Iterative Solvers for Large Linear Systems

Part IIIa: Method of Steepest Descent

Andreas Meister

University of Kassel, Department of Analysis and Applied Mathematics

Outline

- Basics of Iterative Methods
- Splitting-schemes
  - Jacobi- u. Gauß-Seidel-scheme
  - Relaxation methods
- Methods for symmetric, positive definite Matrices
  - Method of steepest descent
  - Method of conjugate directions
  - CG-scheme

Outline

- Multigrid Method
  - Smoother, Prolongation, Restriction
  - Twogrid Method and Extension
- Methods for non-singular Matrices
  - GMRES
  - BiCG, CGS and BiCGSTAB
- Preconditioning
  - ILU, IC, GS, SGS, ...

Projection method & Krylov subspace approach

We consider

$$Ax = b$$

with given data $A \in \mathbb{R}^{n \times n}, \quad b \in \mathbb{R}^n$.

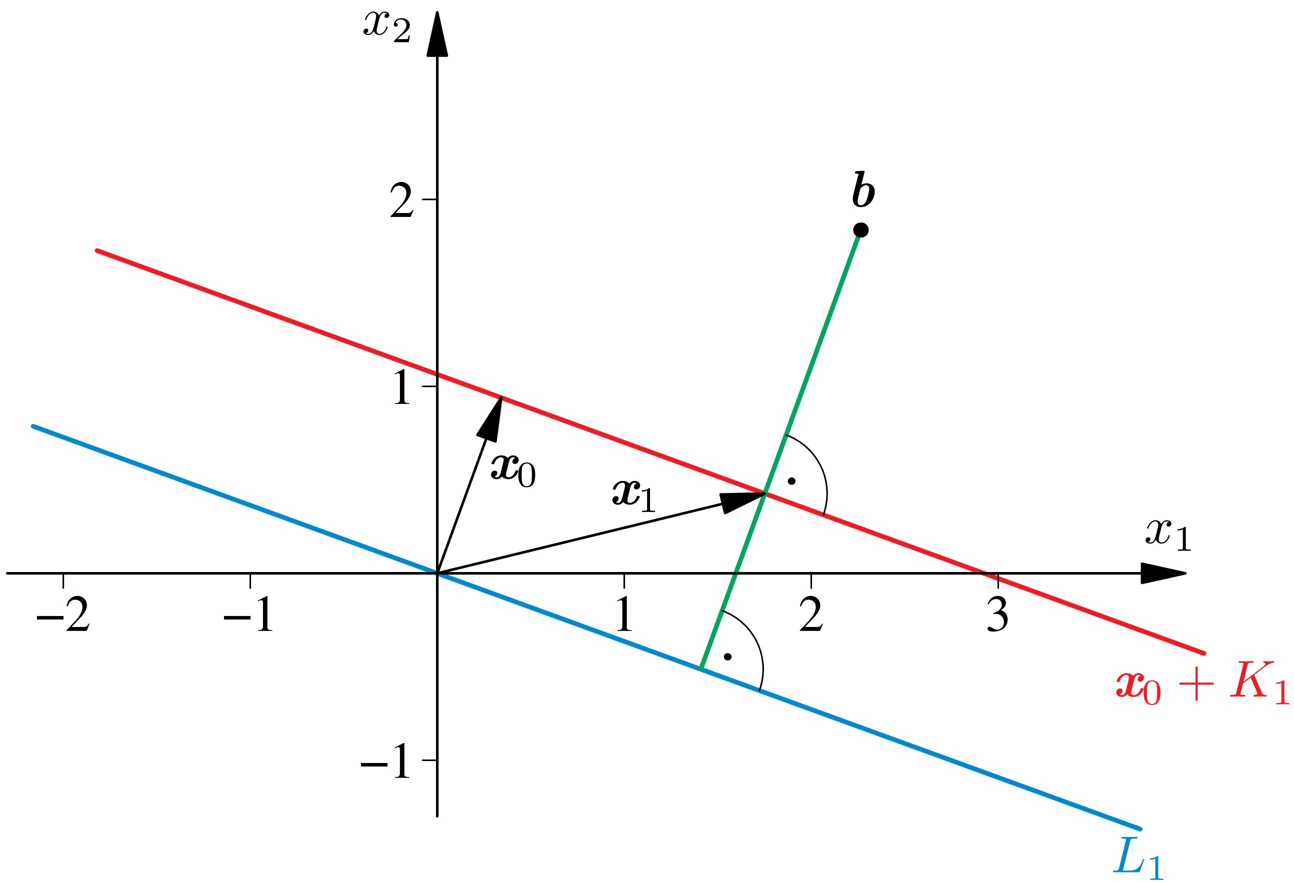
Splitting methods	**Projection methods**
Looking for approximations $x_m \in \mathbb{R}^n$	Looking for approximations $x_m \in x_0 + K_m \subset \mathbb{R}^n$ $\dim K_m = m \leq n$
Numerical algorithm $x_{m+1} = Mx_m + Nb$	Numerical algorithm (orthogonality constraint) $b - Ax_m \perp L_m \subset \mathbb{R}^n$ $\dim L_m = m \leq n$

