§7.3 The Jacobi and Gauss-Siedel Iterative Techniques

- **Problem:** To solve $A\mathbf{x} = \mathbf{b}$ for $A \in \mathbb{R}^{n\times n}$.
- **Methodology:** Iteratively approximate solution $\mathbf{x}$. No GEPP.

§7.3 The Jacobi and Gauss-Siedel Iterative Techniques

- **Problem:** To solve $A\mathbf{x} = \mathbf{b}$ for $A \in \mathbb{R}^{n\times n}$.
- **Methodology:** Iteratively approximate solution $\mathbf{x}$. No GEPP.

Matrix splitting

$$A = \mathbf{diag}\left(a_{1,1}, a_{2,2}, \cdots, a_{n,n}\right) + \begin{pmatrix} 0 & & & & \\ a_{2,1} & 0 & & & \\ \vdots & \vdots & \ddots & & \\ a_{n-1,1} & a_{n-1,2} & \cdots & 0 & \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1} & 0 \end{pmatrix}$$

$$+ \begin{pmatrix} 0 & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ & 0 & \cdots & a_{2,n-1} & a_{2,n} \\ & & \ddots & \vdots & \vdots \\ & & & 0 & a_{n-1,n} \\ & & & & 0 \end{pmatrix}$$

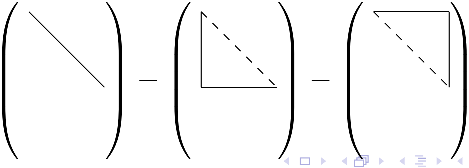
§7.3 The Jacobi and Gauss-Siedel Iterative Techniques

- **Problem:** To solve $A\mathbf{x} = \mathbf{b}$ for $A \in \mathbb{R}^{n \times n}$.
- **Methodology:** Iteratively approximate solution $\mathbf{x}$. No GEPP.

MATRIX SPLITTING

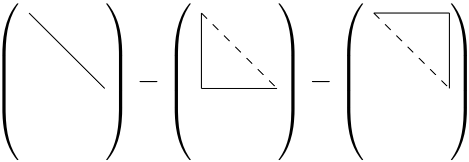
$$
\begin{aligned}
A &= \mathbf{diag}\left(a_{1,1}, a_{2,2}, \cdots, a_{n,n}\right) + \begin{pmatrix} 0 & & & & \\ a_{2,1} & 0 & & & \\ \vdots & \vdots & \ddots & & \\ a_{n-1,1} & a_{n-1,2} & \cdots & 0 & \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1} & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} 0 & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ & 0 & \cdots & a_{2,n-1} & a_{2,n} \\ & & \ddots & \vdots & \vdots \\ & & & 0 & a_{n-1,n} \\ & & & & 0 \end{pmatrix}
\end{aligned}
$$

$$\stackrel{def}{=} \quad D - L - U =$$



.

Ex: Matrix splitting for $A = \begin{pmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{pmatrix}$

$A =$ 

$$= \textbf{diag}\,(10, 11, 10, 8) - \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ -2 & 1 & 0 & \\ 0 & -3 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -2 & 0 \\ & 0 & 1 & -3 \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}$$

The Jacobi and Gauss-Siedel Methods for solving $A\mathbf{x} = \mathbf{b}$

JACOBI METHOD: With matrix splitting $A = D - L - U$, rewrite

$$\mathbf{x} = D^{-1}(L + U)\mathbf{x} + D^{-1}\mathbf{b}.$$

Jacobi iteration with given $\mathbf{x}^{(0)}$,

$$\mathbf{x}^{(k+1)} = D^{-1}(L + U)\mathbf{x}^{(k)} + D^{-1}\mathbf{b}, \quad \text{for} \quad k = 0, 1, 2, \cdots .$$

The Jacobi and Gauss-Siedel Methods for solving $A\mathbf{x} = \mathbf{b}$

JACOBI METHOD: With matrix splitting $A = D - L - U$, rewrite

$$\mathbf{x} = D^{-1}(L + U)\,\mathbf{x} + D^{-1}\,\mathbf{b}.$$

Jacobi iteration with given $\mathbf{x}^{(0)}$,

$$\mathbf{x}^{(k+1)} = D^{-1}(L + U)\,\mathbf{x}^{(k)} + D^{-1}\,\mathbf{b}, \quad \text{for} \quad k = 0, 1, 2, \cdots .$$

