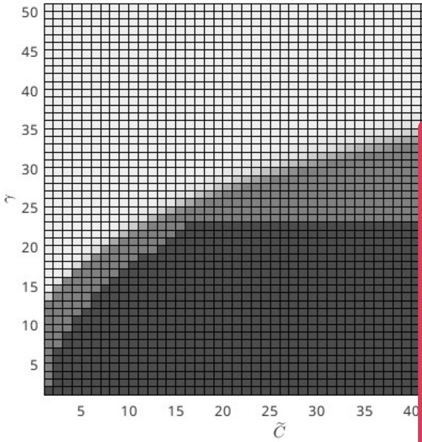

Sparse equations
the conjugate gradients method

Jochewed Schmeck

Otto-von-Guericke-Universität Magdeburg, Germany

Lima, August 19, 2025

Introduction

Definition (Overdetermined linear system): We consider the linear system:

$$Ax = b$$

with given $A \in \mathbb{R}^{m \times n}$, $m \geq n$, and

$b \in \mathbb{R}^n$,

we are looking for the unknown

$$x \in \mathbb{R}^m.$$

Repetition

Definition (QR-Decomposition): For a matrix $A \in \mathbb{R}^{m \times n}$, with $m \geq n$, there exists a decomposition

$$A = QR$$

with

$Q \in \mathbb{R}^{m \times m}$ orthonormal matrix,

$R \in \mathbb{R}^{m \times n}$ upper triangular matrix.

Repetition

We can rewrite the system $Ax = b$ as

$$QRx = b.$$

So we get the solution $x \in \mathbb{R}^n$ by solving the system

$$Rx = Q^T b.$$

Example

- First order difference quotient:

$$u'(x) \approx \frac{u(x+h) - u(x)}{h}$$

- Second order difference quotient:

$$u''(x) \approx$$

Example

- First order difference quotient:

$$u'(x) \approx \frac{u(x+h) - u(x)}{h}$$

- Second order difference quotient:

$$u''(x) \approx \frac{\frac{u(x+h)-u(x)}{h} - \frac{u(x)-u(x-h)}{h}}{h}$$

$$= \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}$$

Example

Example

- For every $i \in [1, n-1]$, there holds:

$$\frac{u(x_i + h) - 2u(x_i) + u(x_i - h)}{h^2} = f_i,$$

with $x_i + h = x_{i+1}$, and $u(x_i) = u_i$, one can rewrite this to

$$f_i = \frac{1}{h^2}u_{i-1} - \frac{2}{h^2}u_i + \frac{1}{h^2}u_{i+1} = \begin{bmatrix} \frac{1}{h^2} & -\frac{2}{h^2} & \frac{1}{h^2} \end{bmatrix} \begin{bmatrix} u_{i-1} \\ u_i \\ u_{i+1} \end{bmatrix}.$$

Example

We can write this in Matrixform:

$$\frac{1}{h^2}\begin{bmatrix} 1 & -2 & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & -2 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & 1 & -2 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \dots \\ \dots \\ u_n \end{bmatrix} = f.$$

Example

With boundary conditions and some cosmetic corrections:

$$Au := \begin{bmatrix} 2 & -1 & 0 & \dots & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & \dots & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & -1 & 2 & -1 & 0 \\ 0 & \dots & \dots & 0 & -1 & 2 & -1 \\ 0 & \dots & \dots & \dots & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \dots \\ \dots \\ \dots \\ \dots \\ u_n \end{bmatrix} = -h^2 f$$

