KUTAISI
INTERNATIONAL
UNIVERSITY

# Numerical Linear Algebra

Ramaz Botchorishvili

Kutaisi International University

December 1, 2022

KUTAISI INTERNATIONAL UNIVERSITY

# Numerical Linear Algebra

Ramaz Botchorishvili

Kutaisi International University

December 1, 2022

# Direct and iterative methods for linear systems

- Termination criteria
- Quadratic functional and linear systems
- Method of minimal residuals
- CP2
- Q & A

# Recap of Previous Lecture

- Recap of Previous Lecture
- Convergence of Richardson's iterations
- Optimal parameter parameter of Richardson's iterations
- Sufficient condition $P > 0.5A$
- Preconditioning matrices

# Termination criteria, 1

## Definition 11.1

Iteration method is linear convergent if

$$\|x^{(k+1)} - x^{(k)}\| \leq r\|x^{(k)} - x^{(k-1)}\|, \quad r < 1$$

$$\lim_{k\to\infty} x^{(k)} = A^{-1}b$$

# Termination criteria, 1

## Definition 11.1

Iteration method is linear convergent if

$$\|x^{(k+1)} - x^{(k)}\| \le r\|x^{(k)} - x^{(k-1)}\|, \quad r < 1$$
$$\lim_{k\to\infty} x^{(k)} = A^{-1}b$$

- Inequality can be easily checked during computation

# Termination criteria, 1

## Definition 11.1

Iteration method is linear convergent if

$$\|x^{(k+1)} - x^{(k)}\| \leq r\|x^{(k)} - x^{(k-1)}\|, \quad r < 1$$
$$\lim_{k\to\infty} x^{(k)} = A^{-1}b$$

- Inequality can be easily checked during computation
- Satisfied after some iterations only

# Termination criteria, 1

## Definition 11.1

Iteration method is linear convergent if

$$\|x^{(k+1)} - x^{(k)}\| \leq r\|x^{(k)} - x^{(k-1)}\|, \quad r < 1$$

$$\lim_{k\to\infty} x^{(k)} = A^{-1}b$$

- Inequality can be easily checked during computation
- Satisfied after some iterations only

$$\Downarrow$$

- $$\lim_{k\to\infty} \frac{\|x^{(k+1)}-x^{(k)}\|}{\|x^{(k)}-x^{(k-1)}\|} = r$$

# Termination criteria, 1

## Definition 11.1

Iteration method is linear convergent if

$$\|x^{(k+1)} - x^{(k)}\| \leq r\|x^{(k)} - x^{(k-1)}\|, \quad r < 1$$

$$\lim_{k\to\infty} x^{(k)} = A^{-1}b$$

- Inequality can be easily checked during computation
- Satisfied after some iterations only

$$\Downarrow$$

- $$\lim_{k\to\infty} \frac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k)} - x^{(k-1)}\|} = r$$

## Theorem 11.2

*linear convergent iteration method satisfy*

$$\|x - x^{(k)}\| \leq \frac{r}{1-r}\|x^{(k)} - x^{(k-1)}\|$$

# Termination criteria, 2

## Proof.

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$\|x^{(k+m)} - x^{(k+m-1)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + ... + \|x^{(k+1)} - x^{(k)}\|$$

# Termination criteria, 2

## Proof.

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$\|x^{(k+m)} - x^{(k+m-1)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + ... + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$r\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + ... + \|x^{(k+1)} - x^{(k)}\|$$

# Termination criteria, 2

## Proof.

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$\|x^{(k+m)} - x^{(k+m-1)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$r\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$(r+1)\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-2)} - x^{(k-m-3)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

# Termination criteria, 2

## Proof.

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$\|x^{(k+m)} - x^{(k+m-1)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$r\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$(r+1)\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-2)} - x^{(k-m-3)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq (r^{m-1} + r^{m-2} + \ldots + r + 1)\|x^{(k+1)} - x^{(k)}\|$$

# Termination criteria, 2

## Proof.

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$\|x^{(k+m)} - x^{(k+m-1)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + ... + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$r\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + ... + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq$$
$$(r+1)\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-2)} - x^{(k-m-3)}\| + ... + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq (r^{m-1} + r^{m-2} + ... + r + 1)\|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \leq \frac{1 - r^{m-1}}{r - 1}\|x^{(k+1)} - x^{(k)}\| \leq \frac{1 - r^{m-1}}{r - 1} r\|x^{(k)} - x^{(k-1)}\|$$

