Numerical Methods of Linear Algebra for Sparse Matrices

Lecture 3

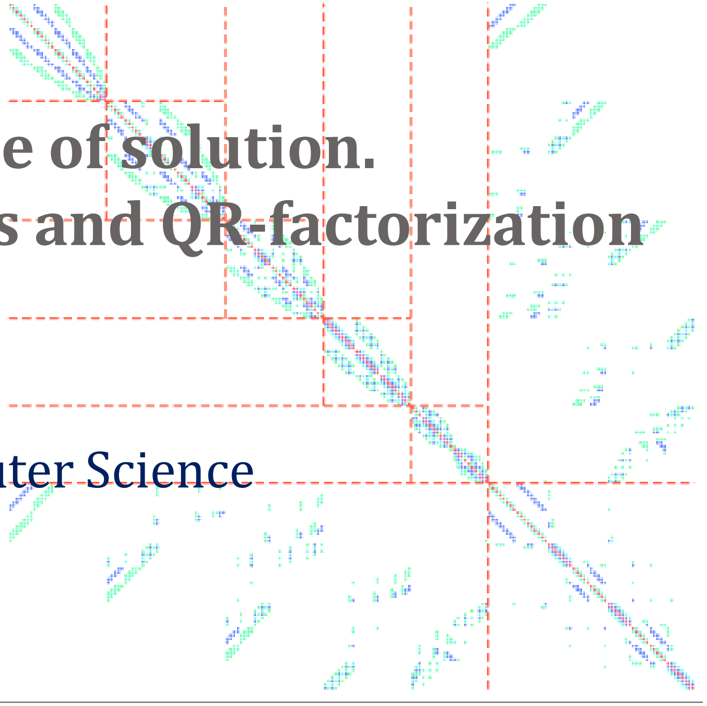
Subspace, range and kernel. Existence of solution. Orthogonality, Gram-Schmidt process and QR-factorization

Anna Nasedkina, PhD, Assoc. prof.

Department of Mathematical Modeling

Institute of Mathematics, Mechanics and Computer Science

Southern Federal University



Outline

- Subspace, range and kernel.
- Existence of solution.
- Orthogonality of vectors and subspaces.
- Gram-Schmidt process
- QR-factorization

Subspace and linear independence

- Subspace $S \subset \mathbb{C}^n$

S is called the subspace of the complex linear vector space $\mathbb{C}^n$,

if S is the subset of $\mathbb{C}^n$ and also a vector space

- Linear combination of vectors

$z = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k$, where $\alpha_i \in \mathbb{C}$, $\{v_i\}_{i=1}^{k} \in \mathbb{C}^n$

Consider a vector set $G = \{v_1, v_2, \ldots, v_k\} \subset \mathbb{C}^n$

- Linear span of the vector set G is the set all linear combinations:

$$span(G) = \left\{ z \in \mathbb{C}^n \,\middle|\, z = \sum_{i=1}^{k} \alpha_i v_i \right\} \text{ where } \alpha_i \in \mathbb{C},\ \{v_i\}_{i=1}^{k} \in \mathbb{C}^n$$

- Linear independent vector set $G = \{v_1, v_2, \ldots, v_k\} \subset \mathbb{C}^n$

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_k v_k = 0 \Rightarrow \alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$$

- Linear dependent vector set $\Rightarrow$ otherwise

Basis and dimension of a subspace

- Basis of a subspace is the vector set of linear independent vectors that spans this subspace (every vector of the subspace can be uniquely represented as a linear combination of the basis vectors)

The vector set G is the basis of $span(G)$, if G is linear independent

- Dimension of a subspace S is the number of vectors in its basis

If $S \subset \mathbb{C}^n$ and $G = \{v_1, v_2, ..., v_k\}$ is the basis in S, then $\dim(S) = k$

For the whole space $\mathbb{C}^n$ $\dim(\mathbb{C}^n) = n$

Important: we can define different bases in S, but every basis of S will have **the same** number of vectors

Sum and direct sum of subspaces

- Sum of two subspaces

$S = S_1 + S_2$ is the sum of S_1 and S_2, if

$\forall v \in S \;\; v = v_1 + v_2,\;\; v_1 \in S_1,\, v_2 \in S_2$

- Direct sum of two subspaces

$S = S_1 \oplus S_2$ is the direct sum of S_1 and S_2, if

1) $\forall v \in S \;\; v = v_1 + v_2,\;\; v_1 \in S_1,\, v_2 \in S_2$

2) Intersection $S_1 \cap S_2 = \{0\}$, otherwise the sum is not direct

If $S = \mathbb{C}^n \Rightarrow \;\exists!\, v_1 \in S_1,\, v_2 \in S_2 :\; v = v_1 + v_2$