Example

The QR-decomposition of the system matrix with $n = 6$ is

$$
A = \begin{bmatrix}
\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 & 0 \\
\frac{3}{\sqrt{70}} & \frac{3\sqrt{2}}{35} & -\frac{\sqrt{5}}{14} & 0 & 0 & 0 & 0 \\
\frac{2}{\sqrt{105}} & \frac{4}{\sqrt{105}} & \frac{2}{3\sqrt{35}} & -\sqrt{\frac{7}{15}} & 0 & 0 & 0 \\
\frac{1}{\sqrt{66}} & \frac{2}{\sqrt{33}} & \frac{3}{\sqrt{22}} & \frac{\sqrt{2}}{\sqrt{33}} & -\sqrt{\frac{6}{11}} & 0 & 0 \\
\frac{6}{\sqrt{5005}} & \frac{12}{\sqrt{5005}} & \frac{18}{\sqrt{5005}} & \frac{24}{\sqrt{5005}} & 6\sqrt{\frac{5}{1001}} & -\sqrt{\frac{55}{91}} & 0 \\
\frac{1}{\sqrt{91}} & \frac{2}{\sqrt{91}} & \frac{3}{\sqrt{91}} & \frac{4}{\sqrt{91}} & \frac{5}{\sqrt{91}} & \frac{6}{\sqrt{91}} & \frac{1}{\sqrt{13}}
\end{bmatrix}
\begin{bmatrix}
\sqrt{5} & -\frac{4}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 & 0 & 0 & 0 \\
0 & \sqrt{\frac{14}{5}} & -8\sqrt{\frac{2}{35}} & -\frac{\sqrt{5}}{14} & 0 & 0 & 0 \\
0 & 0 & \sqrt{\frac{15}{7}} & -4\sqrt{\frac{5}{21}} & \sqrt{\frac{7}{15}} & 0 & 0 \\
0 & 0 & 0 & \sqrt{\frac{11}{6}} & -8\sqrt{\frac{2}{33}} & \sqrt{\frac{6}{11}} & 0 \\
0 & 0 & 0 & 0 & \sqrt{\frac{91}{55}} & -4\sqrt{\frac{35}{143}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{7}{13}}
\end{bmatrix}
$$

Sparse matrices

Definition (Sparse matrices): A matrix $A \in \mathbb{R}^{n\times n}$ with really large n is sparse if there are only $\mathcal{O}(n)$ enties different from zero.

- These matrices can be stored by its sparsity pattern, $B \subset \{1, \ldots n\}^2$, which is the set of indices of the non-zero entries.
- 'Local equations' are sparse.
- Band matrices are sparse

It's a really nice property!

We get a problem, when we lose it...

Iterative method

Definition (Iterative method): An iterative method is defined by

$$x_{k+1} := \Psi(x_k), \qquad k \geq 0,$$

with initial value x_0.

For a given linear system $Ax = b$ with exact solution x^* the error is defined by

$$e^k := x_k - x^*, \qquad k \geq 0.$$

The residual is defined by

$$res := Ax_k - b, \qquad k \geq 0.$$

Conjugate Gradients

- Two non-zero vectors u and v are conjugate with respect to A, if
$$u^T Av = 0.$$
- For a symmetric and positive-definite matrix A, this defines an inner product:
$$u^T Av = \langle u, v\rangle_A := \langle Au, v\rangle = \langle u, A^T v\rangle = \langle u, Av\rangle.$$
- Two vectors are conjugate if and only if they are orthogonal with respect to this inner product.
- Being conjugate is a symmetric relation:
  If u is conjugate to v, then v is conjugate to u.

Conjugate Gradients

- Let $P = \{p_1, \dots, p_n\}$ be a set of n mutually conjugate vectors with respect to A:

$$p_i^T A p_j = 0 \qquad \text{for all} \qquad i \neq j.$$

- Then P forms a basis for $\mathbb{R}^n$,
- One can express the solution x^* of $Ax = b$ in this basis:

$$x^* = \sum_{i=1}^{n} \alpha_i p_i$$

$$b = Ax^* = \sum_{i=1}^{n} \alpha_i A p_i.$$

Conjugate Gradients

- Left-multiplying with the vector p_k^T yields

$$p_k^T b = p_k^T A x^* = \sum_{i=1}^{n} \alpha_i p_k^T A p_i = \sum_{i=1}^{n} \alpha_i \langle p_k, p_i \rangle_A = \alpha_k \langle p_k, p_k \rangle_A .$$

- and so we can describe α_k by:

$$\alpha_k = \frac{\langle p_k, b \rangle}{\langle p_k, p_k \rangle_A} .$$

Conjugate Gradients

- x^* minimizes the quadratic function

$$f(x) = \frac{1}{2}x^T A x - x^T b, \qquad x \in R^n \, .$$

- with Hessian matrix $$H(f(x)) = A.$$

- A is symmetric positive-definite, so a unique minimizer exists
- we use $Df(x) = 0$ to see, that the minimizer solves the initial problem follows from its first derivative

$$\nabla f(x) = Ax - b.$$

Conjugate Gradients

- Calculate the residuum
$$r_k = b - Ax_k.$$
- Calculate the next conjugate vector
$$p_k = r_k - \sum_{i<k} \frac{r_k^T A p_i}{p_i^T A p_i} p_i$$
- α_k can be derived by substituting the expression for x_{k+1} into f and minimizing it with respect to α_k:
$$\alpha_k = \frac{p_k^T (b - Ax_k)}{p_k^T A p_k} = \frac{p_k^T r_k}{p_k^T A p_k},$$
- Update x:
$$x_{k+1} = x_k + \alpha_k p_k.$$