# Termination criteria, 2

**Proof.**

$$\|x^{(k+m)} - x^{(k)}\| \le$$
$$\|x^{(k+m)} - x^{(k+m-1)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \le$$
$$r\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-1)} - x^{(k-m-2)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \le$$
$$(r+1)\|x^{(k+m-1)} - x^{(k-m-2)}\| + \|x^{(k+m-2)} - x^{(k-m-3)}\| + \ldots + \|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \le (r^{m-1} + r^{m-2} + \ldots + r + 1)\|x^{(k+1)} - x^{(k)}\|$$

$$\|x^{(k+m)} - x^{(k)}\| \le \frac{1 - r^{m-1}}{r - 1}\|x^{(k+1)} - x^{(k)}\| \le \frac{1 - r^{m-1}}{r - 1} r\|x^{(k)} - x^{(k-1)}\|$$

$m \to \infty \Rightarrow \|x - x^{(k)}\| \le \frac{r}{1-r}\|x^{(k)} - x^{(k-1)}\|$ □

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$
- Criterion 1

$$\|b - Ax^{(k)}\| < \epsilon$$

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$
- Criterion 1

$$\|b - Ax^{(k)}\| < \epsilon$$

Bad criterion, is not scaling invariant

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$
- Criterion 1
$$\|b - Ax^{(k)}\| < \epsilon$$
Bad criterion, is not scaling invariant
- Criterion 2
$$\frac{\|b - Ax^{(k)}\|}{\|b - Ax^{(0)}\|} < \epsilon$$

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$
- Criterion 1
$$\|b - Ax^{(k)}\| < \epsilon$$
Bad criterion, is not scaling invariant
- Criterion 2
$$\frac{\|b - Ax^{(k)}\|}{\|b - Ax^{(0)}\|} < \epsilon$$
Scaling invariant, drawback = depends on initial estimate $x^{(0)}$

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$
- Criterion 1
$$\|b - Ax^{(k)}\| < \epsilon$$
Bad criterion, is not scaling invariant
- Criterion 2
$$\frac{\|b - Ax^{(k)}\|}{\|b - Ax^{(0)}\|} < \epsilon$$
Scaling invariant, drawback = depends on initial estimate $x^{(0)}$
- Criterion 3, better termination criterion
$$\frac{\|b - Ax^{(k)}\|}{\|b\|} < \epsilon$$

# Termination criteria, 3

- $\|x - x^{(k)}\| \leq \frac{r^k}{1-r}\|x^{(1)} - x^0\|$
- Criterion 1
$$\|b - Ax^{(k)}\| < \epsilon$$
Bad criterion, is not scaling invariant
- Criterion 2
$$\frac{\|b - Ax^{(k)}\|}{\|b - Ax^{(0)}\|} < \epsilon$$
Scaling invariant, drawback = depends on initial estimate $x^{(0)}$
- Criterion 3, better termination criterion
$$\frac{\|b - Ax^{(k)}\|}{\|b\|} < \epsilon$$
- Criterion 4, if $\|A^{-1}\|$ estimate available
$$\frac{\|b - Ax^{(k)}\|}{\|x^{(k)}\|} < \epsilon/\|A^{-1}\|$$

# Termination criteria, 5

## Proposition 11.3

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon / Cond(A)$$

# Termination criteria, 5

## Proposition 11.3

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon / Cond(A)$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

# Termination criteria, 5

## Proposition 11.3

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon / Cond(A)$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

## Proof.

$$Ax = b, A\tilde{x} = \tilde{b}, Right\ perturbation\ theorem$$

# Termination criteria, 5

## Proposition 11.3

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon / Cond(A)$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

## Proof.

$$Ax = b, A\tilde{x} = \tilde{b}, Right\ perturbation\ theorem$$

$$\Downarrow$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq cond(A)\frac{\|b - \tilde{b}\|}{\|b\|}$$

# Termination criteria, 5

## Proposition 11.3

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon / Cond(A)$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

## Proof.

$$Ax = b, A\tilde{x} = \tilde{b}, Right\ perturbation\ theorem$$

$$\Downarrow$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq cond(A) \frac{\|b - \tilde{b}\|}{\|b\|}$$

$$Ax = b, Ax^{(k)} = b + (Ax^{(k)} - b)$$

# Termination criteria, 5

## Proposition 11.3

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon / Cond(A)$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

## Proof.

$$Ax = b, A\tilde{x} = \tilde{b}, Right\ perturbation\ theorem$$

$$\Downarrow$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq cond(A) \frac{\|b - \tilde{b}\|}{\|b\|}$$

$$Ax = b, Ax^{(k)} = b + (Ax^{(k)} - b) \qquad \Downarrow$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq cond(A) \frac{\|b - Ax^{(k)}\|}{\|b\|} \leq \epsilon$$

