Numerical methods

System of linear equations (cont.)

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Determinant of the second order

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}$$

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = a_{11}b_2 - a_{21}b_1$$

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, \qquad x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Determinant of the second order

Definition: Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

be a matrix over a field F.

Then the term $a_{11}a_{22} - a_{12}a_{21}$ is called **determinant** (of second order) of matrix A and we denote it as

$$a_{11}a_{22} - a_{12}a_{21} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = |A|.$$

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

Determinant of the third order

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

Lets multiply the first equation by $a_{22}a_{33} - a_{23}a_{32}$,

the second equation by $a_{13}a_{32} - a_{12}a_{33}$,

the third equation by $a_{12}a_{23} - a_{13}a_{22}$,

and then sum-up the first two with the third one and we get:

$$\begin{aligned}&(a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32}\\ &- a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}) x_1\\ &=\\ &b_1a_{22}a_{33} - b_1a_{23}a_{32} + b_2a_{13}a_{32}\\ &- b_2a_{12}a_{33} + b_3a_{12}a_{23} - b_3a_{13}a_{22}\end{aligned}$$

Determinant of the third order

Definition: Let A is a square matrix of the order of 3 with the elements from the field F.
Then the **determinant** (of the third order) of matrix A is the term

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32}$$
$$- a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = |A|.$$

If we denote

$$A_1 = \begin{pmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix}, \quad A_3 = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{pmatrix},$$

then

$$x_1 = \frac{|A_1|}{|A|}, \quad x_2 = \frac{|A_2|}{|A|}, \quad x_3 = \frac{|A_3|}{|A|}.$$

Cramer's rule

Sarrus' rule

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Sarrus' rule

$$\begin{vmatrix} \mathbf{a_{11}} & a_{12} & a_{13} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ a_{31} & a_{32} & \mathbf{a_{33}} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$\mathbf{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}}$$

Sarrus' rule

$$\begin{vmatrix} \mathbf{a_{11}} & a_{12} & a_{13} \\ \mathbf{a_{21}} & \mathbf{a_{22}} & a_{23} \\ a_{31} & \mathbf{a_{32}} & \mathbf{a_{33}} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$\mathbf{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}}$$

Sarrus' rule

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Sarrus' rule

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Sarrus' rule

$$\begin{vmatrix} a_{11} & a_{12} & \mathbf{a_{13}} \\ a_{21} & \mathbf{a_{22}} & a_{23} \\ \mathbf{a_{31}} & a_{32} & a_{33} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Sarrus' rule

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Sarrus' rule

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{matrix}$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Determinant of a square matrix

If $\mathbf{A} = \left[a_{ij} \right]$ is a square matrix of type $n \times n$,
then the **determinant** is the exactly defined number
we denote as $\left| \mathbf{A} \right|$.

The minor $\mathbf{A}_{ij}$ is defined to be the determinant
of the $(n-1) \times (n-1)$ matrix that results from $\mathbf{A}$ by removing
the ith row and the jth column.

The determinant itself is recursively defined as follows:
If $n = 1$, then the determinant of matrix $\mathbf{A} = \left[a_{11} \right]$ of type 1×1
is simply $\left| \mathbf{A} \right| = a_{11}$.
If $n \geq 2$, then for each row index i it holds:

$$\left| \mathbf{A} \right| = (-1)^{i+1} a_{i1} \left| \mathbf{A}_{i1} \right| + (-1)^{i+2} a_{i2} \left| \mathbf{A}_{i2} \right| + \cdots + (-1)^{i+n} a_{in} \left| \mathbf{A}_{in} \right| \ .$$

This is called

the **Laplace expansion** along the ith row.

Determinant of a square matrix

If we have a **LU** decomposition of matrix **A**,
then

$$|\mathbf{A}| = |\mathbf{L}| \cdot |\mathbf{U}|$$

where

$$|\mathbf{L}| = l_{11} l_{22} l_{33} \ldots l_{nn} \qquad |\mathbf{U}| = u_{11} u_{22} u_{33} \ldots u_{nn}$$

The number of arithmetic operations of LU decomposition
is of order of n^3.
This is much less (in case of determinants of higher order)
much less then $n!$ operation
necessary to perform if we use Lapalce expansion
for calculation of determinant.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Motivation

Many practical problems
require to solve
the large systems of linear equations
$\mathbf{A}\,\mathbf{x} = \mathbf{b}$,
in which the matrix **A** is sparse,
i.e. it has relatively small number of nonzero elements.

Standard elimination methods
are not suitable
for solving such large sparse linear systems.

(*Why ?*)

Because during the elimination process
we fill-up positions of originally zero elements –
matrix is no more sparse.

Iterative methods

We chose an initial vector $\mathbf{x}_0$ and
we generate a sequence of vectors
$$\mathbf{x}_0 \rightarrow \mathbf{x}_1 \rightarrow \mathbf{x}_2 \cdots,$$
which converge to the seeking solution $\mathbf{x}$.

Common feature of all iterative methods is fact,
that each iteration step
$$\mathbf{x}_k \rightarrow \mathbf{x}_{k+1}$$
requires as many operations as
multiplication of matrix $\mathbf{A}$ by a vector,
which is for sparse matrices
relatively small number of operations.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Convergence

The most of classical iterative methods
are based on decomposition of matrix **A** = **M** – **N**,
where **M** is regular matrix.

Then we can define a sequence of $\mathbf{x}_k$ as

$$\mathbf{M}\mathbf{x}_{k+1} = \mathbf{N}\mathbf{x}_k + \mathbf{b},$$

where initial approximation $\mathbf{x}_0$ is prescribed.