$\exists$ stands for "exists", $\exists!$ stands for "exists unique"

Range and kernel of matrix

- Range of a matrix $A \in \mathbb{C}^{n\times m}$

$Ran(A) = \{Ax \mid x \in \mathbb{C}^m\},\ Ran(A) \subset \mathbb{C}^n$

Range is the linear span of matrix columns:

$$Ran(A) = span(a_1, a_2, \ldots, a_m\},\ \text{where } A = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \ldots & a_m \\ | & | & & | \end{pmatrix}$$

$\dim(Ran(A))$ is the number of linear independent columns

- Kernel (null space) of a matrix $A \in \mathbb{C}^{n\times m}$

$Null(A) = \{x \in \mathbb{C}^m \mid Ax = 0\},\ Null(A) \subset \mathbb{C}^m$

Properties

1) $\mathbb{C}^n = Ran(A) \oplus Null(A^T)$

2) $\mathbb{C}^m = Ran(A^T) \oplus Null(A)$

Rank and invariant subspace of matrix

- Rank of a matrix $A \in \mathbb{C}^{n\times m}$ is the number of linear independent columns: $rank(A) = \dim(Ran(A))$

Column rank is equal to row rank (the number of linear independent rows)

- Matrix $A \in \mathbb{C}^{n\times m}$ is the matrix of full rank, if $rank(A) = \min\{m, n\}$

Rank - nullity theorem

For a matrix $A \in \mathbb{C}^{n\times m}$ $\dim(Ran(A)) + \dim(Null(A)) = m$

- Invariant subspace of a matrix $A \in \mathbb{C}^{n\times m}$

If $AS \subset S$, then S is called invariant subspace of A

Property. $Null(A - \lambda I)$ is invariant subspace of A for any eigenvalue λ

- Eigenspace of a matrix $A \in \mathbb{C}^{n\times m}$

$Null(A - \lambda I)$ is called eigensubspace of A

Existence of solution for a linear system

Consider a linear system $Ax = b$ for a square matrix $A \in \mathbb{C}^{n \times n}$

$b \in \mathbb{C}^n$ is the right-hand side vector

$x \in \mathbb{C}^n$ is the vector of unknowns

$r = b - Ax$ is called residual vector

Usually we need residual norm $\|r\|_2$ for estimation.

x is exact solution $\Leftrightarrow r = 0$, x is approximate solution $\Leftrightarrow r \approx 0$

- The system is called consistent, if it has a least one solution.
- The system is called inconsistent, if it has no solutions.

There are three cases of the solution existence:

1) unique solution
2) many solutions
3) no solutions

Existence of solution: unique solution

1) unique solution

when A is nonsingular $(\exists A^{-1},\ \det(A) \neq 0)$

$\exists!\ x = A^{-1}b$ is the solution

Example of consistent system with unique solution.

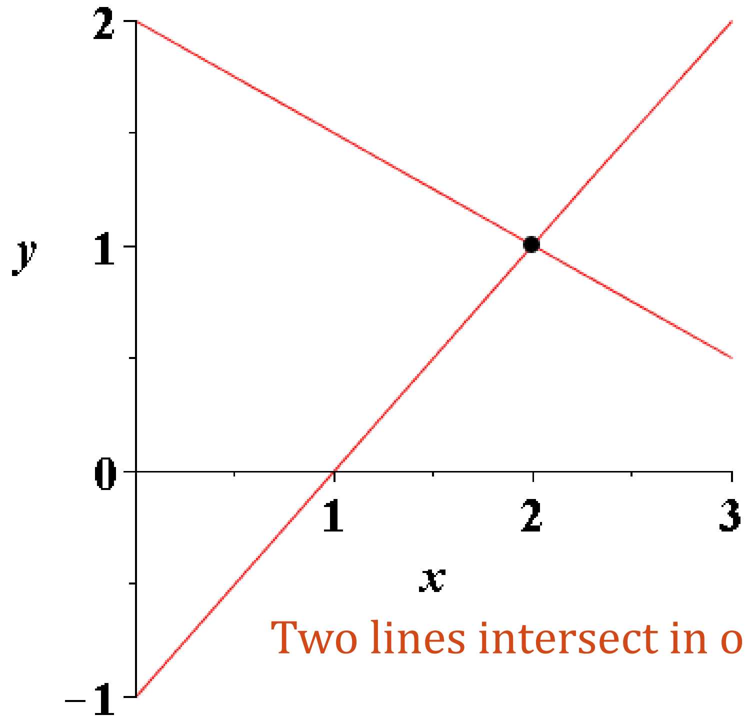
$$\begin{cases} x - y = 1 \\ x + 2y = 4 \end{cases}$$