GAUSS-SIEDEL METHOD: Rewrite

$$\mathbf{x} = (D - L)^{-1}\; U\,\mathbf{x} + (D - L)^{-1}\;\mathbf{b}.$$

Gauss-Siedel iteration with given $\mathbf{x}^{(0)}$,

$$\mathbf{x}^{(k+1)} = (D - L)^{-1}\; U\,\mathbf{x}^{(k)} + (D - L)^{-1}\;\mathbf{b}, \quad \text{for} \quad k = 0, 1, 2, \cdots .$$

Ex: Jacobi Method for $A\mathbf{x} = \mathbf{b}$, with

$$A = \begin{pmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ 25 \\ -11 \\ 15 \end{pmatrix}$$

$$\begin{aligned} A &= D - L - U \\ &= \mathbf{diag}\,(10, 11, 10, 8) - \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ -2 & 1 & 0 & \\ 0 & -3 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -2 & 0 \\ & 0 & 1 & -3 \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \end{aligned}$$

Ex: Jacobi Method for $A\mathbf{x} = \mathbf{b}$, with

$$A = \begin{pmatrix} 10 & -1 & 2 & 0 \\ -1 & 11 & -1 & 3 \\ 2 & -1 & 10 & -1 \\ 0 & 3 & -1 & 8 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ 25 \\ -11 \\ 15 \end{pmatrix}$$

$$\begin{aligned} A &= D - L - U \\ &= \mathbf{diag}\,(10, 11, 10, 8) - \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ -2 & 1 & 0 & \\ 0 & -3 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -2 & 0 \\ & 0 & 1 & -3 \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \end{aligned}$$

Jacobi iteration with $\mathbf{x}^{(0)} = \mathbf{0}$, for $k = 0, 1, 2, \cdots$

$$\begin{aligned} \mathbf{x}_{\mathbf{J}}^{(k+1)} &= D^{-1}(L + U)\,\mathbf{x}_{\mathbf{J}}^{(k)} + D^{-1}\,\mathbf{b} \\ &= \begin{pmatrix} 0 & \frac{1}{10} & -\frac{2}{10} & 0 \\ \frac{1}{11} & 0 & \frac{1}{11} & -\frac{3}{11} \\ -\frac{2}{10} & \frac{1}{10} & 0 & \frac{1}{10} \\ 0 & -\frac{3}{8} & \frac{1}{8} & 0 \end{pmatrix} \mathbf{x}_{\mathbf{J}}^{(k)} + \begin{pmatrix} \frac{6}{10} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{pmatrix} \end{aligned}$$

Ex: Gauss-Siedel Method for $A\mathbf{x} = \mathbf{b}$

$$
\begin{aligned}
A &= D - L - U \\
&= \begin{pmatrix} 10 & & & \\ -1 & 11 & & \\ 2 & -1 & 10 & \\ 0 & 3 & -1 & 8 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -2 & 0 \\ & 0 & 1 & -3 \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}.
\end{aligned}
$$

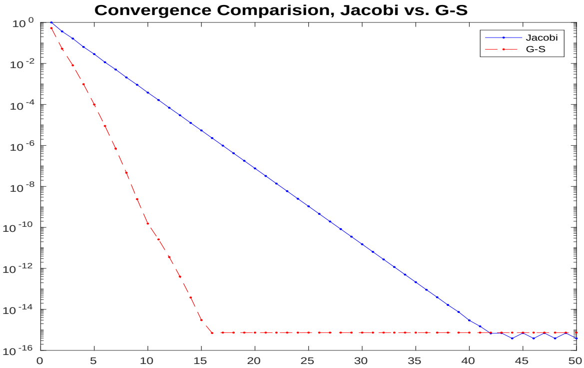
Ex: Gauss-Siedel Method for $A\mathbf{x} = \mathbf{b}$

$$\begin{aligned} A &= D - L - U \\ &= \begin{pmatrix} 10 & & & \\ -1 & 11 & & \\ 2 & -1 & 10 & \\ 0 & 3 & -1 & 8 \end{pmatrix} - \begin{pmatrix} 0 & 1 & -2 & 0 \\ & 0 & 1 & -3 \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}. \end{aligned}$$

