- What is linear algebra and why is to so important to so many applications?
- Basic operations: defined to express linear combination
- Linear operators, Fundamental Theorem of Linear Algebra (FTLA)
- Factorizations: more convenient expression of linear combination

$$\boldsymbol{L}\boldsymbol{U} = \boldsymbol{A}, \boldsymbol{Q}\boldsymbol{R} = \boldsymbol{A}, \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1} = \boldsymbol{A}, \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{A}$$

- Solving linear systems (change of basis) $\boldsymbol{I}\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}$
- Best 2-norm approximation: least squares $\min_{\boldsymbol{x}} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2$
- Exposing the structure of a linear operator between the same sets through eigendecomposition
- Exposing the structure of a linear operator between different sets through the SVD
- Applications: any type of correlation

- Science acquires and organizes knowledge into theories that can be verified by *quantified tests*. Mathematics furnishes the appropriate context through rigorous definition of $\mathbb{N}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$.
- Most areas of science require groups of numbers to describe an observation. To organize knowledge rules on how such groups of numbers may be combined are needed. Mathematics furnishes the concept of a *vector space* $(\mathcal{S}, \mathcal{V}, +)$
  - i formal definition of a single number: scalar, $\alpha, \beta \in \mathcal{S}$
  - ii formal definition of a group of numbers: vector, $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$
  - iii formal definition of a possible way to combine vectors: $\alpha \boldsymbol{u} + \beta \boldsymbol{v}$
- Algebra is concerned with precise definition of ways to combine mathematical objects, i.e., to organize more complex knowledge as a sequence of operations on simpler objects
- Linear algebra concentrates on one particular operation: the *linear combination* $\alpha \boldsymbol{u} + \beta \boldsymbol{v}$
- It turns out that a complete theory can be built around the linear combination, and this leads to the many applications linear algebra finds in all branches of knowledge.

- Group vectors as column vectors into matrices $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{pmatrix} \in \mathbb{R}^{m \times n}$
- Define matrix-vector multiplication to express the basic linear combination operation

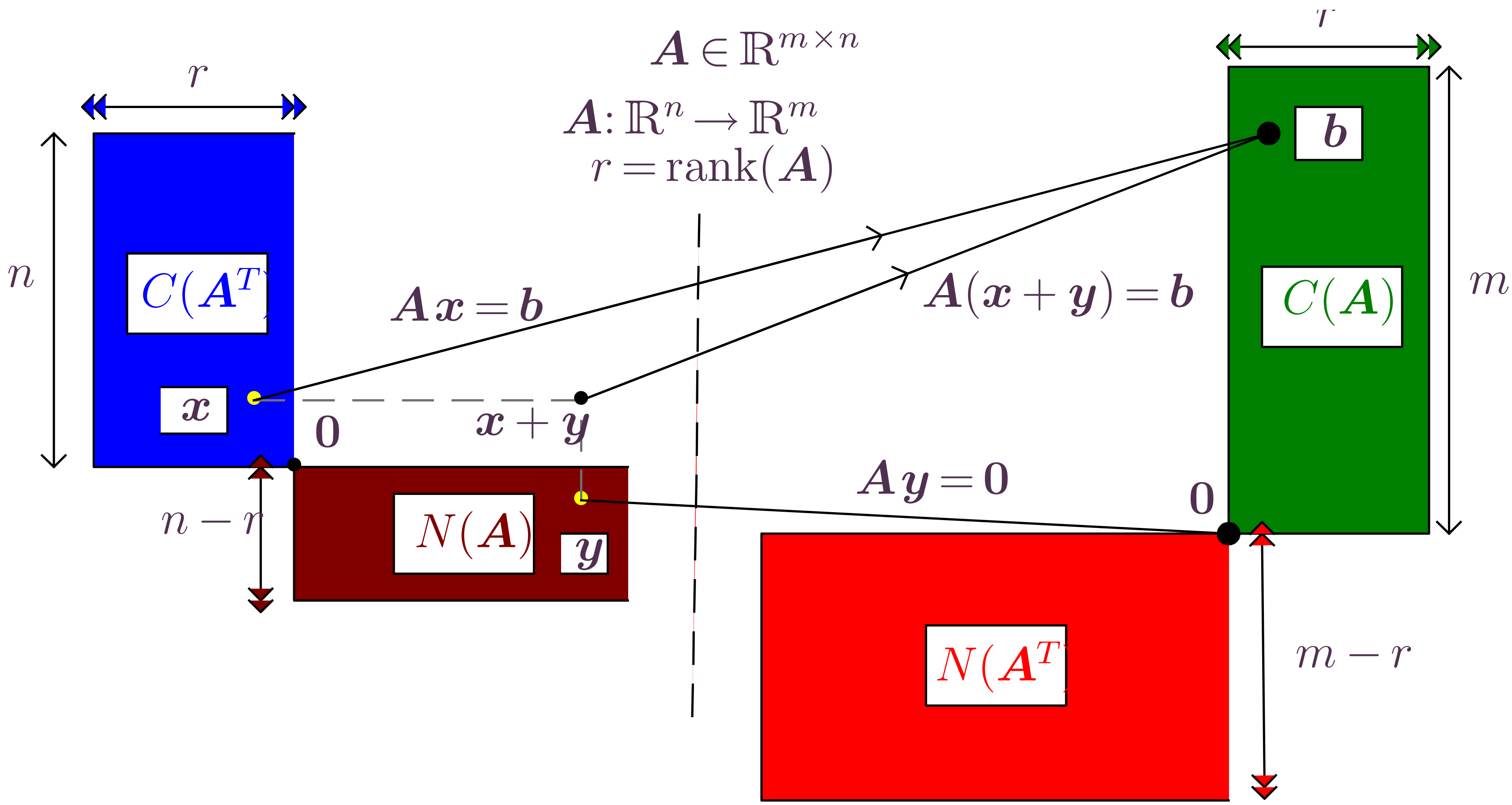
$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} = x_1 \boldsymbol{a}_1 + \ldots + x_n \boldsymbol{a}_n$$