$x = 2;\ y = 1$ is the solution

Matrix form:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix},\ \mathbf{b} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ is the solution}$$



Two lines intersect in one point

Existence of solution: many solutions

2) infinitely many solutions

when A is singular ($\det(A)=0$) and $b \in Ran(A) = \{Au \mid u \in \mathbb{C}^n\}$

$\exists\, x_0 : Ax_0 = b \Rightarrow x_0 + v$ is also the solution for $\forall v \in Null(A) = \{u \in \mathbb{C}^n \mid Au = 0\}$

$\dim(Null(A)) \geq 1$

Example of consistent system with many solutions.

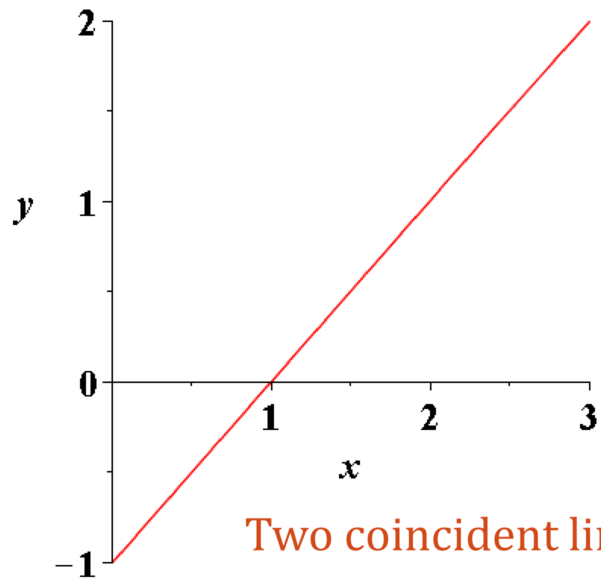
$$\begin{cases} x - y = 1 \\ -2x + 2y = -2 \end{cases}$$

$y = x - 1$ is the solution

Matrix form:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$\mathbf{x} = \begin{pmatrix} x \\ x-1 \end{pmatrix}$ is the solution



Two coincident lines

Existence of solution: no solutions

3) no solutions (inconsistent system)

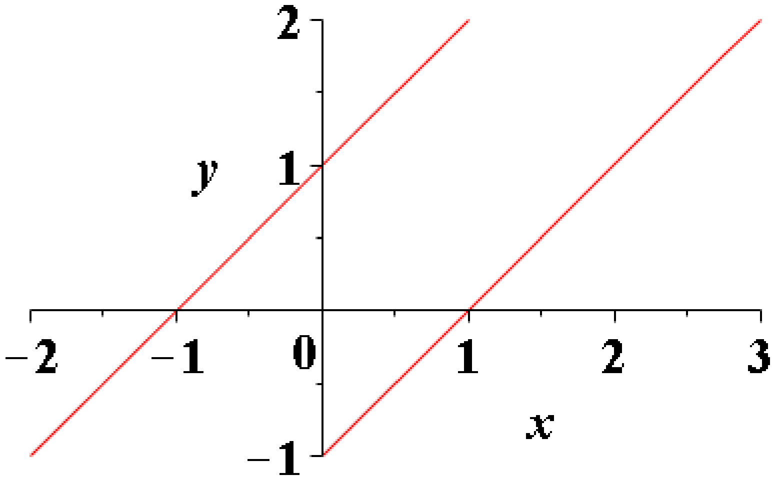
when A is singular $(\det(A) = 0)$ and $b \notin Ran(A) = \{Au \mid u \in \mathbb{C}^n\}$

Example of inconsistent system (no solutions).