We say, that the **iterative method converge,**
and we write $\mathbf{x}_k \longrightarrow \mathbf{x}$,
if numeral sequence

$$\left\|\mathbf{x}_k - \mathbf{x}\right\| \longrightarrow 0.$$

We denote $\mathbf{e}_k = \mathbf{x}_k - \mathbf{x}$ the error of k th iteration.
Because $\mathbf{M}\mathbf{x} = \mathbf{N}\mathbf{x} + \mathbf{b}$, we get

$$\mathbf{M}\left(\mathbf{x}_{k+1} - \mathbf{x}\right) = \mathbf{N}\left(\mathbf{x}_k - \mathbf{x}\right) \quad \text{or} \quad \mathbf{e}_{k+1} = \mathbf{M}^{-1}\mathbf{N}\mathbf{e}_k \quad .$$

Convergence

If we denote $\mathbf{T} = \mathbf{M}^{-1}\mathbf{N}$ then we get

$$\left\|\mathbf{e}_{k+1}\right\| \leq \left\|\mathbf{T}\right\| \cdot \left\|\mathbf{e}_{k}\right\| \leq \left\|\mathbf{T}\right\|^{2} \cdot \left\|\mathbf{e}_{k-1}\right\| \leq \cdots \leq \left\|\mathbf{T}\right\|^{k+1} \cdot \left\|\mathbf{e}_{0}\right\| .$$

Convergence of iterative method
is assured for any initial vector
if

$$\left\|\mathbf{T}\right\| < 1 ,$$

where $\mathbf{T} = \mathbf{M}^{-1}\mathbf{N}$ is **iteration matrix**.

$\left\|\mathbf{T}\right\|$ is an arbitrary matrix norm

Recall – matrix norms

“Entrywise” norm of matrix

$$\|\mathbf{A}\|_p = \|vec(\mathbf{A})\|_p = \left(\sum_{j=1}^{n} \sum_{j=1}^{n} |a_{ij}|^p \right)^{\frac{1}{p}}$$

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2}$$

Frobenius norm (consistent with l_2)

$$\|\mathbf{A}\|_\infty = \max_{i,j} |a_{ij}|$$

max norm

Conditions for stopping iterations

How to decide,
whether $\mathbf{x}_{k+1}$ is good enough approximation
of solution $\mathbf{x}$?

Usually we test one of the conditions:

1. $\left\|\mathbf{x}_{k+1} - \mathbf{x}_k\right\| \leq \varepsilon \left\|\mathbf{x}_k\right\|$
2. $\left\|\mathbf{r}_{k+1}\right\| \leq \varepsilon \left(\left\|\mathbf{A}\right\| \cdot \left\|\mathbf{x}_{k+1}\right\| + \left\|\mathbf{b}\right\|\right)$

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Jacobi iterative method

Suppose that **A** = **L** + **D** + **U**,
where **D** is diagonal matrix,
which has the same diagonal as **A**,
and where **L** and **U** are strictly lower and upper
triangular parts of matrix **A**, i.e.

$$\mathbf{D} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

$$\mathbf{L} = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ a_{21} & 0 & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & & 0 & a_{n-1,n} \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Jacobi iterative method

Jacobi method is
based on decomposition
A = M – N,
where **M = D** and **N = -(L+U)**
and we write it as

$$\mathbf{D}\mathbf{x}_{k+1} = \mathbf{b} - \left(\mathbf{L}+\mathbf{U}\right)\mathbf{x}_k$$

This system is easy to solve.
If we write it in component form then

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_j^{(k)}\right), \qquad i = 1,2,...,n \ .$$

Analysis of properties
of iteration matrix $\mathbf{T} = -\mathbf{D}^{-1}\left(\mathbf{L}+\mathbf{U}\right)$ leads to the statement, that
Jacobi method converge,
if A is diagonally dominant.

Recall

Definition: We say that square matrix $\mathbf{A} = \{a_{ij}\}_{i,j=1}^{n}$ is **diagonally dominant** if

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \qquad i = 1, 2, \ldots, n$$

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Gauss-Seidel method

Recall the Jacobi method

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)} \right) , \qquad i = 1, 2, ..., n \ .$$

If we use $x_i^{(k+1)}$ instead of $x_i^{(k)}$

we get

Gauss-Seidel method

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right) , \qquad i = 1, 2, ..., n \ .$$

Gauss-Seidel method

In a matrix form

$$\left(\mathbf{D}+\mathbf{L}\right)\mathbf{x}_{k+1} = \mathbf{b} - \mathbf{U}\mathbf{x}_k$$

Definition: Symmetric matrix $\mathbf{A}$ is **positive-definite** if

for any non-zero vector x it holds

$$x^T \cdot \mathbf{A} \cdot x > 0$$

Analysis of properties
of iteration matrix $\mathbf{T} = -\left(\mathbf{D}+\mathbf{L}\right)^{-1}\mathbf{U}$ leads to the statement, that

Gauss-Seidel method converge,

if A is diagonally dominant

or

positive-definite.

Checking, whether the matrix is positive-definite
is usually problematic.

If we multiply any regular matrix **A** from left by
their transpose matrix,
the final matrix

$$\mathbf{A}^T\mathbf{A}$$

is symmetric and positive-definite.

Therefore, for system

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

it is assured that Gauss-Seidel method converge.

In such a case the convergence could be very slow.

Jacobi method vs Gauss-Seidel method

- Convergence of Gauss-Seidel method
  is usually faster then
  convergence of Jacobi method

- There are matrices,
  for which the Gauss-Seidel method converge
  but Jacobi not
  and vice versa

- Jacobi method allows parallel processing,
  while the Gauss-Seidel method
  is sequential from their core

Summary

- Gauss elimination method and Cramer's rule lead to the solution.
  (Without round off errors we could find the exact solution.)

- The base of GEM is the modification of matrix to triangular form. (using elementary row operations)

- The influence of round-off errors on GEM could be considerable, therefore we use partial pivoting.

- The GEM is demanding from time and memory aspects. It is best suited for not very large systems with dense matrix.

- The Cramer's rule is suitable only for very small systems.

Summary

- Using the iterative method we usually find only approximate solution.
- At the beginning we chose the initial approximation of solution and we refine the solution by repeatedly inserting it into iteration formula.
  The computation is usually finished when the norm of difference of two consecutive iterations is small enough.
- Iterative methods could also diverge.
  This depend on the properties of the matrix.
- Iterative methods are suitable for solving large systems with sparse matrix.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Stationary iterative methods

Jacobi and Gauss-Seidel methods are so called
stationary iterative methods.