- Introduce a way to switch between column and row storage through the transposition operation $\boldsymbol{A}^T$. $(\boldsymbol{A} + \boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$, $(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T \boldsymbol{A}^T$
- Transform between one set of basis vectors and another $\boldsymbol{b}\boldsymbol{I} = \boldsymbol{A}\boldsymbol{x}$
- *Linear independence* establishes when a vector cannot be described as a linear combination of other vectors, i.e., if *the only way* to satisfy $x_1 \boldsymbol{a}_1 + \ldots + x_n \boldsymbol{a}_n = \boldsymbol{0}$ is for $x_1 = \ldots = x_n = 0$, then the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are linearly independent
- The *span* $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle = \{\boldsymbol{b} \mid \exists \boldsymbol{x} \in \mathbb{R}^n \, \text{such that} \, \boldsymbol{b} = x_1 \boldsymbol{a}_1 + \ldots x_n \boldsymbol{a}_n\}$ is the set of all vectors is reachable by linear combination of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$
- The set of vectors $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is a *basis* of a vector space $\mathcal{V}$ if $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle = \mathcal{V}$, and $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are linearly independent
- The number of vectors in a basis is the *dimension* of a vector space.

- Any linear operator $T\colon \mathcal{D} \to \mathcal{C}$, $T(\alpha \boldsymbol{u} + \beta \boldsymbol{v}) = \alpha T(\boldsymbol{u}) + \beta T(\boldsymbol{v})$ can be characterized by a matrix $\boldsymbol{A} = [\ T(\boldsymbol{e}_1)\ \ T(\boldsymbol{e}_2)\ \ \ldots\ \ T(\boldsymbol{e}_n)\ ]$
- For each matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ there exist four fundamental subspaces:
  1. *Column space*, $C(\boldsymbol{A}) = \{\boldsymbol{b} \in \mathbb{R}^m | \exists \boldsymbol{x} \in \mathbb{R}^n \text{ such that } \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x}\} \subseteq \mathbb{R}^m$, the part of $\mathbb{R}^m$ *reachable* by linear combination of columns of $\boldsymbol{A}$
  2. *Left null space*, $N(\boldsymbol{A}^T) = \{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{A}^T \boldsymbol{y} = 0\} \subseteq \mathbb{R}^m$, the part of $\mathbb{R}^m$ *not reachable* by linear combination of columns of $\boldsymbol{A}$
  3. *Row space*, $R(\boldsymbol{A}) = C(\boldsymbol{A}^T) = \{\boldsymbol{c} \in \mathbb{R}^n | \exists \boldsymbol{y} \in \mathbb{R}^m \text{ such that } \boldsymbol{c} = \boldsymbol{A}^T \boldsymbol{y}\} \subseteq \mathbb{R}^n$, the part of $\mathbb{R}^m$ *reachable* by linear combination of rows of $\boldsymbol{A}$
  4. *Null space*, $N(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{A}\boldsymbol{x} = 0\} \subseteq \mathbb{R}^n$, the part of $\mathbb{R}^m$ *not reachable* by linear combination of rows of $\boldsymbol{A}$