# Termination criteria, 6

## Proposition 11.4

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|x^{(k)}\|} \leq \epsilon / \|A^{-1}\|$$

# Termination criteria, 6

## Proposition 11.4

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|x^{(k)}\|} \leq \epsilon / \|A^{-1}\|$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \approx \frac{\|x - x^{(k)}\|}{\|x^{(k)}\|}$$

# Termination criteria, 6

## Proposition 11.4

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|x^{(k)}\|} \leq \epsilon / \|A^{-1}\|$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \approx \frac{\|x - x^{(k)}\|}{\|x^{(k)}\|}$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

# Termination criteria, 6

## Proposition 11.4

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|x^{(k)}\|} \leq \epsilon / \|A^{-1}\|$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \approx \frac{\|x - x^{(k)}\|}{\|x^{(k)}\|}$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

## Proof.

$$\frac{\|x - x^{(k)}\|}{\|x^{(k)}\|} = \frac{\|A^{-1}(b - Ax^{(k)})\|}{\|x^{(k)}\|}$$

# Termination criteria, 6

## Proposition 11.4

*if*

$$\frac{\|b - Ax^{(k)}\|}{\|x^{(k)}\|} \leq \epsilon / \|A^{-1}\|$$

$$\frac{\|x - x^{(k)}\|}{\|x\|} \approx \frac{\|x - x^{(k)}\|}{\|x^{(k)}\|}$$

*then*

$$\frac{\|x - x^{(k)}\|}{\|x\|} \leq \epsilon$$

## Proof.

$$\frac{\|x - x^{(k)}\|}{\|x^{(k)}\|} = \frac{\|A^{-1}(b - Ax^{(k)})\|}{\|x^{(k)}\|}$$

$$\frac{\|x - x^{(k)}\|}{\|x^{(k)}\|} \leq \frac{\|A^{-1}\|\|b - Ax^{(k)}\|}{\|x^{(k)}\|} \leq \epsilon$$

□

# Quadratic functional & $Ax = b$, 1

### Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $\Downarrow$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2,3):$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2, 3)$ :
  1. $Ay = b$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2, 3)$ :
  1. $Ay = b$
  2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n\times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2, 3)$ :
  1. $Ay = b$
  2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
  3. $f(x) = 0.5(Ax, x) - (b, x)$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n\times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2,3)$ :
  1. $Ay = b$
  2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
  3. $f(x) = 0.5(Ax,x) - (b,x)$

## Proof.

- two steps

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2, 3)$ :
  1. $Ay = b$
  2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
  3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- two steps
- $1 \Rightarrow (2, 3)$

# Quadratic functional & $Ax = b$, 1

## Theorem 11.5

- $A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $\Downarrow$
- $1 \equiv (2, 3)$ :
  1. $Ay = b$
  2. $y = arg \min_{x \in \mathbb{R}^n} f(x)$
  3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- two steps
- $1 \Rightarrow (2, 3)$
- $(2, 3) \Rightarrow 1$

$\square$

# Quadratic functional $\&Ax = b$, 2

**1,2,3**

1. $Ay = b$
2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

# Quadratic functional & $Ax = b$, 2

**1,2,3**

1. $Ay = b$
2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

**Proof.**

- $(2, 3) \Rightarrow 1$

# Quadratic functional &$Ax = b$, 2

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $(2, 3) \Rightarrow 1$
- $f(x) = 0.5(Ax, x) - (b, x) = 0.5\sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}x_j)x_i - \sum_{i=1}^{n} b_i x_i$

# Quadratic functional $\& Ax = b$, 2

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $(2, 3) \Rightarrow 1$
- $f(x) = 0.5(Ax, x) - (b, x) = 0.5 \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} x_j) x_i - \sum_{i=1}^{n} b_i x_i$
- $\frac{\partial f(x)}{\partial x_k} = 0.5(\sum_{i=1}^{n} a_{ik} x_i + \sum_{j=1}^{n} a_{kj} x_j) - b_k$

# Quadratic functional $\& Ax = b$, 2

## 1,2,3

1. $Ay = b$
2. $y = arg \min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $(2, 3) \Rightarrow 1$
- $f(x) = 0.5(Ax, x) - (b, x) = 0.5 \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} x_j) x_i - \sum_{i=1}^{n} b_i x_i$
- $\frac{\partial f(x)}{\partial x_k} = 0.5(\sum_{i=1}^{n} a_{ik} x_i + \sum_{j=1}^{n} a_{kj} x_j) - b_k$
- $\nabla f(x) = 0.5(A^T + Ax) - b = Ax - b$