Stationary iterative methods
solve a linear system with an operator
approximating the original one;
and based on a measurement of the error in the result,
form a "correction equation"
for which this process is repeated.

While these methods
are simple to derive, implement, and analyze,
convergence is only guaranteed
for a limited class of matrices.

Linear stationary iterative methods are also called
relaxation methods.

Krylov subspace methods
work by forming a basis of the sequence
of successive matrix powers
times the initial residual (the **Krylov sequence**).

The approximations to the solution are then formed
by minimizing the residual over the subspace formed.
The prototypical method in this class is
the **conjugate gradient method** (CG).

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
   - Conjugate gradient method
   - Generalizations of CG method
   - Convergence of CG method
6. Preconditioning
7. Singular value decomposition

Conjugate gradient method (CG)

The **conjugate gradient method**
is implemented as an iterative algorithm,
applicable to large sparse systems
with
symmetric, positive-definite matrix.

We say that two non-zero vectors
u and **v** are conjugate
(with respect to **A**) if

$$\mathbf{u}^{\mathrm{T}}\mathbf{A}\mathbf{v} = 0$$

Conjugate gradient method (CG)

The iterative algorithms is described by this formulas:

$$\mathbf{v}^{(1)} = \mathbf{r}^{(1)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(1)} \quad ,$$

$$\alpha_k = \frac{\mathbf{v}^{(k)T}\mathbf{r}^{(k)}}{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{v}^{(k)}} \quad ,$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{v}^{(k)} \quad ,$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{A}\mathbf{v}^{(k)} \quad ,$$

$$\beta_k = -\frac{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{r}^{(k+1)}}{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{v}^{(k)}} \quad ,$$

$$\mathbf{v}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{v}^{(k)} \quad .$$

Conjugate gradient method (CG)

From the fundamentals of algorithm it follows
that after n iterations
we obtain exact solution of the system
and therefore **it is not** an iterative method
in a strict sense.

This would be true only if there are
no round-off errors.

Therefore we have to look at CG method
as an iterative method
and we have to define stop criteria.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
   - Conjugate gradient method
   - Generalizations of CG method
   - Convergence of CG method
6. Preconditioning
7. Singular value decomposition

Generalizations of CG method

In case of nonsymmetrical and
not necessary positive-definite matrices
we use
biconjugate gradient method
(e.g. subroutine **`linbcg`** from *Numerical Recipes*).

The ordinary conjugate gradient method
is their special case.

The other variant suitable for symmetric
but not positive-definite matrices $\mathbf{A}$
we obtain using exchange of all matrix multiplications
$\mathbf{a} \,.\, \mathbf{b}$ for $\mathbf{a} \,.\, \mathbf{A} \,.\, \mathbf{b}$.
This method is called **algorithm of minimal residual**,
because we minimize the form

$$Q(\mathbf{x}) = \frac{1}{2}\mathbf{r}^T \cdot \mathbf{r} = \frac{1}{2}\left|\mathbf{A}\cdot\mathbf{x} - \mathbf{b}\right|^2 \quad .$$

Generalization for nonsymmetrical matrices leads to
the **generalized minimal residual method (GMRES).**

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
   - Conjugate gradient method
   - Generalizations of CG method
   - Convergence of CG method
6. Preconditioning
7. Singular value decomposition

Convergence of CG method

Let $\mathbf{x}^{(k)}$ be an approximate solution in the k-th step of CG and let $\mathbf{x}^*$ be the exact solution.

For symmetric, positive-definite matrix we define **condition number of matrix** $\kappa(\mathbf{A})$ and **A-norm** of an arbitrary vector $\mathbf{z}$:

$$\kappa(\mathbf{A}) := \frac{\lambda_{\max}(\mathbf{A})}{\lambda_{\min}(\mathbf{A})}, \qquad \|\mathbf{z}\|_{\mathbf{A}} := (\mathbf{A}\mathbf{z}, \mathbf{z})^{1/2} .$$

Then

$$\left\|\mathbf{x}^* - \mathbf{x}^{(k)}\right\|_{\mathbf{A}} \leq 2\left[\frac{\sqrt{\kappa(\mathbf{A})}-1}{\sqrt{\kappa(\mathbf{A})}+1}\right]^k \left\|\mathbf{x}^* - \mathbf{x}^{(1)}\right\|_{\mathbf{A}}$$

If $\kappa(\mathbf{A}) >> 1$ the convergence is **very slow**.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition

Preconditioning

The aim of preconditioning is
to speed-up the convergence of iterative method in such a way,
that we solve an alternative system of linear equations
in which the coefficient matrix has

lower condition number κ

then the original coefficient matrix.

Let $\mathbf{A}.\mathbf{x} = \mathbf{b}$ be the original linear system.

Left preconditioning is defined as:
Let matrix $\mathbf{M}$ is regular and „close“ to matrix $\mathbf{A}$.
Then we solve the system

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{x} = \mathbf{M}^{-1}\mathbf{b}$$

Right preconditioning is defined as:
Let matrix $\mathbf{M}$ is regular. Then we solve system

$$\mathbf{A}\mathbf{M}^{-1}\mathbf{u} = \mathbf{b}, \quad \mathbf{x} = \mathbf{M}^{-1}\mathbf{u} \ .$$