  The fundamental theorem of linear algebra (FTLA) states

$$C(\boldsymbol{A}), N(\boldsymbol{A}^T) \leq \mathbb{R}^m, C(\boldsymbol{A}) \perp N(\boldsymbol{A}^T),\ \ C(\boldsymbol{A}) \cap N(\boldsymbol{A}^T) = \{\boldsymbol{0}\},\ \ C(\boldsymbol{A}) \oplus N(\boldsymbol{A}^T) = \mathbb{R}^m$$
$$C(\boldsymbol{A}^T), N(\boldsymbol{A}) \leq \mathbb{R}^n, C(\boldsymbol{A}^T) \perp N(\boldsymbol{A}),\ \ C(\boldsymbol{A}^T) \cap N(\boldsymbol{A}) = \{\boldsymbol{0}\},\ \ C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A}) = \mathbb{R}^n$$



$\mathbb{R}^n = C(\boldsymbol{A}^T) \oplus N(\boldsymbol{A})$ usually: $m \geqslant n$ $\mathbb{R}^m = N(\boldsymbol{A}^T) \oplus C(\boldsymbol{A})$

$C(\boldsymbol{A}^T) \perp N(\boldsymbol{A})$ $N(\boldsymbol{A}^T) \perp C(\boldsymbol{A})$

- $\boldsymbol{L}\boldsymbol{U} = \boldsymbol{A}$, (or $\boldsymbol{L}\boldsymbol{U} = \boldsymbol{P}\boldsymbol{A}$ with $\boldsymbol{P}$ a permutation matrix) Gaussian elimination, solving linear systems. Given $\boldsymbol{A} \in \mathbb{R}^{m\times m}, \boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{b} \in C(\boldsymbol{A})$, find $\boldsymbol{x} \in \mathbb{R}^m$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} = \boldsymbol{I}\boldsymbol{b}$ by:
  1. Factorize, $\boldsymbol{L}\boldsymbol{U} = \boldsymbol{P}\boldsymbol{A}$
  2. Solve lower triangular system $\boldsymbol{L}\boldsymbol{y} = \boldsymbol{P}\boldsymbol{b}$ by forward substitution
  3. Solve upper triangular system $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}$ by backward substitution
- $\boldsymbol{Q}\boldsymbol{R} = \boldsymbol{A}$, (or $\boldsymbol{Q}\boldsymbol{R} = \boldsymbol{P}\boldsymbol{A}$ with $\boldsymbol{P}$ a permutation matrix) Gram-Schmidt, solving least squares problem. Given $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $n \leqslant m$, solve $\min_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|$ by:
  1. Factorize, $\boldsymbol{Q}\boldsymbol{R} = \boldsymbol{P}\boldsymbol{A}$
  2. Solve upper triangular system $\boldsymbol{R}\boldsymbol{x} = \boldsymbol{Q}^T\boldsymbol{b}$ by forward substitution
- $\boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1} = \boldsymbol{A}$, eigendecomposition of $\boldsymbol{A} \in \mathbb{R}^{m\times m}$ ($\boldsymbol{X}$ invertible if $\boldsymbol{A}$ is normal, i.e., $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A}$)
- $\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T = \boldsymbol{A}$, orthogonal eigendecomposition when $\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{A}^T\boldsymbol{A}$ (normal)
- $\boldsymbol{Q}\boldsymbol{T}\boldsymbol{Q}^T = \boldsymbol{A}$, Schur decomposition of $\boldsymbol{A} \in \mathbb{R}^{m\times m}$, $\boldsymbol{Q}$ orthogonal matrix, $\boldsymbol{T}$ triangular matrix, decomposition always exists
- $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{A}$, Singular value decomposition of $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{U} \in \mathbb{R}^{m\times m}$, $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ orthogonal matrices, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots) \in \mathbb{R}_+^{m\times n}$, decomposition always exists

