- New concepts:
  - An orthogonal eigenvalue-revealing decomposition: Schur
  - Computability
  - Non-computability of polynomial roots
  - The need for an additional orthogonal decomposition
  - Interpreting the eigenvalue decomposition
  - Motivating the singular value decomposition
  - The singular value decomposition (SVD)

- Review of alread encountered 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)

Theorem. *(Schur) Any square matrix* $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ *can be decomposed as* $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^T$, *with* $\boldsymbol{T} \in \mathbb{R}^{m \times m}$ *upper triangular* ($t_{ij} = 0$ *for* $i > j$) *and* $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ *orthogonal* ($\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}$).

- The eigenvalues of $\boldsymbol{T}$ triangular are its diagonal elements
- $\boldsymbol{A}$ is similar to $\boldsymbol{T}$: $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \Rightarrow \boldsymbol{T}\boldsymbol{y} = \lambda \boldsymbol{y}$ and $\boldsymbol{y} = \boldsymbol{Q}^T \boldsymbol{y}$

Computability.

– roots of first degree polynomial: $p_1(x) = ax + b = 0 \Rightarrow x = -b/a \, (a \neq 0)$

– roots of second degree polynomia: $p_2(x) = ax^2 + bx + c \Rightarrow$

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

– roots of third degree polynomial (Cardano's formulas ~1520's)

– roots of fourth degree polynomial (Ferrari's formulas ~1540's, irrespective of Inquisitor Torquemada forbidding Valmes such knowledge)

– fifth degree polynomial: no formula possible (Galois, Abel Ruffini, 1820's)

- If $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ is normal ($\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A}$) then it has an orthogonal eigendecomposition $\exists \boldsymbol{Q}$, $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T, \boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$
- How does $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$ work?

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T\boldsymbol{x} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{w}$$

$$\boldsymbol{c} = \boldsymbol{Q}^T\boldsymbol{b} \Leftrightarrow \boldsymbol{Q}\boldsymbol{c} = \boldsymbol{I}\boldsymbol{b}, \boldsymbol{y} = \boldsymbol{Q}^T\boldsymbol{x} \Leftrightarrow \boldsymbol{Q}\boldsymbol{y} = \boldsymbol{I}\boldsymbol{x}$$

- In the $\boldsymbol{I} = [\ \boldsymbol{e}_1 \ \ \dots \ \ \boldsymbol{e}_m\ ]$ basis set $\boldsymbol{b} = x_1\boldsymbol{a}_1 + \cdots + x_m\boldsymbol{a}_m$ implies all components of $\boldsymbol{a}_1, \dots,$ $\boldsymbol{a}_m$ influence each component of $\boldsymbol{b}$

$$b_i = \sum_{j=1}^{m} a_{ij}x_j$$

- In the $\boldsymbol{Q}$ basis set $\boldsymbol{w} = \boldsymbol{\Lambda}\boldsymbol{y}$, component i of $\boldsymbol{w}$ influenced only by component i of $\boldsymbol{y}$

$$w_i = \lambda_i y_i$$

- For $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, $\boldsymbol{b}=\boldsymbol{A}\boldsymbol{x}$, a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ try to define:
  - an orthonormal basis $\boldsymbol{V}$ in $\mathbb{R}^n$, $\boldsymbol{V}\in\mathbb{R}^{n\times n}$, $\boldsymbol{V}\boldsymbol{V}^T=\boldsymbol{V}^T\boldsymbol{V}=\boldsymbol{I}_n$

$$\boldsymbol{I}\boldsymbol{x}=\boldsymbol{V}\boldsymbol{y}\Rightarrow\boldsymbol{y}=\boldsymbol{V}^T\boldsymbol{x}$$

  - an orthonormal basis $\boldsymbol{U}$ in $\mathbb{R}^m$, $\boldsymbol{U}\in\mathbb{R}^{m\times m}$, $\boldsymbol{U}\boldsymbol{U}^T=\boldsymbol{U}^T\boldsymbol{U}=\boldsymbol{I}_m$

$$\boldsymbol{I}\boldsymbol{b}=\boldsymbol{U}\boldsymbol{c}\Rightarrow\boldsymbol{c}=\boldsymbol{U}^T\boldsymbol{b}$$

  - impose that the action of $\boldsymbol{A}$ in the new bases is a simple component scaling

$$\boldsymbol{c}=\boldsymbol{\Sigma}\boldsymbol{y}\Rightarrow\boldsymbol{U}^T\boldsymbol{b}=\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{x}\Rightarrow\boldsymbol{b}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{x}\Rightarrow$$