Projection method & Krylov subspace approach

Example

$A = I \in \mathbb{R}^{2\times 2}$

$\circ K_1 = L_1$ (Orthogonal projection method)



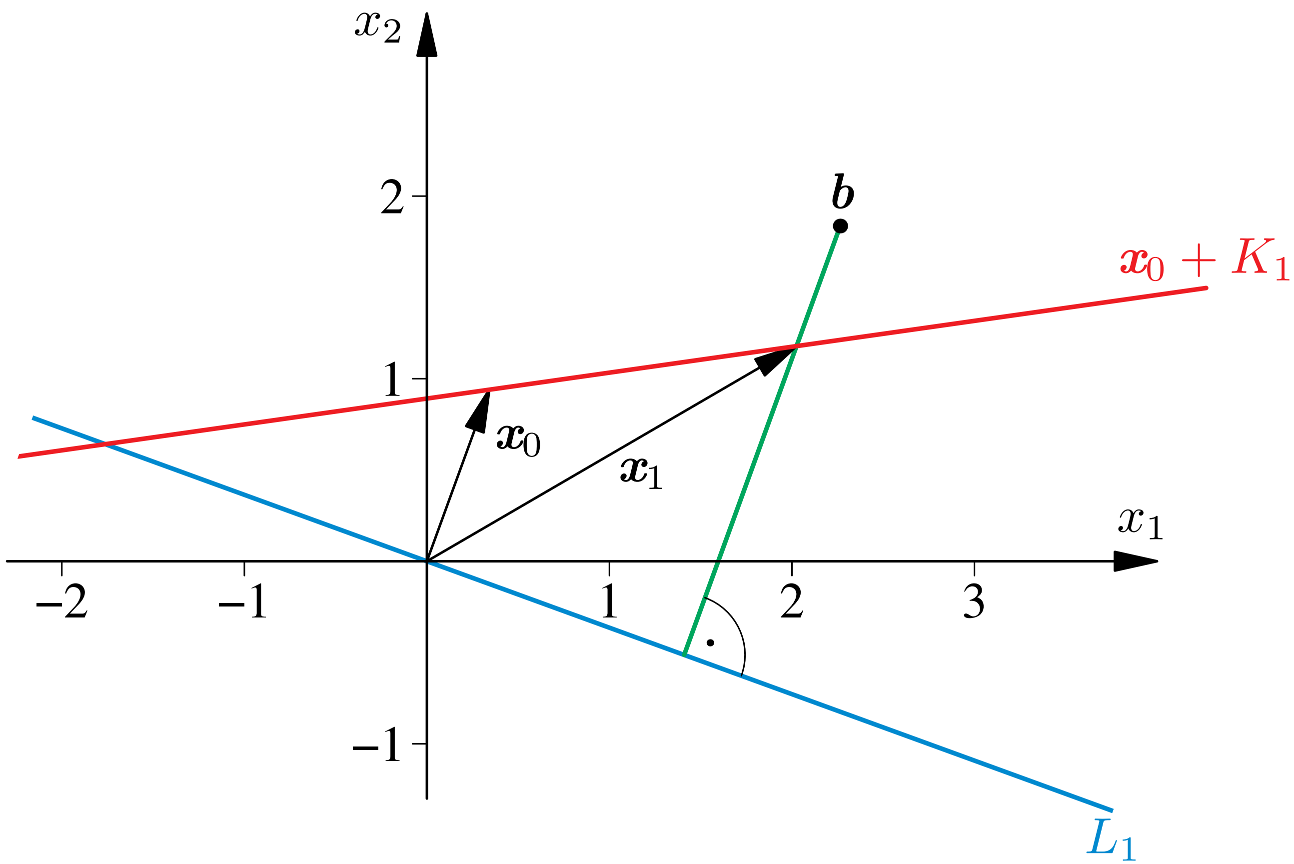
Projection method & Krylov subspace approach

Example

$A = I \in \mathbb{R}^{2\times 2}$

$\circ K_1 = L_1$ (Orthogonal projection method)