Gauss-Siedel iteration with $\mathbf{x}^{(0)} = \mathbf{0}$, for $k = 0, 1, 2, \cdots$

$$\begin{aligned} \mathbf{x}_{\mathbf{GS}}^{(k+1)} &= (D - L)^{-1} \, U \, \mathbf{x}_{\mathbf{GS}} + (D - L)^{-1} \, \mathbf{b} \\ &= \begin{pmatrix} 10 & & & \\ -1 & 11 & & \\ 2 & -1 & 10 & \\ 0 & 3 & -1 & 8 \end{pmatrix}^{-1} \left(\begin{pmatrix} 0 & 1 & -2 & 0 \\ & 0 & 1 & -3 \\ & & 0 & 1 \\ & & & 0 \end{pmatrix} \mathbf{x}_{\mathbf{GS}}^{(k)} \right) \\ &\quad + \begin{pmatrix} \frac{6}{10} \\ \frac{25}{11} \\ -\frac{11}{10} \\ \frac{15}{8} \end{pmatrix}. \end{aligned}$$

Jacobi vs. Gauss-Siedel: **solution** $x = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}$.



General Iteration Methods

To solve $A\mathbf{x} = \mathbf{b}$ with matrix splitting $A = D - L - U$,

- Jacobi Method:

$$\mathbf{x}_{\mathbf{J}}^{(k+1)} = D^{-1}(L+U)\,\mathbf{x}_{\mathbf{J}}^{(k)} + D^{-1}\,\mathbf{b}.$$

- Gauss-Siedel Method:

$$\mathbf{x}_{\mathbf{GS}}^{(k+1)} = (D-L)^{-1}\;U\,\mathbf{x}_{\mathbf{GS}}^{(k)} + (D-L)^{-1}\;\mathbf{b}.$$

General Iteration Method: for $k = 0, 1, 2, \cdots$

$$\mathbf{x}^{(k+1)} = T\,\mathbf{x}^{(k)} + \mathbf{c}.$$

Next: convergence analysis on General Iteration Method

General Iteration: $\mathbf{x}^{(k+1)} = T\,\mathbf{x}^{(k)} + \mathbf{c}$ for $k = 0, 1, 2, \cdots$

Thm: The following statements are equivalent

- $\rho(T) < 1$.
- The equation
$$\mathbf{x} = T\,\mathbf{x} + \mathbf{c} \qquad (1)$$
has a unique solution and $\{\mathbf{x}^{(k)}\}$ converges to this solution from any $\mathbf{x}^{(0)}$.

General Iteration: $\mathbf{x}^{(k+1)} = T\,\mathbf{x}^{(k)} + \mathbf{c}$ for $k = 0, 1, 2, \cdots$

Thm: The following statements are equivalent

- $\rho(T) < 1$.
- The equation
$$\mathbf{x} = T\,\mathbf{x} + \mathbf{c} \qquad (1)$$
has a unique solution and $\{\mathbf{x}^{(k)}\}$ converges to this solution from any $\mathbf{x}^{(0)}$.

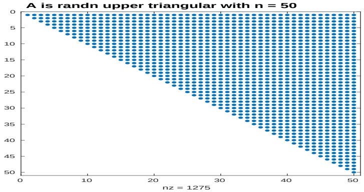
Proof: Assume $\rho(T) < 1$. Then (1) has unique solution $\mathbf{x}^{(*)}$.

$$\begin{aligned}
\mathbf{x}^{(k+1)} - \mathbf{x}^{(*)} &= T\left(\mathbf{x}^{(k)} - \mathbf{x}^{(*)}\right) = T^2\left(\mathbf{x}^{(k-1)} - \mathbf{x}^{(*)}\right) \\
&= \cdots = T^{k+1}\left(\mathbf{x}^{(0)} - \mathbf{x}^{(*)}\right) \Longrightarrow \mathbf{0}.
\end{aligned}$$

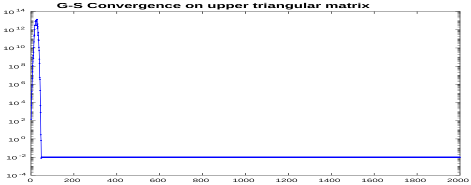
Conversely, if $\cdots$ (omitted)

Jacobi on random upper triangular matrix

- $A = D - U$. $T = D^{-1}U$ with $\rho(T) = 0$.