- Gaussian elimination produces a sequence matrices similar to $\boldsymbol{A}\in\mathbb{R}^{m\times m}$

$$\boldsymbol{A}=\boldsymbol{A}^{(0)}\sim\boldsymbol{A}^{(1)}\sim\dots\sim\boldsymbol{A}^{(k)}\sim\dots\sim\boldsymbol{A}^{(m-1)}$$

- Step k produces zeros underneath diagonal position (k,k)
- Step k can be represented as multiplication by matrix

$$\boldsymbol{A}^{(k)}=\boldsymbol{L}_k\boldsymbol{A}^{(k-1)},\boldsymbol{L}_k=\begin{pmatrix} 1 & \dots & 0 & \dots & 1\\ 0 & \ddots & 0 & \dots & 0\\ 0 & \dots & 1 & \dots & 0\\ 0 & \dots & -l_{k+1,k} & \dots & 0\\ \vdots & \dots & \vdots & \ddots & \vdots\\ 0 & \dots & -l_{m,k} & \dots & 1 \end{pmatrix}, l_{j,k}=\frac{a_{j,k}^{(k-1)}}{a_{k,k}^{(k-1)}}, \boldsymbol{A}^{(k)}=[a_{i,j}^{(k)}]$$

- All $m-1$ steps produce an upper triangular matrix

$$\boldsymbol{L}_{m-1}\dots\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A}=\boldsymbol{U}\Rightarrow\boldsymbol{A}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\dots\boldsymbol{L}_{m-1}^{-1}\boldsymbol{U}=\boldsymbol{L}\boldsymbol{U}$$

- With permutations $\boldsymbol{P}\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$ (Matlab `[L,U,P]=lu(A), A=P'*L*U`)

- With known LU-factorization: $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Rightarrow (\boldsymbol{L}\boldsymbol{U})\boldsymbol{x} = \boldsymbol{P}\boldsymbol{b} \Rightarrow \boldsymbol{L}(\boldsymbol{U}\boldsymbol{x}) = \boldsymbol{P}\boldsymbol{b}$
- To solve $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$:
  1. Carry out LU-factorization: $\boldsymbol{P}^T\boldsymbol{L}\boldsymbol{U} = \boldsymbol{A}$
  2. Solve $\boldsymbol{L}\boldsymbol{y} = \boldsymbol{c} = \boldsymbol{P}\boldsymbol{b}$ by forward substitution to find $\boldsymbol{y}$
  3. Solve $\boldsymbol{U}\boldsymbol{x} = \boldsymbol{y}$ by backward substitution
- FLOP = floating point operation = one multiplication and one addition
- Operation counts: how many FLOPS in each step?
  1. Each $\boldsymbol{L}_k\boldsymbol{A}^{(k-1)}$ costs $(m-k)^2$ FLOPS. Overall

$$(m-1)^2 + (m-2)^2 + \cdots + 1^2 = \frac{m(m-1)(2m-1)}{6} \approx \frac{m^3}{3}$$

  2. Forward substitution step k costs k flops

$$1 + 2 + \cdots + m = \frac{m(m+1)}{2} \approx \frac{m^2}{2}$$

  3. Backward substitution cost is identical $m(m+1)/2 \approx m^2/2$/

- Orthonormalization of columns of $\boldsymbol{A}$ is also a factorization

$$\boldsymbol{A} = [\ \boldsymbol{a}_1\ \ \boldsymbol{a}_2\ \ \ldots\ \ \boldsymbol{a}_n\ ] = [\ \boldsymbol{q}_1\ \ \boldsymbol{q}_2\ \ \ldots\ \ \boldsymbol{q}_n\ ] \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ 0 & r_{22} & \ldots & r_{2n} \\ \vdots & & \ddots & \\ 0 & 0 & \ldots & r_{nn} \end{bmatrix} = \boldsymbol{Q}\boldsymbol{R}$$