# Quadratic functional $\&Ax = b$, 2

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $(2,3) \Rightarrow 1$
- $f(x) = 0.5(Ax, x) - (b, x) = 0.5\sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}x_j)x_i - \sum_{i=1}^{n} b_i x_i$
- $\frac{\partial f(x)}{\partial x_k} = 0.5(\sum_{i=1}^{n} a_{ik}x_i + \sum_{j=1}^{n} a_{kj}x_j) - b_k$
- $\nabla f(x) = 0.5(A^T + Ax) - b = Ax - b$
- $Ay = b \Rightarrow \nabla f(y) = 0$

# Quadratic functional & $Ax = b$, 2

## 1,2,3

1. $Ay = b$
2. $y = arg\,\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $(2, 3) \Rightarrow 1$
- $f(x) = 0.5(Ax, x) - (b, x) = 0.5 \sum_{i=1}^{n} (\sum_{j=1}^{n} a_{ij} x_j) x_i - \sum_{i=1}^{n} b_i x_i$
- $\frac{\partial f(x)}{\partial x_k} = 0.5(\sum_{i=1}^{n} a_{ik} x_i + \sum_{j=1}^{n} a_{kj} x_j) - b_k$
- $\nabla f(x) = 0.5(A^T + Ax) - b = Ax - b$
- $Ay = b \Rightarrow \nabla f(y) = 0$
- $\nabla^2 f(x) = A > 0 \Rightarrow (2, 3) \Rightarrow 1$

$\square$

# Quadratic functional & $Ax = b$, 3

**1,2,3**

1. $Ay = b$
2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

# Quadratic functional & $Ax = b$, 3

**1,2,3**

1. $Ay = b$
2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

**Proof.**

- $1 \Rightarrow (2, 3)$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$

# Quadratic functional & $Ax = b$, 3

**1,2,3**

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

**Proof.**

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$
- $f(x) = 0.5(Ay, y) - (b, y) + 0.5(Az, z) + (Ay, z) - (b, z)$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x \in \mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$
- $f(x) = 0.5(Ay, y) - (b, y) + 0.5(Az, z) + (Ay, z) - (b, z)$
- $f(x) = f(y) + 0.5\|z\|_A + (Ay - b, z) = f(y) + 0.5\|z\|_A$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$
- $f(x) = 0.5(Ay, y) - (b, y) + 0.5(Az, z) + (Ay, z) - (b, z)$
- $f(x) = f(y) + 0.5\|z\|_A + (Ay - b, z) = f(y) + 0.5\|z\|_A$
- $f(x) = f(y) + 0.5\|z\|_A$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$
- $f(x) = 0.5(Ay, y) - (b, y) + 0.5(Az, z) + (Ay, z) - (b, z)$
- $f(x) = f(y) + 0.5\|z\|_A + (Ay - b, z) = f(y) + 0.5\|z\|_A$
- $f(x) = f(y) + 0.5\|z\|_A$
- $f(x) - f(y) = 0.5\|z\|_A \geq 0$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$
- $f(x) = 0.5(Ay, y) - (b, y) + 0.5(Az, z) + (Ay, z) - (b, z)$
- $f(x) = f(y) + 0.5\|z\|_A + (Ay - b, z) = f(y) + 0.5\|z\|_A$
- $f(x) = f(y) + 0.5\|z\|_A$
- $f(x) - f(y) = 0.5\|z\|_A \geq 0 \Rightarrow f(x) \geq f(y)\ \forall x$

# Quadratic functional & $Ax = b$, 3

## 1,2,3

1. $Ay = b$
2. $y = arg\min_{x\in\mathbb{R}^n} f(x)$
3. $f(x) = 0.5(Ax, x) - (b, x)$

## Proof.

- $1 \Rightarrow (2, 3)$
- $f(x) = f(y + x - y) = f(x + z) = 0.5(A(y + z), y + z) - (b, y + z)$
- $f(x) = 0.5[(Ay, y) + (Az, z) + (Ay, z) + (Az, y)] - (b, y) - (b, z)$
- $f(x) = 0.5(Ay, y) - (b, y) + 0.5(Az, z) + (Ay, z) - (b, z)$
- $f(x) = f(y) + 0.5\|z\|_A + (Ay - b, z) = f(y) + 0.5\|z\|_A$
- $f(x) = f(y) + 0.5\|z\|_A$
- $f(x) - f(y) = 0.5\|z\|_A \geq 0 \Rightarrow f(x) \geq f(y)\ \forall x \ \ \Rightarrow 1 \Rightarrow (2, 3)$