- Convergence plot



§7.4 Relaxation Techniques for Solving Linear Systems

To solve $A\mathbf{x} = \mathbf{b}$ with matrix splitting $A = D - L - U$, rewrite

$$\begin{aligned} D\,\mathbf{x} &= D\,\mathbf{x}, \\ \omega\,L\,\mathbf{x} &= \omega\,(D - U)\,\mathbf{x} - \omega\,\mathbf{b}, \quad \text{for any } \omega. \end{aligned}$$

§7.4 Relaxation Techniques for Solving Linear Systems

To solve $A\,\mathbf{x} = \mathbf{b}$ with matrix splitting $A = D - L - U$, rewrite

$$\begin{aligned} D\,\mathbf{x} &= D\,\mathbf{x}, \\ \omega\,L\,\mathbf{x} &= \omega\,(D - U)\,\mathbf{x} - \omega\,\mathbf{b}, \quad \text{for any } \omega. \end{aligned}$$

Taking difference of two equations,

$$(D - \omega\,L)\,\mathbf{x} = ((1-\omega)\,D + \omega\,U)\,\mathbf{x} + \omega\,\mathbf{b}.$$

§7.4 Relaxation Techniques for Solving Linear Systems

To solve $A\mathbf{x} = \mathbf{b}$ with matrix splitting $A = D - L - U$, rewrite

$$\begin{aligned} D\mathbf{x} &= D\mathbf{x}, \\ \omega L\mathbf{x} &= \omega (D - U)\mathbf{x} - \omega \mathbf{b}, \quad \text{for any } \omega. \end{aligned}$$

Taking difference of two equations,

$$(D - \omega L)\mathbf{x} = ((1-\omega) D + \omega U)\mathbf{x} + \omega \mathbf{b}.$$

Successive Over-Relaxation (SOR), for $k = 0, 1, 2, \cdots$

$$\begin{aligned} \mathbf{x}_{\mathbf{SOR}}^{(k+1)} &= (D - \omega L)^{-1}((1-\omega) D + \omega U)\mathbf{x}_{\mathbf{SOR}}^{(k)} + \omega (D - \omega L)^{-1}\mathbf{b} \\ &\stackrel{def}{=} T_{\mathbf{SOR}}\mathbf{x}_{\mathbf{SOR}}^{(k)} + \mathbf{c}_{\mathbf{SOR}}. \end{aligned}$$

§7.4 Relaxation Techniques for Solving Linear Systems

To solve $A\mathbf{x} = \mathbf{b}$ with matrix splitting $A = D - L - U$, rewrite

$$\begin{aligned} D\,\mathbf{x} &= D\,\mathbf{x}, \\ \omega\,L\,\mathbf{x} &= \omega\,(D - U)\,\mathbf{x} - \omega\,\mathbf{b}, \quad \text{for any } \omega. \end{aligned}$$

Taking difference of two equations,

$$(D - \omega\,L)\,\mathbf{x} = ((1-\omega)\,D + \omega\,U)\,\mathbf{x} + \omega\,\mathbf{b}.$$

Successive Over-Relaxation (SOR), for $k = 0, 1, 2, \cdots$

$$\begin{aligned} \mathbf{x}_{\mathbf{SOR}}^{(k+1)} &= (D - \omega\,L)^{-1}\left((1-\omega)\,D + \omega\,U\right)\mathbf{x}_{\mathbf{SOR}}^{(k)} + \omega\,(D - \omega\,L)^{-1}\,\mathbf{b} \\ &\stackrel{def}{=} T_{\mathbf{SOR}}\,\mathbf{x}_{\mathbf{SOR}}^{(k)} + \mathbf{c}_{\mathbf{SOR}}. \end{aligned}$$

converges if $\rho\left(T_{\mathbf{SOR}}\right) < 1$.