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

  - Note that $\boldsymbol{\Sigma}\in\mathbb{R}^{m\times n}$

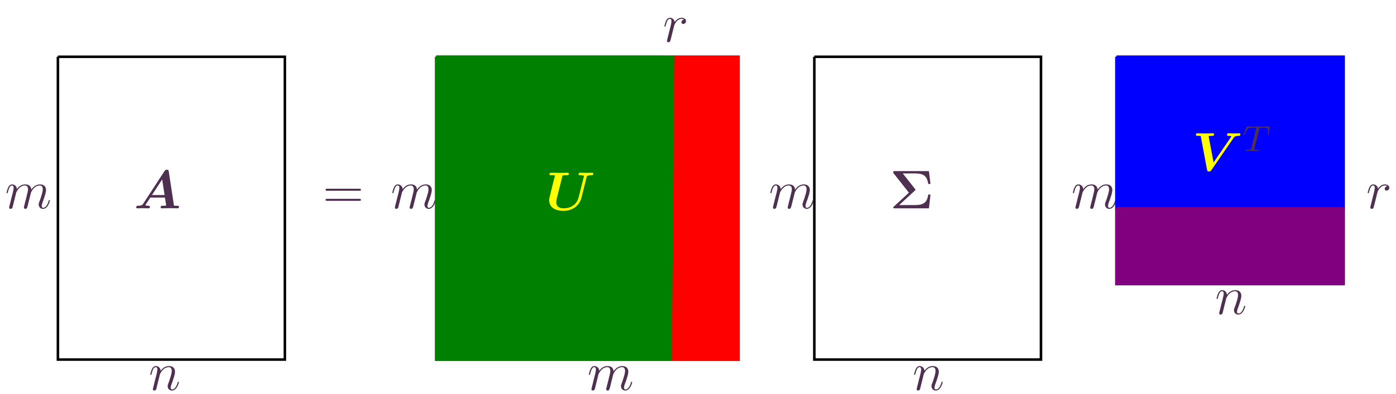
Theorem. *(SVD) For any* $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^T$, *with* $\boldsymbol{U} \in \mathbb{R}^{m \times m}$, $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ *orthogonal,* $\boldsymbol{\Sigma} \in \mathbb{R}_+^{m \times n}$ *pseudo-diagonal* $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \ldots, \sigma_r, \ldots, 0)$, $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$, $r \leqslant \min(m, n)$. $r = \operatorname{rank}(\boldsymbol{A})$.

The SVD is determined by eigendecomposition of $\boldsymbol{A}^T \boldsymbol{A}$, and $\boldsymbol{A} \boldsymbol{A}^T$

- $\boldsymbol{A}^T \boldsymbol{A} = (\boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^T)^T (\boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^T) = \boldsymbol{V} (\boldsymbol{\Sigma}^T \boldsymbol{\Sigma}) \boldsymbol{V}^T$, an eigendecomposition of $\boldsymbol{A}^T \boldsymbol{A}$. The columns of $\boldsymbol{V}$ are eigenvectors of $\boldsymbol{A}^T \boldsymbol{A}$ and called *right singular vectors* of $\boldsymbol{A}$
- $\boldsymbol{A} \boldsymbol{A}^T = (\boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^T)(\boldsymbol{U} \boldsymbol{\Sigma}^T \boldsymbol{V}^T)^T = \boldsymbol{U} (\boldsymbol{\Sigma} \boldsymbol{\Sigma}^T) \boldsymbol{U}^T$, an eigendecomposition of $\boldsymbol{A} \boldsymbol{A}^T$. The columns of $\boldsymbol{U}$ are eigenvectors of $\boldsymbol{A} \boldsymbol{A}^T$ and called *left singular vectors* of $\boldsymbol{A}$
- The matrix $\boldsymbol{\Sigma}$ has zero elements except for the diagonal that contains σ_i, the *singular values* of $\boldsymbol{A}$, computed as the square roots of the eigenvalues of $\boldsymbol{A}^T \boldsymbol{A}$ (or $\boldsymbol{A} \boldsymbol{A}^T$)

The theorem also holds for complex matrices with transposition replaced by taking the adjoint, $\boldsymbol{A} \in \mathbb{C}^{m \times n}$, $\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^*$, with $\boldsymbol{U} \in \mathbb{C}^{m \times m}$, $\mathbb{C} \in \mathbb{R}^{n \times n}$ unitary.

- SVD of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ reveals: $\operatorname{rank}(\boldsymbol{A})$, bases for $C(\boldsymbol{A}), N(\boldsymbol{A}^T), C(\boldsymbol{A}^T), N(\boldsymbol{A})$



$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_r & \boldsymbol{u}_{r+1} & \dots & \boldsymbol{u}_m \end{pmatrix} \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^T \\ \vdots \\ \boldsymbol{v}_r^T \\ \boldsymbol{v}_{r+1}^T \\ \vdots \\ \boldsymbol{v}_n^T \end{pmatrix}$$