$$\begin{aligned} \boldsymbol{a}_1 &= r_{11}\,\boldsymbol{q}_1 \\ \boldsymbol{a}_2 &= r_{12}\,\boldsymbol{q}_1 + r_{22}\,\boldsymbol{q}_2 \\ \boldsymbol{a}_3 &= r_{13}\,\boldsymbol{q}_1 + r_{23}\,\boldsymbol{q}_2 + r_{33}\,\boldsymbol{q}_3 \\ &\vdots \\ \boldsymbol{a}_n &= r_{1n}\,\boldsymbol{q}_1 + r_{2n}\,\boldsymbol{q}_2 + r_{3n}\,\boldsymbol{q}_3 + \cdots + r_{nn}\boldsymbol{q}_n \end{aligned}$$

$$\begin{aligned} \boldsymbol{q}_1 &= \boldsymbol{a}_1 / r_{11} \\ \boldsymbol{q}_2 &= (\boldsymbol{a}_2 - r_{12}\,\boldsymbol{q}_1) / r_{22} \\ \boldsymbol{q}_3 &= (\boldsymbol{a}_3 - r_{13}\,\boldsymbol{q}_1 - r_{23}\,\boldsymbol{q}_2) / r_{33} \\ &\vdots \end{aligned}$$

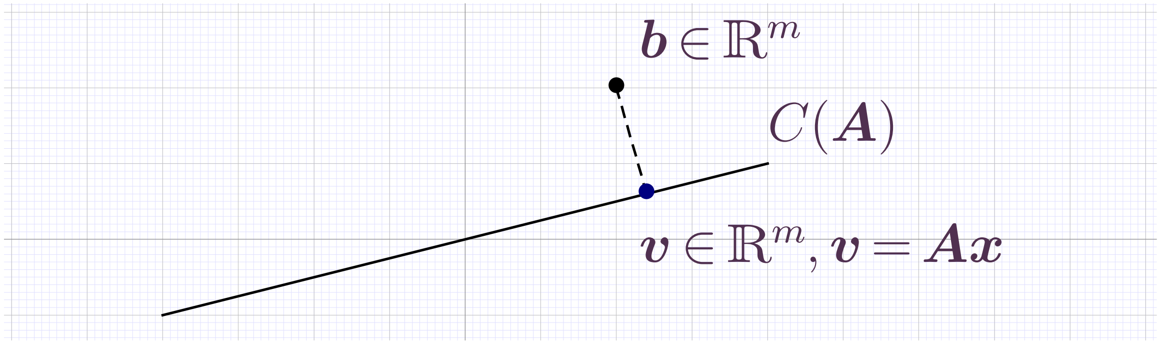
- Operation count:
  - $r_{jk} = \boldsymbol{q}_j^T \boldsymbol{a}_k$ costs m FLOPS
  - There are $1 + 2 + \cdots + n$ components in $\boldsymbol{R}$, Overall cost $n(n+1)m/2$
- With permutations $\boldsymbol{A}\boldsymbol{P} = \boldsymbol{Q}\boldsymbol{R}$ (Matlab `[Q,R,P]=qr(A)`)

- With known QR-factorization: $\boldsymbol{Ax}=\boldsymbol{b}\Rightarrow(\boldsymbol{QRP}^T)\boldsymbol{x}=\boldsymbol{b}\Rightarrow\boldsymbol{Ry}=\boldsymbol{Q}^T\boldsymbol{b}$
- To solve $\boldsymbol{Ax}=\boldsymbol{b}$:
  1. Carry out QR-factorization: $\boldsymbol{QRP}^T=\boldsymbol{A}$
  2. Compute $\boldsymbol{c}=\boldsymbol{Q}^T\boldsymbol{b}$
  3. Solve $\boldsymbol{Ry}=\boldsymbol{c}$ by backward substitution
  4. Find $\boldsymbol{x}=\boldsymbol{P}^T\boldsymbol{y}$
- Operation counts: how many FLOPS in each step?
  1. QR-factorization $m^2(m+1)/2\approx m^3/2$
  2. Compute $\boldsymbol{c}$, m^2
  3. Backward substitution $m(m+1)/2\approx m^2/2$