Good choice of ω is tricky, but critical for accelerated convergence

Optimal **SOR** parameters

Thm: If A is symmetric positive definite and tridiagonal, then

$$\rho\left(T_{GS}\right) = \left(\rho\left(T_J\right)\right)^2 < 1,$$

and the optimal choice of ω for the **SOR** method is

$$\omega_{\mathbf{OPT}} = \frac{2}{1 + \sqrt{1 - \left(\rho\left(T_J\right)\right)^2}}$$

with

$$\rho\left(T_{\mathbf{SOR}}\right) = \omega_{\mathbf{OPT}} - 1 = \left(\frac{\rho\left(T_J\right)}{1 + \sqrt{1 - \left(\rho\left(T_J\right)\right)^2}}\right)^2. \qquad \square$$

$$A = \begin{pmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{x} = \tfrac{1}{3} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

- If A *is* symmetric positive definite and tridiagonal,

$$\mathbf{det}\,(A) = 24, \quad \mathbf{det} \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} = 7, \quad 4 > 0.$$

$$A = \begin{pmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{x} = \frac{1}{3}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

- If A *is* symmetric positive definite and tridiagonal,

$$\mathbf{det}\,(A) = 24, \quad \mathbf{det}\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} = 7, \quad 4 > 0.$$

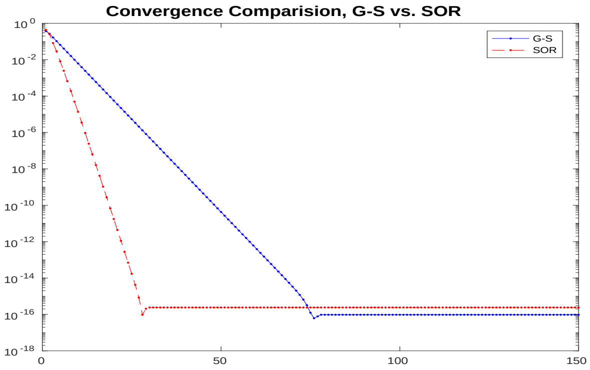
- $\rho(T_{\mathbf{J}}) = \sqrt{0.625}$

$$T_{\mathbf{J}} = D^{-1}(L+U) = \frac{1}{4}\begin{pmatrix} 0 & -3 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

- Optimal ω:

$$\omega_{\mathbf{OPT}} = \frac{2}{1+\sqrt{1-(\rho(T_J))^2}} = \omega_{\mathbf{OPT}} = \frac{2}{1+\sqrt{0.375}} \approx 1.24.$$

$$A = \begin{pmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x} = \frac{1}{3}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$



§7.5 Error Bounds and Iterative Refinement

Assume that $\widehat{\mathbf{x}}$ is an approximation to the solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$.

- RESIDUAL $\widehat{\mathbf{r}} \stackrel{def}{=} \mathbf{b} - A\widehat{\mathbf{x}} = A\left(\mathbf{x} - \widehat{\mathbf{x}}\right)$. Thus small $\|\mathbf{x} - \widehat{\mathbf{x}}\|$ implies small $\|\widehat{\mathbf{r}}\|$.

§7.5 Error Bounds and Iterative Refinement

Assume that $\widehat{\mathbf{x}}$ is an approximation to the solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$.