$\square$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k(b - Ax^{(k)})$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, ...$$

- $x^{(k+1)} = x^{(k)} + \alpha_k(b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k (b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \dots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k(b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?
- $A \in \mathbb{R}^{n \times n}, A = A^T > 0,\ Ae^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r_k$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k(b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?
- $A \in \mathbb{R}^{n\times n}, A = A^T > 0,\ Ae^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r_k$
- "Best" $\equiv$ minimal residual $\|r^{(k+1)}\|_2$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \dots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k (b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?
- $A \in \mathbb{R}^{n \times n}, A = A^T > 0,\ Ae^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r_k$
- "Best" $\equiv$ minimal residual $\|r^{(k+1)}\|_2$
- $r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r_k) = b - Ax^{(k)} - \alpha_k A r_k$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k(b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?
- $A \in \mathbb{R}^{n \times n}, A = A^T > 0,\ Ae^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r_k$
- "Best" $\equiv$ minimal residual $\|r^{(k+1)}\|_2$
- $r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r_k) = b - Ax^{(k)} - \alpha_k Ar_k$
- $r^{(k+1)} = r_k - \alpha_k Ar_k$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k(b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?
- $A \in \mathbb{R}^{n\times n}, A = A^T > 0,\ Ae^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r_k$
- "Best" $\equiv$ minimal residual $\|r^{(k+1)}\|_2$
- $r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r_k) = b - Ax^{(k)} - \alpha_k A r_k$
- $r^{(k+1)} = r_k - \alpha_k A r_k$
- $\varphi(\alpha) = \|r_k - \alpha A r_k\|_2 \Rightarrow$

# Method of minimal residuals, Derivation, 1

- Non-stationary Richardson method

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b, \quad k = 0, 1, 2, \ldots$$

- $x^{(k+1)} = x^{(k)} + \alpha_k (b - Ax^{(k)})$
- $x^{(k+1)} = x^{(k)} + \alpha_k r_k$
- What is "best" $\alpha_k$?
- $A \in \mathbb{R}^{n \times n}, A = A^T > 0,\ Ae^{(k)} = A(x - x^{(k)}) = b - Ax^{(k)} = r_k$
- "Best" $\equiv$ minimal residual $\|r^{(k+1)}\|_2$
- $r^{(k+1)} = b - Ax^{(k+1)} = b - A(x^{(k)} + \alpha_k r_k) = b - Ax^{(k)} - \alpha_k A r_k$
- $r^{(k+1)} = r_k - \alpha_k A r_k$
- $\varphi(\alpha) = \|r_k - \alpha A r_k\|_2 \Rightarrow$ "Best" $\alpha_{opt} = arg \min \varphi(\alpha)$

# Method of minimal residuals, Derivation, 2

- $f(x) = 0.5(Ax, x) - (b, x)$, $A \in \mathbb{R}^{n \times n}, A = A^T > 0$

# Method of minimal residuals, Derivation, 2

- $f(x) = 0.5(Ax, x) - (b, x)$, $A \in \mathbb{R}^{n\times n}, A = A^T > 0$
-

$$
\begin{aligned}
\varphi^2(\alpha) = \|r_k - \alpha A r_k\|_2 = (r_k - \alpha A r_k, r_k - \alpha A r_k) = \\
= (r_k, r_k) - \alpha(r_k, Ar_k) - \alpha(Ar_k, r_k) + \alpha^2(Ar_k, Ar_k) \\
= \|r_k\|_2^2 - 2\alpha(r_k, Ar_k) + \alpha^2\|Ar_k\|_2^2
\end{aligned}
$$

# Method of minimal residuals, Derivation, 2

- $f(x) = 0.5(Ax, x) - (b, x)$, $A \in \mathbb{R}^{n\times n}, A = A^T > 0$
- $$\begin{aligned} \varphi^2(\alpha) = \|r_k - \alpha A r_k\|_2 = (r_k - \alpha A r_k, r_k - \alpha A r_k) = \\ = (r_k, r_k) - \alpha(r_k, A r_k) - \alpha(A r_k, r_k) + \alpha^2 (A r_k, A r_k) \\ = \|r_k\|_2^2 - 2\alpha(r_k, A r_k) + \alpha^2 \|A r_k\|_2^2 \end{aligned}$$
- "Best" $\alpha_{opt} = arg\min \varphi(\alpha)$