- Consider approximating $\boldsymbol{b} \in \mathbb{R}^m$ by linear combination of n vectors, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$
- Make approximation error $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{v} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ as small as possible

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2$$

Error is measured in the 2-norm $\Rightarrow$ the *least squares problem* (LSQ)



- Solution is the projection of $\boldsymbol{b}$ onto $C(\boldsymbol{A})$

$$\boldsymbol{Q}\boldsymbol{R} = \boldsymbol{A}, \boldsymbol{P}_{C(\boldsymbol{A})} = \boldsymbol{Q}\boldsymbol{Q}^T, \boldsymbol{v} = (\boldsymbol{Q}\boldsymbol{Q}^T)\boldsymbol{b}$$

- The vector $\boldsymbol{x}$ is found by back-substitution from

$$\boldsymbol{v} = (\boldsymbol{Q}\boldsymbol{Q}^T)\boldsymbol{b} = (\boldsymbol{Q}\boldsymbol{R})\boldsymbol{x} \Rightarrow \boldsymbol{R}\boldsymbol{x} = \boldsymbol{Q}^T\boldsymbol{b}.$$

- For square matrix $\boldsymbol{A} \in \mathbb{R}^{m \times m}$ find *non-zero* vectors whose *directions* are not changed by multiplication by $\boldsymbol{A}$, $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$, λ is scalar, the *eigenvalue problem*.
- Consider the eigenproblem $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$ for $\boldsymbol{A} \in \mathbb{R}^{m \times m}$. Rewrite as

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \Rightarrow (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}.$$

  Since $\boldsymbol{x} \neq \boldsymbol{0}$, a solution to eigenproblem exists only if $\boldsymbol{A} - \lambda \boldsymbol{I}$ is singular.
- $\boldsymbol{A} - \lambda \boldsymbol{I}$ singular implies $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$, $\boldsymbol{x} \in N(\boldsymbol{A} - \lambda \boldsymbol{I})$
- Investigate form of $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & a_{2m} \\ a_{31} & a_{32} & a_{33} - \lambda & \dots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mm} - \lambda \end{vmatrix}$$

- $p_m(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A})$, an m^{th}-degree polynomial in λ, *characteristic polynomial* of $\boldsymbol{A}$, with m roots, $\lambda_1, \lambda_2, \dots, \lambda_m$, the eigenvalues of $\boldsymbol{A}$

- $\boldsymbol{A} \in \mathbb{R}^{m \times m}$, eigenvalue problem $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$ ($\boldsymbol{x} \neq \boldsymbol{0}$) in matrix form:

$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{\Lambda}$$

$$\boldsymbol{X} = [\, \boldsymbol{x}_1 \ \ \ldots \ \ \boldsymbol{x}_m \,], \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_m) = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_m \end{bmatrix}.$$

- $\boldsymbol{X}$ is the *eigenvector matrix*, $\boldsymbol{\Lambda}$ is the (diagonal) *eigenvalue matrix*
- If column vectors of $\boldsymbol{X}$ are linearly independent, then $\boldsymbol{X}$ is invertible

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{\Lambda}\boldsymbol{X}^{-1},$$

  the *eigendecomposition* of $\boldsymbol{A}$ (compare to $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{U}$, $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$)
- Rule "determinant of product = product of determinants" implies

$$\det(\boldsymbol{A}\boldsymbol{X}) = \det(\boldsymbol{X}\boldsymbol{\Lambda}) \Rightarrow \det(\boldsymbol{A}) = \det(\boldsymbol{\Lambda}) \ (\text{for} \det(\boldsymbol{X}) \neq 0).$$

Definition 1. *The algebraic multiplicity of an eigenvalue λ is the number of times it appears as a repeated root of the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$*