$\circ\ K_1 \neq L_1$ (Skew projection method)



Krylov subspace approach:

Projection method based on

$$K_m = K_m(A, r_0) = \text{span}\{r_0, Ar_0, \ldots, A^{m-1} r_0\},$$

with $r_0 = b - Ax_0$ is called Krylov subspace method

Methods for symmetric, positive definite matrices

Basic idea:

Minimize the function

$$F(x) = \frac{1}{2}(Ax, x) - (b, x)$$

with respect to specific search directions

$$p_0, p_1, \ldots \in \mathbb{R}^n \backslash \{0\}.$$

Procedure:

- Choose $x_0 \in \mathbb{R}^n$ and $p_0, p_1, \ldots \in \mathbb{R}^n \backslash \{0\}$.
- For $m = 0, 1, \ldots$ we calculate x_{m+1} such that

$$F(x_{m+1}) \leq F(y) \quad \forall y \in x_m + \text{span}\{p_m\}$$

$$\Longrightarrow x_{m+1} = \arg \min_{\lambda \in \mathbb{R}} \underbrace{F(x_m + \lambda p_m)}_{= f_{x_m, p_m}(\lambda)}$$

Methods for symmetric, positive definite matrices

<u>Questions:</u>

1. Does $x^\star = A^{-1}b$ represent the global minimum of F? Yes
2. How do we calculate $\lambda \in \mathbb{R}$?

Concerning 1)

$$
\begin{aligned}
& & F(x) &= \frac{1}{2}(Ax, x) - (b, x) \\
&\Longrightarrow & \nabla F(x) &= \frac{1}{2}(A + A^T)x - b \\
& & &\overset{\text{A symm.}}{=} Ax - b
\end{aligned}
$$

$$\Longrightarrow \nabla^2 F(x) = A \overset{\text{A pos.def.}}{\Longrightarrow} \text{F is a convex mapping}$$

$$\nabla F(x) = 0 \iff x = A^{-1}b$$

Methods for symmetric, positive definite matrices

<u>Questions:</u>

1. Does $x^\star = A^{-1}b$ represent the global minimum of F? Yes
2. How do we calculate $\lambda \in \mathbb{R}$? $\lambda = \dfrac{(b - Ax_m, p_m)}{(Ap_m, p_m)}$

Conc. 2)

$$
\begin{aligned}
f_{x,p}(\lambda) &= \frac{1}{2}(A(x+\lambda p), x+\lambda p) - (b, x+\lambda p) \\
&= F(x) + \lambda(Ax - b, p) + \frac{1}{2}\lambda^2(Ap, p) \\
f'_{x,p}(\lambda) &= (Ax - b, p) + \lambda(Ap, p) \\
f''_{x,p}(\lambda) &= (Ap, p) > 0 \quad \text{für} \quad p \neq 0
\end{aligned}
$$

- Thus, $f_{x,p}$ is convex and the optimal λ is given in the form

$$
f'_{x,p}(\lambda) = 0 \Longleftrightarrow \lambda = \frac{(b - Ax, p)}{(Ap, p)}.
$$

Methods for symmetric, positive definite matrices

Residual

The vector $r = b - Ax$ is called residual (vector).

Algorithm:

- Choose $x_0 \in \mathbb{R}^n$ and $p_0, p_1, \ldots \in \mathbb{R}^n \backslash \{0\}$
- For $m = 0, 1, \ldots$

$$
\begin{aligned}
r_m &= b - Ax_m \\
\lambda_m &= \frac{(r_m, p_m)}{(Ap_m, p_m)} \\
x_{m+1} &= x_m + \lambda_m p_m
\end{aligned}
$$

Problem:

Specification of the search direction $p_0, p_1, \ldots$.

Method of steepest descent

Basic idea:

Choose the optimal search direction in the local sense

$$\tilde{p}_m = -\nabla F(x_m) = -(Ax_m - b) = r_m$$

Normalizing the search direction:

$$p_m = \frac{\tilde{p}_m}{\|\tilde{p}_m\|_2} = \frac{r_m}{\|r_m\|_2}$$

Stopping criterion: $r_m = 0$

Method of steepest descent

Algorithm:

- Choose $x_0 \in \mathbb{R}^n$
- For $m = 0, 1, \ldots$

$$r_m = b - Ax_m$$

If $r_m \neq 0$

$$\begin{aligned} \lambda_m &= \frac{\|r_m\|_2^2}{(Ar_m, r_m)} \\ x_{m+1} &= x_m + \lambda_m r_m \end{aligned}$$