Preconditioning in CG method

Original algorithms CG:

$$\mathbf{r}^{(1)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(1)}$$

$$\mathbf{v}^{(1)} = \mathbf{r}^{(1)}$$

$$\alpha_k = \frac{\mathbf{v}^{(k)T}\mathbf{r}^{(k)}}{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{v}^{(k)}}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{v}^{(k)}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{A}\mathbf{v}^{(k)}$$

$$\beta_k = -\frac{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{r}^{(k+1)}}{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{v}^{(k)}}$$

$$\mathbf{v}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{v}^{(k)}$$

Algorithm of preconditioned CG:

$$\mathbf{r}^{(1)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(1)}, \qquad \mathbf{z}^{(1)} = \mathbf{M}^{-1}\mathbf{r}^{(1)}$$

$$\mathbf{v}^{(1)} = \mathbf{z}^{(1)}$$

$$\alpha_k = \frac{\mathbf{z}^{(k)T}\mathbf{r}^{(k)}}{\mathbf{v}^{(k)T}\mathbf{A}\mathbf{v}^{(k)}}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{v}^{(k)}$$

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k \mathbf{A}\mathbf{v}^{(k)}, \quad \mathbf{z}^{(k+1)} = \mathbf{M}^{-1}\mathbf{r}^{(k+1)}$$

$$\beta_k = \frac{\mathbf{z}^{(k+1)T}\mathbf{r}^{(k+1)}}{\mathbf{z}^{(k)T}\mathbf{r}^{(k)}}$$

$$\mathbf{v}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta_k \mathbf{v}^{(k)}$$

Jacobi preconditioner

One of the simplest forms of preconditioning,
is obtained by the choosing the preconditioner
to be the diagonal of the matrix matice $\mathbf{A}$

$$M_{ij} = \begin{cases} A_{ii} & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Then

$$M_{ij}^{-1} = \frac{\delta_{ij}}{A_{ii}} \ .$$

This preconditioning we call
Jacobi preconditioner
or
diagonal scaling.

Advantages of Jacobi preconditioner are
the easy implementation and
low memory requirements.

Other types of preconditioning

More sophisticated choices of preconditioner
have to be a compromise
between the reduction of condition number of the matrix
and therefore faster convergence
and
the time necessary to calculate inverse matrix $\mathbf{M}^{-1}$.

More information here

M. Benzi (2002):
Preconditioning Techniques for Large Linear Systems: A Survey
Journal of Computational Physics **182**, 418-477
doi: 10.1006/jcph.2002.7176