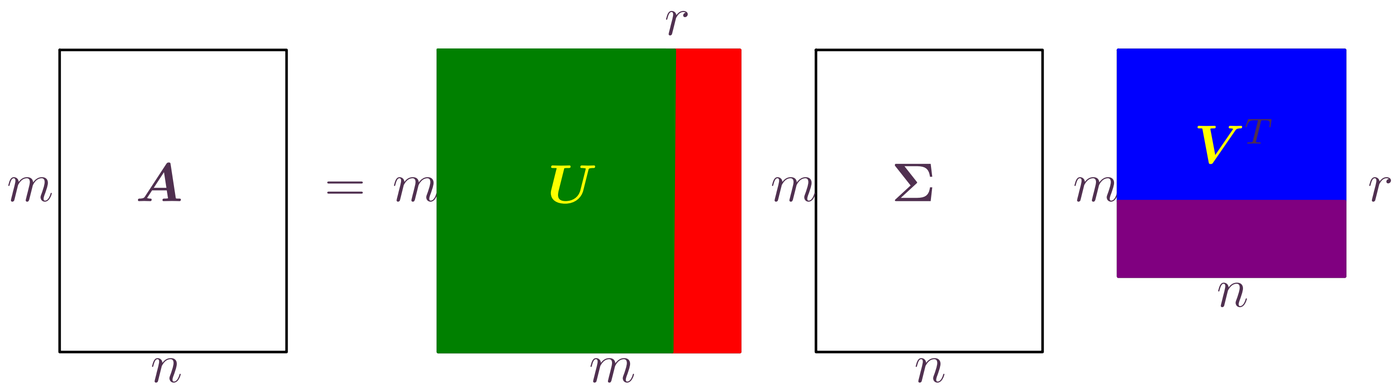
Example. $p(\lambda) = \lambda(\lambda - 1)(\lambda - 2)^2$ has two single roots $\lambda_1 = 0$, $\lambda_2 = 1$ and a repeated root $\lambda_{3,4} = 2$. The eigenvalue $\lambda = 2$ has an algebraic multiplicity of 2

Definition 2. *The geometric multiplicity of an eigenvalue λ is the dimension of the null space of $A - \lambda I$*

Definition 3. *An eigenvalue for which the geometric multiplicity is less than the algebraic multiplicity is said to be defective*

Theorem. *A matrix is diagonalizable if the geometric multiplicity of each eigenvalue is equal to the algebraic multiplicity of that eigenvalue.*

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



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

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

$$\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\boldsymbol{U}^T = \boldsymbol{U}\boldsymbol{\Lambda}_m\boldsymbol{U}^T \Rightarrow \boldsymbol{B}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Lambda}_m$$

  Singular vectors $\boldsymbol{U}$ are eigenvectors of $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^T$

$$\boldsymbol{C} = \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{V}\boldsymbol{\Lambda}_n\boldsymbol{V}^T \Rightarrow \boldsymbol{C}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}_n$$

  Singular vectors $\boldsymbol{V}$ are eigenvectors of $\boldsymbol{C} = \boldsymbol{A}^T\boldsymbol{A}$
- SVD relevant for correlations: N measurements of $\boldsymbol{x}(t) \in \mathbb{R}^n$ at $t_1, \ldots, t_N$

$$\boldsymbol{X} = [\ \boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{x}_n\ ] \in \mathbb{R}^{N \times n}$$

  Choose origin such that $E[\boldsymbol{x}] = \mathbf{0}$, construct *covariance matrix*

$$\boldsymbol{C}_{\boldsymbol{X}} = \boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_n^T \end{bmatrix} [\ \boldsymbol{x}_1\ \ \boldsymbol{x}_2\ \ \ldots\ \ \boldsymbol{x}_n\ ] = \begin{bmatrix} \boldsymbol{x}_1^T\boldsymbol{x}_1 & \boldsymbol{x}_1^T\boldsymbol{x}_2 & \ldots & \boldsymbol{x}_1^T\boldsymbol{x}_n \\ \boldsymbol{x}_2^T\boldsymbol{x}_1 & \boldsymbol{x}_2^T\boldsymbol{x}_2 & \ldots & \boldsymbol{x}_2^T\boldsymbol{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{x}_n^T\boldsymbol{x}_1 & \boldsymbol{x}_n^T\boldsymbol{x}_2 & \ldots & \boldsymbol{x}_n^T\boldsymbol{x}_n \end{bmatrix}$$

  Eigenvectors of $\boldsymbol{C}$ are singular vectors $\boldsymbol{V}$ of $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T \Rightarrow$ image compression, graph partition, social networks, data analysis