$$\begin{cases} x - y = 1 \\ x - y = -1 \end{cases}$$

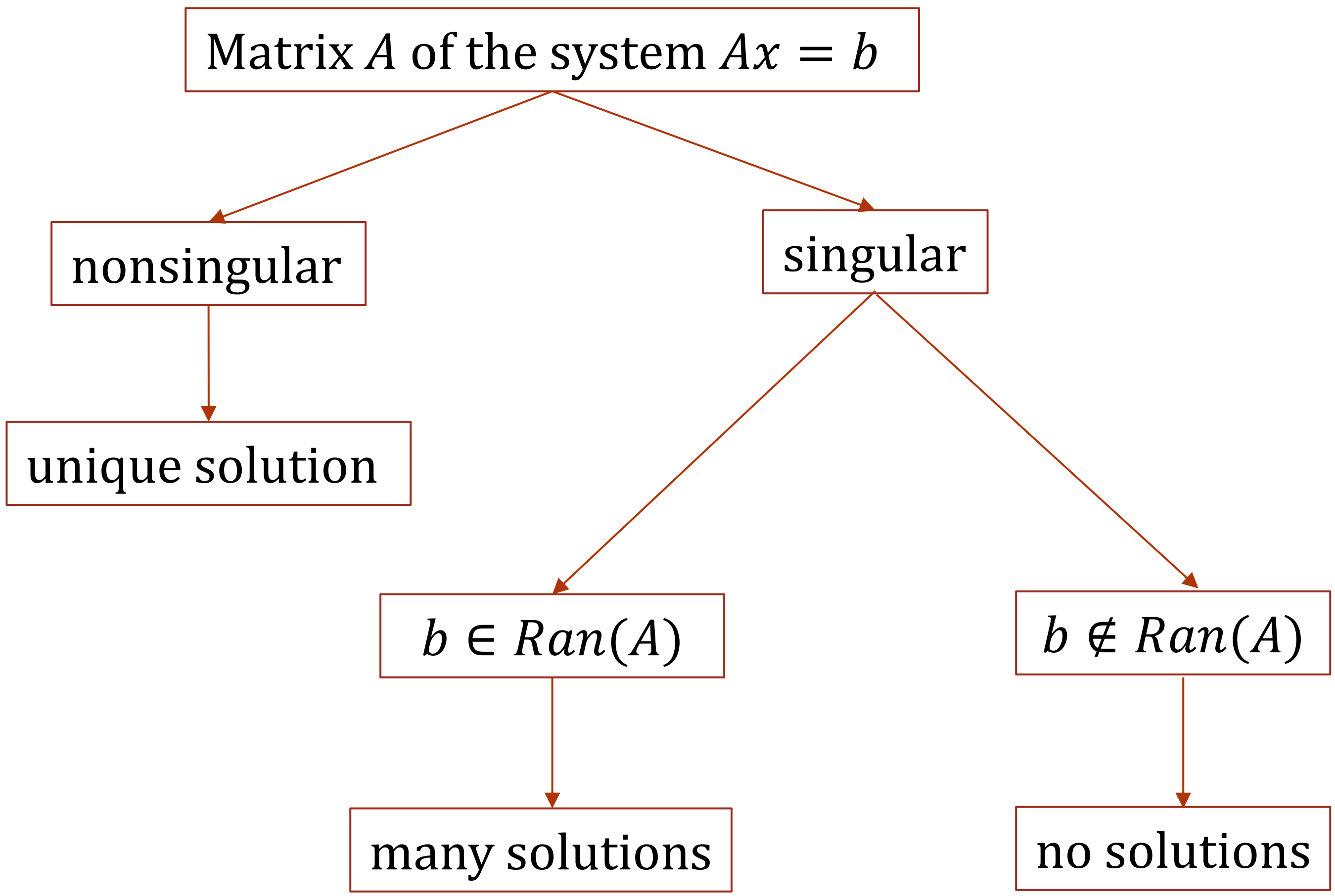
Matrix form:

$$\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$



Two parallel lines

Existence of solution: summary



Existence of solution for a system with rectangular matrix

Consider a linear system $Ax = b$ for *rectangular* matrix $A \in \mathbb{C}^{n \times m}$

$b \in \mathbb{C}^n$ is the right-hand side vector

$x \in \mathbb{C}^m$ is the vector of unknowns

n rows $\Rightarrow$ n equations

m columns $\Rightarrow$ m unknowns

- The system is called overdefined, if $n > m$.
- The system is called underdefined, if $n < m$.

For the case $n > m$ we can find solutions of a least square problem:

find $x \in \mathbb{C}^m$: $\|r\|_2 = \|b - Ax\|_2 \to \min$ (minimize the residual norm)

Orthogonality of vectors and subspaces

- orthogonal vectors

The set of vectors $G = \{v_1, v_2, \ldots, v_k\} \subset \mathbb{C}^n$ is called *orthogonal*, if $(v_i, v_j) = 0 \;\forall i \neq j$.

Notation: $v_i \perp v_j$

- orthonormal vectors

$$(v_i, v_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

$\|v_i\|_2 = 1$ (the Eucledian norm $\|v_i\|_2 = \sqrt{(v_i, v_i)}$ of every vector equals one)

- orthogonality of vector to a subspace $x \perp S$

$\forall s \in S \; x \perp s \Leftrightarrow (x, s) = 0$

- orthogonal complement

$S^\perp = \{x \mid x \perp s \;\; \forall s \in S\}$

Property. $\mathbb{C}^n = S \oplus S^\perp$

- orthogonal projector

is the mapping $S \to S^\perp$

- normalization of a vector

$$\frac{v}{\|v\|_2} = q, \text{ where } \|q\|_2 = 1$$

Gram-Schmidt process

This is a process of orthonormalization for a set of vectors.