Example

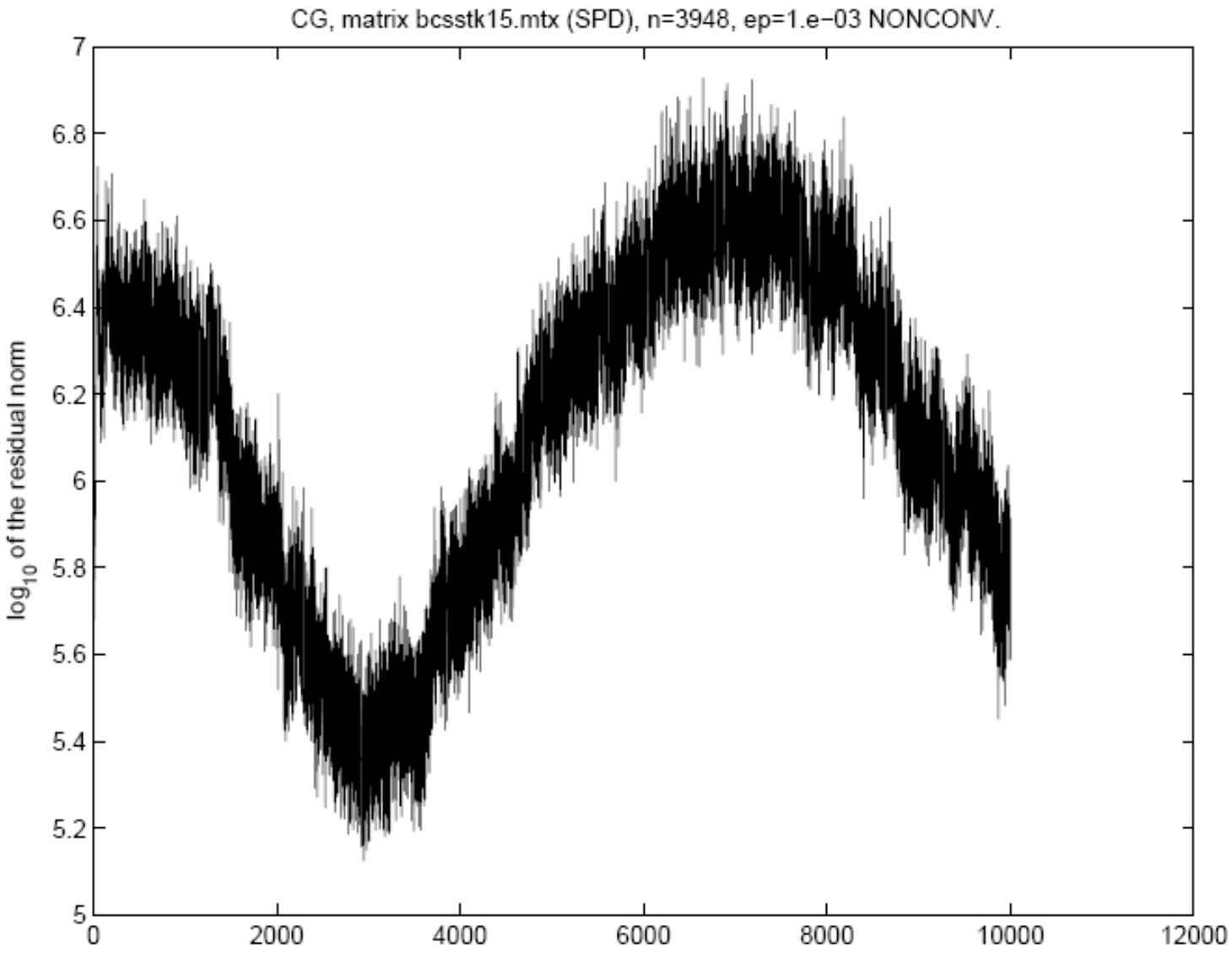


Figure: CG, $\kappa(A) = 8 * 10^9$

Example

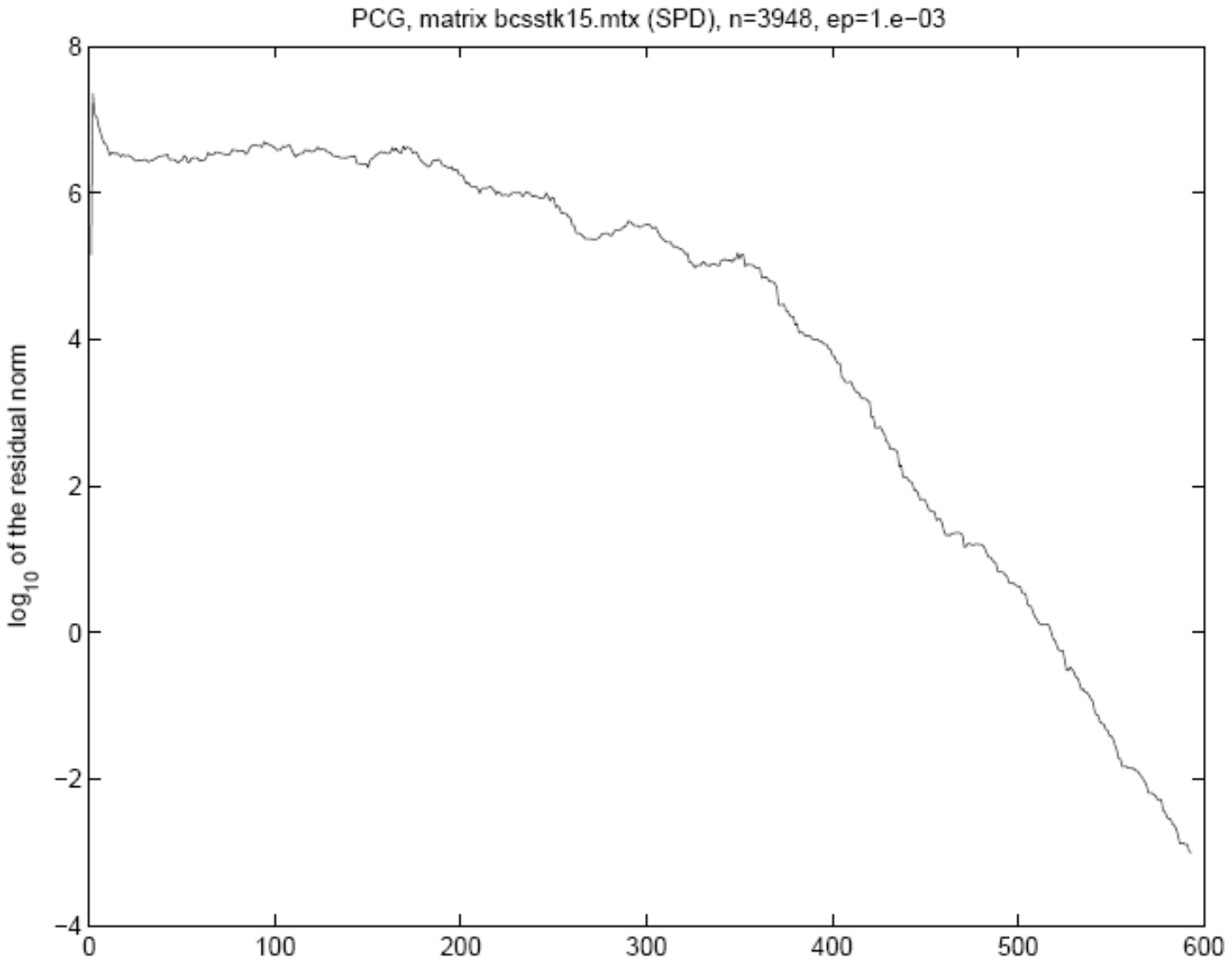


Figure: PCG, $\kappa(A) = 8 * 10^9$

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition
   - Introduction
   - SVD of a square matrix
   - SVD for less equations than unknowns
   - SVD for more equations than unknowns

Terminology and basic relations

Lets have a square matrix $\mathbf{A}$.

Its **eigenvalues** will be denoted as λ_n ,

right and **left eigenvectors** as $\mathbf{x}$ and $\mathbf{y}$,

and it holds

$$\det\left(\mathbf{A}-\lambda\mathbf{I}\right)=0$$

$$\mathbf{A}\mathbf{x}_i=\lambda_i\mathbf{x}_i$$

$$\mathbf{y}_i^T\mathbf{A}=\lambda_i\mathbf{y}_i^T$$

If $\mathbf{A}\in\Re$ and is symmetric,

then λ_i, $\mathbf{x}_i\equiv\mathbf{y}_i\in\Re$ and

eigenvectors generate the orthogonal base.

Affine transform $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$

does not change the eigenvalues of matrix $\mathbf{A}$.

Real **orthogonal matrix** is such a square matrix $\mathbf{Q}$,
for which its transpose is equal to its inverse.

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{I}$$

If $\mathbf{Q}$ is not a square matrix,
then conditions

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I} \qquad \text{and} \qquad \mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{I}$$

are not equivalent.

Condition $\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I}$ says that
matrix $\mathbf{Q}$ is **matrix with orthogonal columns.**

Condition $\mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{I}$ says that
matrix $\mathbf{Q}$ is **matrix with orthogonal rows.**

Singular value decomposition - introduction

If the number of equations M is less than the number of unknowns N,
or if $M = N$ but equations are linearly dependent,
then the system has no solution
or
it has more than one solution.

In the later case the space of solutions is given by
a particular solution
added to the any linear combination of $N - M$ vectors.

The task to find the space of solution for matrix $\mathbf{A}$
is possible to solve using
singular value decomposition of matrix $\mathbf{A}$.

Singular value decomposition - introduction

If the number of equations M is greater than the number of unknowns N,
then in general there is no solution vector and
the system is called **overdetermined**.

However, we could find a best “compromise” solution,
which is “the closest” solution
that satisfy all equations.