# Method of minimal residuals, Derivation, 2

- $f(x) = 0.5(Ax, x) - (b, x)$, $A \in \mathbb{R}^{n\times n}, A = A^T > 0$
-

$$\begin{aligned}\varphi^2(\alpha) = \|r_k - \alpha A r_k\|_2 = (r_k - \alpha A r_k, r_k - \alpha A r_k) = \\ = (r_k, r_k) - \alpha(r_k, Ar_k) - \alpha(Ar_k, r_k) + \alpha^2(Ar_k, Ar_k) \\ = \|r_k\|_2^2 - 2\alpha(r_k, Ar_k) + \alpha^2\|Ar_k\|_2^2\end{aligned}$$

- "Best" $\alpha_{opt} = arg\min \varphi(\alpha)$
- $\varphi'_\alpha = 2\alpha(Ar_k, Ar_k) - 2\alpha(Ar_k, r_k)$

# Method of minimal residuals, Derivation, 2

- $f(x) = 0.5(Ax, x) - (b, x),\ A \in \mathbb{R}^{n \times n}, A = A^T > 0$
- $$\begin{aligned} \varphi^2(\alpha) = \|r_k - \alpha A r_k\|_2 = (r_k - \alpha A r_k, r_k - \alpha A r_k) = \\ = (r_k, r_k) - \alpha(r_k, A r_k) - \alpha(A r_k, r_k) + \alpha^2 (A r_k, A r_k) \\ = \|r_k\|_2^2 - 2\alpha(r_k, A r_k) + \alpha^2 \|A r_k\|_2^2 \end{aligned}$$
- "Best" $\alpha_{opt} = arg \min \varphi(\alpha)$
- $\varphi'_\alpha = 2\alpha(Ar_k, Ar_k) - 2\alpha(Ar_k, r_k)$
- $\varphi'_\alpha = 0 \Rightarrow \alpha_{opt} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$Ax = b, A = A^T > 0$

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$$Ax = b, A = A^T > 0$$

- Choose $x^{(0)}$, set $k = 0$

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$Ax = b, A = A^T > 0$

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$$Ax = b, A = A^T > 0$$

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:
  1. $r_k = b - Ax^{(k)}$

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$$Ax = b, A = A^T > 0$$

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:
  1. $r_k = b - Ax^{(k)}$
  2. $\alpha_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$$Ax = b, A = A^T > 0$$

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:
  1. $r_k = b - Ax^{(k)}$
  2. $\alpha_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
  3. $x^{(k+1)} = x^{(k)} + \alpha_k r_k$

# Method of minimal residuals, derivation, 3

Algorithm of the method of minimal residuals for solving
$$Ax = b, A = A^T > 0$$

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:
  1. $r_k = b - Ax^{(k)}$
  2. $\alpha_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
  3. $x^{(k+1)} = x^{(k)} + \alpha_k r_k$

# Method of minimal residuals, convergence, 1

**Method of minimal residuals** for solving $Ax = b, A = A^T > 0$ is **non-stationary Richardson**

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:
    1. $r_k = b - Ax^{(k)}$
    2. $\alpha_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
    3. $x^{(k+1)} = x^{(k)} + \alpha_k r_k$

# Method of minimal residuals, convergence, 1

**Method of minimal residuals** for solving $Ax = b, A = A^T > 0$ is **non-stationary Richardson**

- Choose $x^{(0)}$, set $k = 0$
- Do until termination criterion is satisfied:
  1. $r_k = b - Ax^{(k)}$
  2. $\alpha_k = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
  3. $x^{(k+1)} = x^{(k)} + \alpha_k r_k$

$$\Downarrow$$

$$x^{(k+1)} = x^{(k)} + \alpha_k r_k = x^{(k)} + \alpha_k (b - Ax^{(k)})$$

$$\Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k (b - Ax^{(k)})$$

$$\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b$$

$$B_{minres} = B_{R,\alpha_k}, \quad x^{(k+1)} = (I - \alpha_k A)x^{(k)} + b$$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k\to\infty} x_{minres}^{(k)} = A^{-1}b$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > \dots > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k \to \infty} x_{minres}^{(k)} = A^{-1}b$
- $\|r_{minres,k}\|_2 \leq \rho_{opt}^k \|r_{minres,0}\|_2$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k\to\infty} x_{minres}^{(k)} = A^{-1}b$
- $\|r_{minres,k}\|_2 \leq \rho_{opt}^k \|r_{minres,0}\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k\to\infty} x_{minres}^{(k)} = A^{-1}b$
- $\|r_{minres,k}\|_2 \leq \rho_{opt}^k \|r_{minres,0}\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}$