Consider a set of vectors $\{x_1, x_2, ..., x_k\} \in \mathbb{C}^n$

The goal is to build an *orthonormal* set of vectors

$$\{q_1, q_2, ..., q_k\} \in \mathbb{C}^n, \text{ where } (q_i, q_j) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

If $\{x_i\}_{i=1}^k$ contains *linear dependent* vectors, then the resulting orthonormal set will have less number of vectors: $\{q_i\}_{i=1}^p,\ p < k$

Get the formalas to compute q_i

$$\begin{cases} x_1 = r_{11} q_1 \\ x_2 = r_{12} q_1 + r_{22} q_2 \\ \dots \\ x_k = r_{1k} q_1 + r_{2k} q_2 + \dots + r_{kk} q_k \end{cases}$$

Gram-Schmidt process

Form matrices that contain the columns of initial and resulting vectors

$$X = \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \dots & x_k \\ | & | & & | \end{pmatrix}$$

$$Q = \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \dots & q_k \\ | & | & & | \end{pmatrix}$$

$$R = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ 0 & r_{22} & & r_{2k} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & r_{kk} \end{pmatrix},$$

where $R \in \mathbb{C}^{k \times k}$, R is upper triangular

The entries of R are computed as $r_{ij} = (x_j, q_i)$

1) Take $x_1 \;\Rightarrow\; q_1 = \dfrac{x_1}{\|x_1\|_2}$

2) Take $x_2 \;\Rightarrow\; x_2 := x_2 - (x_2, q_1)q_1;$

$$q_2 = \frac{x_2}{\|x_2\|_2}$$

i) Take $x_i := x_i - (x_i, q_{i-1})q_{i-1} =$

$$= x_i - r_{i-1,i}q_{i-1};\; q_i = \frac{x_i}{\|x_i\|_2}$$

This process gives $X = QR$

Classic Gram-Schmidt algorithm

1) $r_{11} := \|x_1\|_2$. If $r_{11}=0 \Rightarrow$ stop, else $q_1 := \dfrac{x_1}{r_{11}}$

2) Loop for j from 2 to k.

2.1) $r_{ij} := (x_j, q_i),\ i = \overline{1, j-1}$

2.2) $\hat{q} := x_j - \sum_{i=1}^{j-1} r_{ij} q_i$

$\hat{q}$ is auxilary vector othogonal to all previous vectors

2.3) $r_{jj} := \|\hat{q}\|_2$. If $r_{jj}=0 \Rightarrow$ stop, else

2.4) $q_j := \dfrac{\hat{q}}{r_{jj}}$

Note that in this algorithm all diagonal entries r_{jj} of R are computed as norms. Therefore $r_{jj} \geq 0$

Modified Gram-Schmidt algorithm

For actual programming use modified G-S algorithm

- In modified G-S the sum at step 2.2 is computed in iterative process
- Modified G-S is more stable than classic G-S when initial set of vectors is almost linear dependent

1) $r_{11} := \|x_1\|_2$. If $r_{11} = 0 \Rightarrow$ stop, else $q_1 := \dfrac{x_1}{r_{11}}$

2) Loop for j from 2 to k.

2.1) $q := x_j$

2.2) Loop for i from 1 to $j-1$ $r_{ij} := (q, q_i)$, $q := q - r_{ij} q_i$

2.3) $r_{jj} := \|q\|_2$. If $r_{jj} = 0 \Rightarrow$ stop, else

2.4) $q_j := \dfrac{q}{r_{jj}}$

QR-factorization

- Factorization of a matrix is its representation as a product of several matrices, usually two or three
- QR-factorization of a matrix $A \in \mathbb{C}^{n \times m}$, $n \geq m$

$A = QR,$ where Q is unitary and R is upper triangular

The sizes of these two matrices can be diffent

- Thin (reduced) QR-factorization is

$A = QR,$ where $Q \in \mathbb{C}^{n \times m}$ and $R \in \mathbb{C}^{m \times m}$

- Full QR-factorization is

$A = QR,$ where $Q \in \mathbb{C}^{n \times n}$ and $R \in \mathbb{C}^{n \times m}$

- G-S process applied to the columns of A gives a thin QR-factorization

$$A = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \dots & a_m \\ | & | & & | \end{pmatrix}, a_j \in \mathbb{C}^n, j = \overline{1, m}$$