If „the closest“ we define in a sense of least square,
i.e. the sum of square of residuals is the smallest possible,
then the overdetermined system is reduced to
solvable problem
called the **least square method**.

Singular value decomposition - introduction

Reduced system of equations could be written as
system N x N equation

$$\left(\mathbf{A}^T \cdot \mathbf{A}\right) \cdot \mathbf{x} = \left(\mathbf{A}^T \cdot \mathbf{b}\right) \ .$$

This equations we call **normal equations**
of a least square problem.

Singular value decomposition has many common features
with the least square method,
which we show later.

Direct solution of normal equations is in general
not the best way to find a solution of least square problem.

Singular value decomposition

In many cases,
when GEM or LU decomposition fail,
singular value decomposition (SVD)
precisely diagnose, where is the problem
and in many cases it also offer
a suitable numerical solution.

SVD is also method
to solve many least square problems.

Singular value decomposition

SVD is based on the following theorem of linear algebra:

Each matrix $\mathbf{A}$ of type M x N,

which the number of rows M is greater or equal to the number of columns N,

could be decomposed to a product of

matrix with orthogonal columns $\mathbf{U}$ of type M x N,

diagonal matrix $\mathbf{W}$ of type N x N

with positive or zero entries (**singular values**)

and transpose orthogonal matrix $\mathbf{V}$ of type N x N.

$$\begin{pmatrix} \\ \mathbf{A} \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \mathbf{U} \\ \\ \end{pmatrix} \cdot \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \cdots & \\ & & \cdots & \\ & & & w_N \end{pmatrix} \cdot \begin{pmatrix} \\ \mathbf{V}^T \\ \\ \end{pmatrix}$$

Singular value decomposition

Orthogonality of matrices $\mathbf{U}$ and $\mathbf{V}$ could be written as

$$\begin{pmatrix} & \mathbf{U}^T & \end{pmatrix} \cdot \begin{pmatrix} & \mathbf{U} & \end{pmatrix} = \begin{pmatrix} & \mathbf{V}^T & \end{pmatrix} \cdot \begin{pmatrix} & \mathbf{V} & \end{pmatrix}$$

$$= \begin{pmatrix} & \mathbf{1} & \end{pmatrix}$$

Singular value decomposition

SVD could be done also if $M < N$.

In such a case the singular values w_j for $j = M+1, \ldots, N$ are all zero

as well as corresponding columns of matrix $\mathbf{U}$.

There is a lot of algorithms of SVD,
proven is subroutine **svdcmp** from *Numerical Recipes*.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition
   - Introduction
   - SVD of a square matrix
   - SVD for less equations than unknowns
   - SVD for more equations than unknowns

Singular value decomposition of a square matrix

If $\mathbf{A}$ is a square matrix of type N x N,

then $\mathbf{U}$, $\mathbf{W}$ and $\mathbf{V}$ are all square matrices of type N x N.

Because $\mathbf{U}$ and $\mathbf{V}$ are orthogonal,

their inverse are equal to transpose.

Then we can write a formula for inverse matrix $\mathbf{A}$

$$\mathbf{A}^{-1} = \mathbf{V} \cdot \left[\mathrm{diag}\left(1/w_j\right) \right] \cdot \mathbf{U}^T$$

SVD offers clear diagnosis of situation,

if some of singular values are zero

or close to zero.

Singular value decomposition of a square matrix

Recall one of definition of
condition number of matrix $\kappa(\mathbf{A})$:

$$\kappa(\mathbf{A}) := \|\mathbf{A}\|\left\|\mathbf{A}^{-1}\right\| = \frac{\max\{w_j\}}{\min\{w_j\}}$$

The matrix is **singular** if its condition umber is infinity.

The matrix is **ill-conditioned**
if reciprocal value of its condition number
is close to the machine epsilon of a computer,
i.e. less than 10^{-6} for single precision or
10^{-12} for double precision.

Singular value decomposition of a square matrix

The case of system

$$\mathbf{A}\cdot\mathbf{x}=\mathbf{b}\ ,$$

in which the coefficient matrix $\mathbf{A}$ is singular:

At first, lets have a look at the homogeneous case,
i.e. the case of $\mathbf{b}=\mathbf{0}$.

In other words
we are looking for null space of $\mathbf{A}$

$$null\left(\mathbf{A}\right)=\left\{\mathbf{x}\in\mathbb{R}^n:\mathbf{A}\mathbf{x}=\mathbf{0}\right\}=\left\{\mathbf{x}\in\mathbb{R}^n:\mathbf{U}\mathbf{w}\mathbf{V}^T\mathbf{x}=\mathbf{0}\right\}$$

$$=\left\{\mathbf{x}\in\mathbb{R}^n:\mathbf{U}^T\mathbf{U}\mathbf{w}\mathbf{V}^T\mathbf{x}=\mathbf{0}\right\}=\left\{\mathbf{x}\in\mathbb{R}^n:\mathbf{w}\mathbf{V}^T\mathbf{x}=\mathbf{0}\right\}$$

Singular value decomposition of a square matrix

The case of system

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \ ,$$

in which the coefficient matrix $\mathbf{A}$ is singular:

At first, lets have a look at the homogeneous case,
i.e. the case of $\mathbf{b}=\mathbf{0}$.

In other words
we are looking for null space of $\mathbf{A}$

SVD gives direct solution – each column of matrix $\mathbf{V}$,

which corresponding singular value w_j is zero
is a solution.

Singular value decomposition of a square matrix

The case of system

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \ ,$$

in which the coefficient matrix $\mathbf{A}$ is singular:

Now lets have a look at range of matrix $\mathbf{A}$

$$range(\mathbf{A}) = \left\{ \mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{R}^n \right\} = \left\{ \mathbf{U}\mathbf{w}\mathbf{V}^T\mathbf{x} : \mathbf{x} \in \mathbb{R}^n \right\} = \left\{ \mathbf{U}\mathbf{w}\mathbf{y} : \mathbf{y} \in \mathbb{R}^n \right\}$$

$$\mathbf{y} = \mathbf{V}^T\mathbf{x}$$

The case of system

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \ ,$$

in which the coefficient matrix $\mathbf{A}$ is singular:

Now lets have a look at range of matrix $\mathbf{A}$

The range of matrix $\mathbf{A}$ is composed by span
of columns of matrix **U**,
which corresponding singular value w_j is nonzero.