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k\to\infty} x_{minres}^{(k)} = A^{-1}b$
- $\|r_{minres,k}\|_2 \leq \rho_{opt}^k \|r_{minres,0}\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}$

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k,r_k)}{(Ar_k,Ar_k)}$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k \to \infty} x^{(k)}_{minres} = A^{-1}b$
- $\|r_{minres,k}\|_2 \leq \rho^k_{opt} \|r_{minres,0}\|_2$
- $\rho_{opt} = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)}$

## Proof.

- Richardson: $\frac{x^{(k+1)} - x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A) + \lambda_n(A)}$

# Method of minimal residuals, convergence, 2

## Theorem 11.6

- $Ax = b, A = A^T > 0$
- $\lambda_1(A) > \lambda_2(A) > ... > \lambda_n(A) > 0$
- $\Downarrow$
- $\lim_{k\to\infty} x_{minres}^{(k)} = A^{-1}b$
- $\|r_{minres,k}\|_2 \leq \rho_{opt}^k \|r_{minres,0}\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)}$

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$

$\square$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k,r_k)}{(Ar_k,Ar_k)}$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k,r_k)}{(Ar_k,Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$

# Method of minimal residuals, convergence, 3

Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)}$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k,r_k)}{(Ar_k,Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k A r_k$

# Method of minimal residuals, convergence, 3

Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k Ar_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k Ar_k$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k A r_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k A r_k$
- $r_{minres,k+1} = (I - \alpha_{minres,k} A) r_{minres,k} = B_{R_{minres,k}} r_{minres,k}$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k A r_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k A r_k$
- $r_{minres,k+1} = (I - \alpha_{minres,k}A) r_{minres,k} = B_{R_{minres,k}} r_{minres,k}$
- $r_{opt,k+1} = (I - \alpha_{opt}A) r_{opt,k} = B_{R_{opt,k}} r_{opt,k} \equiv B_{R_{opt,k}} r_k$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k Ar_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k Ar_k$
- $r_{minres,k+1} = (I - \alpha_{minres,k}A)r_{minres,k} = B_{R_{minres,k}} r_{minres,k}$
- $r_{opt,k+1} = (I - \alpha_{opt}A)r_{opt,k} = B_{R_{opt,k}} r_{opt,k} \equiv B_{R_{opt,k}} r_k$
- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k A r_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k A r_k$
- $r_{minres,k+1} = (I - \alpha_{minres,k} A) r_{minres,k} = B_{R_{minres,k}} r_{minres,k}$
- $r_{opt,k+1} = (I - \alpha_{opt} A) r_{opt,k} = B_{R_{opt,k}} r_{opt,k} \equiv B_{R_{opt,k}} r_k$
- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k, r_k)}{(Ar_k, Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k Ar_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k Ar_k$
- $r_{minres,k+1} = (I - \alpha_{minres,k}A)r_{minres,k} = B_{R_{minres,k}} r_{minres,k}$
- $r_{opt,k+1} = (I - \alpha_{opt}A)r_{opt,k} = B_{R_{opt,k}} r_{opt,k} \equiv B_{R_{opt,k}} r_k$
- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt} \|r_k\|_2$

# Method of minimal residuals, convergence, 3

## Proof.

- Richardson: $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} + Ax^{(k)} = b$
- **Minimal residuals** $\alpha_k = \alpha_{minres,k} = \frac{(Ar_k,r_k)}{(Ar_k,Ar_k)}$
- **Richardson with optimal step** $\alpha_k = \alpha_{opt} = \frac{2}{\lambda_1(A)+\lambda_n(A)}$
- Approach: $\alpha_{minres,k}$ vs $\alpha_{opt}$ for the same $r_k, x_k$
- $\frac{x^{(k+1)}-x^{(k)}}{\alpha_k} = b - Ax^{(k)} \Rightarrow x^{(k+1)} - x^{(k)} = \alpha_k r_k \Rightarrow$
  $Ax^{(k+1)} - Ax^{(k)} = \alpha_k A r_k \Rightarrow -(b - Ax^{(k+1)}) + (b - Ax^{(k)}) = \alpha_k A r_k$
- $r_{minres,k+1} = (I - \alpha_{minres,k}A) r_{minres,k} = B_{R_{minres,k}} r_{minres,k}$
- $r_{opt,k+1} = (I - \alpha_{opt}A) r_{opt,k} = B_{R_{opt,k}} r_{opt,k} \equiv B_{R_{opt,k}} r_k$
- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt}\|r_k\|_2$
- $\|r_{minres,k+1}\| \leq \rho_{opt}\|r_{minres,k}\|_2$

