduction to disjoint eigenspaces)
  - $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T=\boldsymbol{A}$, singular value decomposition (SVD, reduction to diagonal form, but with different bases in the domain, codomain)

Yet another factorization? Why?

- Eigendecomposition $\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1}$ is very useful in solving systems of ODEs, even more so when $\boldsymbol{A}$ is unitarily diagonalizable $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T$.
  - ODE system $\boldsymbol{u}' = \boldsymbol{A}\boldsymbol{u}$, $\boldsymbol{u}(0) = \boldsymbol{u}_0$ can be rewritten as $\boldsymbol{v}' = \boldsymbol{\Lambda}\boldsymbol{v}$, with $\boldsymbol{v} = \boldsymbol{Q}^T\boldsymbol{u}$
  - Solution of $\boldsymbol{v}' = \boldsymbol{\Lambda}\boldsymbol{v}$ is $\boldsymbol{v}(t) = e^{\boldsymbol{\Lambda}t}\boldsymbol{v}(0) = e^{\boldsymbol{\Lambda}t}\boldsymbol{Q}^T\boldsymbol{u}(0)$, hence $\boldsymbol{u}(t) = \boldsymbol{Q}e^{\boldsymbol{\Lambda}t}\boldsymbol{Q}^T\boldsymbol{u}(0)$
  - Since $e^{\boldsymbol{A}t} = \boldsymbol{I} + \boldsymbol{A}t + \frac{1}{2}\boldsymbol{A}^2t^2 + \ldots = \boldsymbol{Q}\left(\boldsymbol{I} + \boldsymbol{\Lambda}t + \frac{1}{2}\boldsymbol{\Lambda}^2t^2 + \ldots\right)\boldsymbol{Q}^T = \boldsymbol{Q}e^{\boldsymbol{\Lambda}t}\boldsymbol{Q}^T$

$$\boldsymbol{u}(t) = e^{\boldsymbol{A}t}\boldsymbol{u}(0)$$

- Above is an elegant procedure to solve ODEs, but only works if $\boldsymbol{A}$ is diagonalizable
- $\boldsymbol{A}$ can always be reduced to triangular form $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^T$, and introducing $\boldsymbol{w} = \boldsymbol{Q}^T\boldsymbol{u}$ leads to $\boldsymbol{w}' = \boldsymbol{T}\boldsymbol{w}$, but this triangular system is much more difficult to solve than $\boldsymbol{v}' = \boldsymbol{\Lambda}\boldsymbol{v}$
- A natural question arises: if $\boldsymbol{A}$ is not diagonalizable, how close can we get to a diagonal form?
- Answer: the Jordan decomposition $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$

- Any matrix $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ can be factorized as $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$, $\boldsymbol{P} \in \mathbb{R}^{m \times m}$ nonsingular,

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_p \end{pmatrix},$$

with p the number of distinct eigenvalues of $\boldsymbol{A}$, and $\boldsymbol{J}_k \in \mathbb{R}^{n_k \times n_k}$ given by

$$\boldsymbol{J}_k = \begin{pmatrix} \lambda_k & 1 & & \\ & \lambda_k & 1 & \\ & & \lambda_k & 1 \\ & & & \lambda_k \end{pmatrix},$$

with n_k the algebraic multiplicity of eigenvalue λ_k, $\sum_{k=1}^{p} n_k = m$.

- Singular value decomposition (SVD), for any $A \in \mathbb{R}^{m\times n}$, $A = U\Sigma V^T$, with $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ orthogonal, $\Sigma \in \mathbb{R}_+^{m\times n}$ diagonal
- The SVD is determined by eigendecomposition of A^TA, and AA^T
  - $A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V(\Sigma^T\Sigma)V^T$, an eigendecomposition of A^TA. The columns of V are eigenvectors of A^TA and called right singular vectors of A
  - $AA^T = (U\Sigma V^T)(U\Sigma^T V^T)^T = U(\Sigma\Sigma^T)U^T$, an eigendecomposition of A^TA. The columns of U are eigenvectors of AA^T and called left singular vectors of A
  - The matrix Σ has form

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \end{pmatrix} \in \mathbb{R}_+^{m\times n}$$

and σ_i are the singular values of A.

SVD properties, applications

- The singular value decomposition (SVD) furnishes complete information about A
  - $\mathrm{rank}(A) = r$ (the number of non-zero singular values)
  - U, V are orthogonal basis for the domain and codomain of A
- The SVD has numerous applications. As a representative example, principal component analysis is used to discover inherent natural descriptions of phenomena
- Let $X \in \mathbb{R}^{m \times n}$ represent n observations of a phenomenon characterized by m variables. Example: $n = 365$ daily measurements of temperatures at $m = 1000$ geographical locations
- Ask whether there are inherent patterns in the data by carrying out the SVD $X = U \Sigma V^T$
- The first few columns of U, V reflect dominant 'modes', or 'principal components' in the data

- Rewrite SVD as

$$A = U\Sigma V^T = \begin{pmatrix} U_1 & \dots & U_m \end{pmatrix} \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} V_1^T \\ \vdots \\ V_n^T \end{pmatrix} = \sum_{i=1}^{r} \sigma_i U_i V_i^T$$

- Each matrix $U_i V_i^T$ is a rank-one matrix
- The singular values are always given ordered $\sigma_1 \geqslant \sigma_2 \geqslant \dots \geqslant \sigma_r > 0$
- A reduced representation can be obtained by using fewer terms in the sum

$$A \cong \sum_{i=1}^{s} \sigma_i U_i V_i, s < r$$

Here's a sequence of images (each a matrix) that correspond various increasing values of s