Singular value decomposition of a square matrix

The solution of a system with nonzero right-hand-side using SVD is this:

- we exchange $1/w_j$ by zero if $w_j=0$
- then we calculate (from right to left)

$$\mathbf{x} = \mathbf{V} \cdot \left[\mathrm{diag}\left(1/w_j \right) \right] \cdot \left(\mathbf{U}^T \cdot \mathbf{b} \right)$$

If a particular solution lies in range of $\mathbf{A}$,

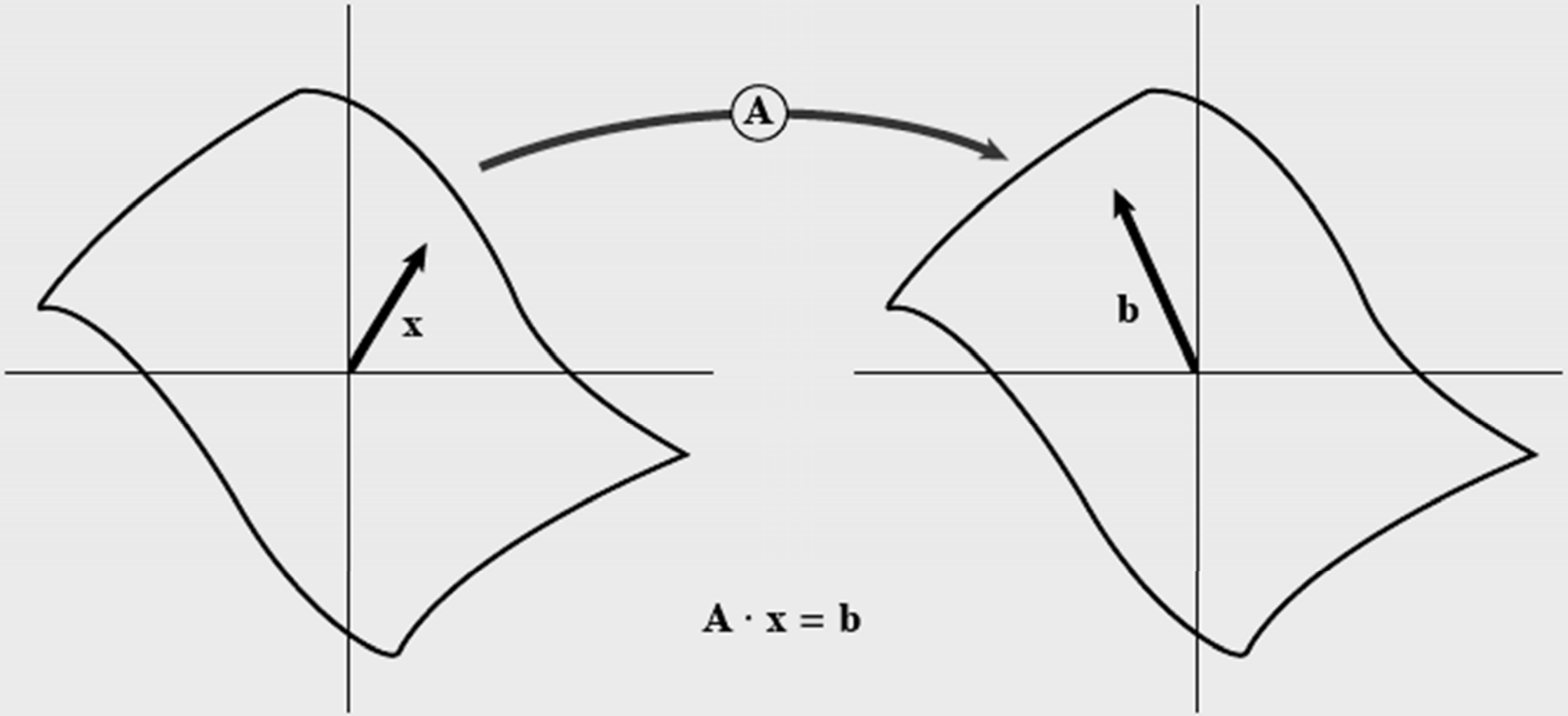
then it has the smallest size $|\mathbf{x}|^2$.

If a particular solution does not lies in range of $\mathbf{A}$,

then $\mathbf{x}$ minimize the residuum of solution $r := |\mathbf{A} \cdot \mathbf{x} - \mathbf{b}|$.