# Method of minimal residuals, convergence, 4

Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$

# Method of minimal residuals, convergence, 4

## Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)} < 1$

# Method of minimal residuals, convergence, 4

**Proof.**

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt} \|r_k\|_2$

# Method of minimal residuals, convergence, 4

## Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt}\|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}\|r_{minres,k}\|_2$

# Method of minimal residuals, convergence, 4

## Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \le \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \le \|r_{opt,k+1}\|_2 \le \rho_{opt} \|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \le \rho_{opt} \|r_{minres,k}\|_2$
- $\|r_{minres,k+1}\|_2 \le \rho_{opt}^2 \|r_{minres,k-1}\|_2$

# Method of minimal residuals, convergence, 4

Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt} \|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt} \|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt} \|r_{minres,k}\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}^2 \|r_{minres,k-1}\|_2$
- ...
- $\|r_{minres,k+1}\| \leq \rho_{opt}^{k+1} \|r_{minres,0}\|_2$

# Method of minimal residuals, convergence, 4

**Proof.**

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt}\|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}\|r_{minres,k}\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}^2\|r_{minres,k-1}\|_2$
- ...
- $\|r_{minres,k+1}\| \leq \rho_{opt}^{k+1}\|r_{minres,0}\|_2$
- $\rho_{opt} < 1 \Rightarrow lim_{k\to\infty}\|r_{minres,k}\|_2 = 0$

# Method of minimal residuals, convergence, 4

## Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A) - \lambda_n(A)}{\lambda_1(A) + \lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt}\|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}\|r_{minres,k}\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}^2\|r_{minres,k-1}\|_2$
- ...
- $\|r_{minres,k+1}\| \leq \rho_{opt}^{k+1}\|r_{minres,0}\|_2$
- $\rho_{opt} < 1 \Rightarrow lim_{k\to\infty}\|r_{minres,k}\|_2 = 0$
- $e^{(k)} = x - x^{(k)} = A^{-1}b - x^{(k)} = A^{-1}(b - Ax^{(k)}) = A^{-1}r_k$

# Method of minimal residuals, convergence, 4

## Proof.

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt}\|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}\|r_{minres,k}\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}^2\|r_{minres,k-1}\|_2$
- ...
- $\|r_{minres,k+1}\| \leq \rho_{opt}^{k+1}\|r_{minres,0}\|_2$
- $\rho_{opt} < 1 \Rightarrow lim_{k\to\infty}\|r_{minres,k}\|_2 = 0$
- $e^{(k)} = x - x^{(k)} = A^{-1}b - x^{(k)} = A^{-1}(b - Ax^{(k)}) = A^{-1}r_k$
- $\|e_{minres}^{(k)}\|_2 \leq \|A^{-1}\|_2\|r_{minres,k}\|_2 \to^{k\to\infty} 0$

# Method of minimal residuals, convergence, 4

**Proof.**

- $\|r_{opt,k+1}\|_2 = \|B_{R_{opt,k}} r_{opt,k}\|_2 \leq \|B_{R_{opt,k}}\|_2 \|r_{opt,k}\|_2 = \rho_{opt}\|r_k\|_2$
- $\rho_{opt} = \frac{\lambda_1(A)-\lambda_n(A)}{\lambda_1(A)+\lambda_n(A)} < 1$
- $\|r_{minres,k+1}\|_2 \leq \|r_{opt,k+1}\|_2 \leq \rho_{opt}\|r_k\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}\|r_{minres,k}\|_2$
- $\|r_{minres,k+1}\|_2 \leq \rho_{opt}^2\|r_{minres,k-1}\|_2$
- ...
- $\|r_{minres,k+1}\| \leq \rho_{opt}^{k+1}\|r_{minres,0}\|_2$
- $\rho_{opt} < 1 \Rightarrow lim_{k\to\infty}\|r_{minres,k}\|_2 = 0$
- $e^{(k)} = x - x^{(k)} = A^{-1}b - x^{(k)} = A^{-1}(b - Ax^{(k)}) = A^{-1}r_k$
- $\|e_{minres}^{(k)}\|_2 \leq \|A^{-1}\|_2\|r_{minres,k}\|_2 \to^{k\to\infty} 0$

$\square$

Q & A