Singular value decomposition of a square matrix

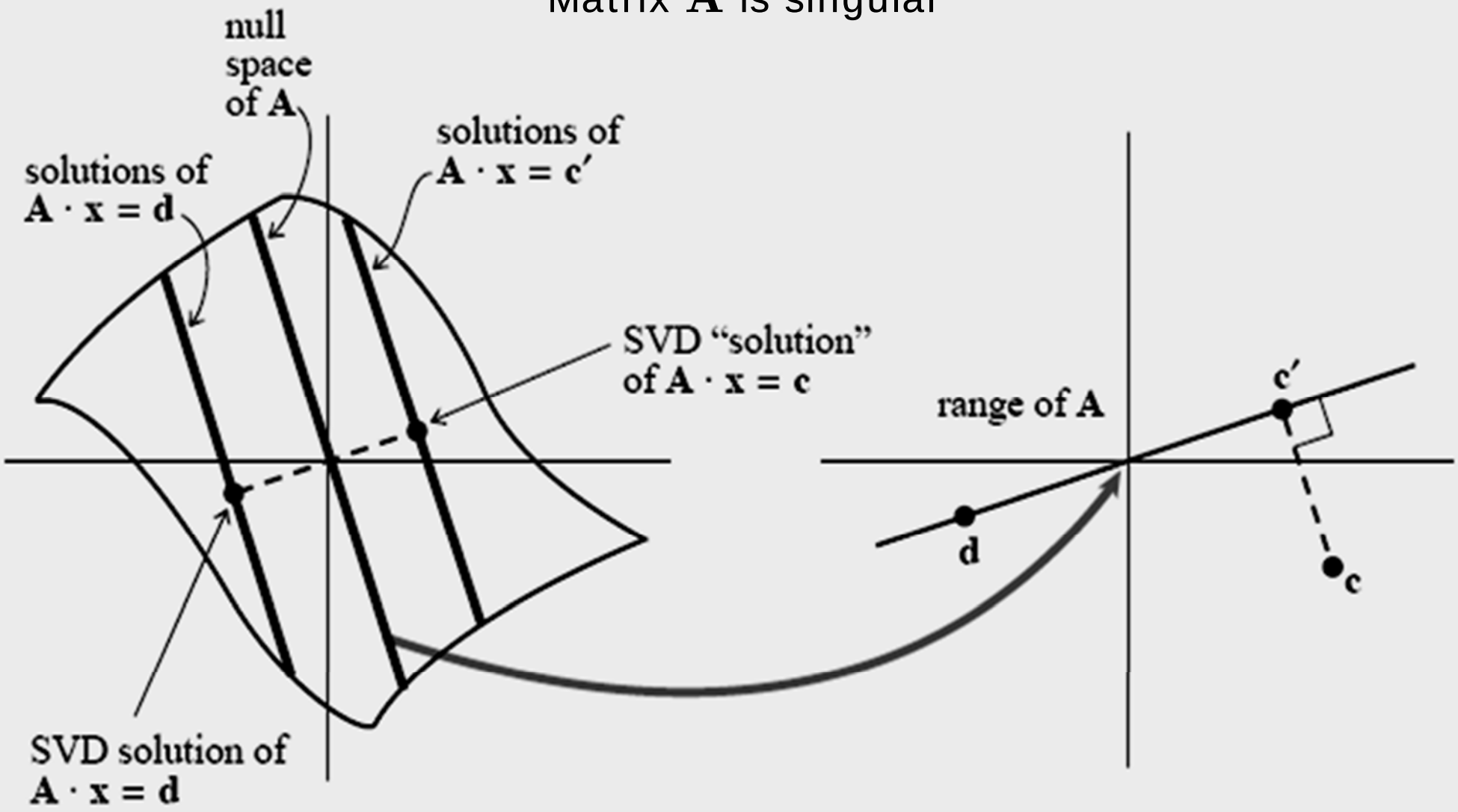
Matrix $\mathbf{A}$ is not singular



$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

Singular value decomposition of a square matrix

Matrix $\mathbf{A}$ is singular



Singular value decomposition of a square matrix

Up till now we considered only extreme cases
that the coefficient matrix is or is not singular.

There is often the case
that the singular values w_j are very small but nonzero,
so the matrix is ill-conditioned.

In such a case direct methods
can offer formally the solution,
but the solution vector has unreasonably large entries,
which during the algebraic manipulations with matrix $\mathbf{A}$
leads to very bad approximation of the right-hand-side vector.

At that time is better small values w_j set to zero and
the solution calculate using
(with replace of $1/w_j$ by zero if $w_j=0$)

$$\mathbf{x} = \mathbf{V} \cdot \left[\mathrm{diag}\left(1/w_j\right) \right] \cdot \left(\mathbf{U}^T \cdot \mathbf{b} \right)$$

We have to be cautious and
we have to chose the good threshold level for zeroise of w_j

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition
   - Introduction
   - SVD of a square matrix
   - SVD for less equations than unknowns
   - SVD for more equations than unknowns

Singular value decomposition for less equations than unknowns

If we have less equation than unknowns
we expect $N - M$ dimensional space of solutions.

In this case SVD offers
$N - M$ zero or negligible small values of w_j.
If some of M equations degenerate,
then we could have additional zero-valued w_j .

Then those columns of $\mathbf{V}$,
which corresponds to zero singular value w_j
makes a basis vectors of seeking solution space.

Particular solution could by find using

$$\mathbf{x} = \mathbf{V} \cdot \left[\mathrm{diag}\left(1/w_j\right) \right] \cdot \left(\mathbf{U}^T \cdot \mathbf{b} \right)$$

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
6. Preconditioning
7. Singular value decomposition
   - Introduction
   - SVD of a square matrix
   - SVD for less equations than unknowns
   - SVD for more equations than unknowns

Singular value decomposition for more equations than unknowns

If we have more equations than unknowns
we are seeking solution
in a least square sense.

We are solving system written as:

$$\begin{pmatrix} & \\ & \mathbf{A} \\ & \end{pmatrix} \cdot \begin{pmatrix} \\ \mathbf{x} \\ \end{pmatrix} = \begin{pmatrix} \\ \mathbf{b} \\ \end{pmatrix}$$

Singular value decomposition for more equations than unknowns

After SVD of matrix $\mathbf{A}$ we have

the solution in a form

$$\left(\mathbf{x}\right) = \left(\quad \mathbf{V} \quad\right) \cdot \left(\mathrm{diag}(1/w_j)\right) \cdot \left(\quad \mathbf{U}^T \quad\right) \cdot \left(\mathbf{b}\right)$$

In this case usually it is not necessary set to zero values w_j
however the unusually small values indicate
that the data are not sensitive to some parameters.

OUTLINE

1. Determinants and Cramer's rule
2. Iterative method for solving linear systems
3. Convergence, iteration matrix, stop criteria
4. Jacobi method, Gauss-Seidel method
5. Krylov subspace methods
   - Conjugate gradient method
   - Generalizations of CG method
   - Convergence of CG method
6. Preconditioning
7. Singular value decomposition
   - Introduction
   - SVD of a square matrix
   - SVD for less equations than unknowns
   - SVD for more equations than unknowns