- RESIDUAL $\widehat{\mathbf{r}} \stackrel{def}{=} \mathbf{b} - A\widehat{\mathbf{x}} = A(\mathbf{x} - \widehat{\mathbf{x}})$. Thus small $\|\mathbf{x} - \widehat{\mathbf{x}}\|$ implies small $\|\widehat{\mathbf{r}}\|$.
- However, big $\|\mathbf{x} - \widehat{\mathbf{x}}\|$ *can* still lead to small $\|\widehat{\mathbf{r}}\|$. **Ex:**

$$\begin{pmatrix} 1 & 2 \\ 1 + 10^{-\tau} & 2 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 3 \\ 3 + 10^{-\tau} \end{pmatrix}.$$

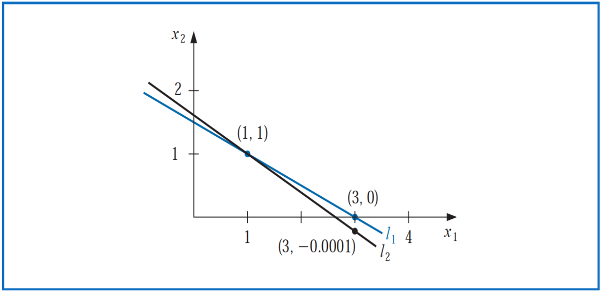
Exact solution $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Bad approximation $\widehat{\mathbf{x}} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ has a small residual for large τ:

$$\widehat{\mathbf{r}} = \begin{pmatrix} 3 \\ 3 + 10^{-\tau} \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 1 + 10^{-\tau} & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \times 10^{-\tau} \end{pmatrix}.$$

Near Linear Dependence

For $\tau = 4$, equations define two *nearly* parallel lines

$$\ell_1 : \quad x_1 + 2\,x_2 = 3, \quad \text{and} \quad \ell_2 : \quad 1.0001\,x_1 + 2\,x_2 = 3.0001.$$



Parallel lines do not have intersections.

Let $A\mathbf{x} = \mathbf{b}$ with non-singular A and non-zero $\mathbf{b}$

Thm: Assume $\widehat{\mathbf{x}}$ is an approximate solution with $\widehat{\mathbf{r}} = b - A\widehat{\mathbf{x}}$. Then for any natural norm,

$$\begin{aligned} \|\widehat{\mathbf{x}} - \mathbf{x}\| &\leq \left\|A^{-1}\right\| \|\widehat{\mathbf{r}}\|, \\ \frac{\|\widehat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} &\leq \kappa(A) \frac{\|\widehat{\mathbf{r}}\|}{\|\mathbf{b}\|}, \end{aligned}$$

where $\kappa(A) \stackrel{def}{=} \|A\| \left\|A^{-1}\right\|$ is the CONDITION NUMBER of A $\square$

Let $A\mathbf{x} = \mathbf{b}$ with non-singular A and non-zero $\mathbf{b}$

Thm: Assume $\widehat{\mathbf{x}}$ is an approximate solution with $\widehat{\mathbf{r}} = b - A\widehat{\mathbf{x}}$. Then for any natural norm,

$$\begin{aligned} \|\widehat{\mathbf{x}} - \mathbf{x}\| &\leq \left\|A^{-1}\right\| \, \|\widehat{\mathbf{r}}\|, \\ \frac{\|\widehat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} &\leq \kappa(A) \, \frac{\|\widehat{\mathbf{r}}\|}{\|\mathbf{b}\|}, \end{aligned}$$

where $\kappa(A) \stackrel{def}{=} \|A\| \left\|A^{-1}\right\|$ is the CONDITION NUMBER of A □

- A is WELL-CONDITIONED if $\kappa(A) = O(1)$:
  small residual implies small solution error.
- A is ILL-CONDITIONED if $\kappa(A) \gg 1$:
  small residual may still allow large solution error.

Ex: Condition Number for $A = \begin{pmatrix} 1 & 2 \\ 1 + 10^{-\tau} & 2 \end{pmatrix}$

Solution: For $\tau > 0$, $\|A\|_\infty = 3 + 10^{-\tau}$. Since

$$A^{-1} = -\frac{10^\tau}{2} \begin{pmatrix} 2 & -2 \\ -(1 + 10^{-\tau}) & 1 \end{pmatrix},$$

we have $\left\|A^{-1}\right\|_\infty = 2 \times 10^\tau$. Thus

$$\kappa(A) = \|A\|_\infty \left\|A^{-1}\right\|_\infty = 6 \times 10^\tau + 2.$$

$\kappa(A)$ grows exponentially in τ. A is ill-conditioned for large τ.

Iterative Refinement (I)

- Let $A\mathbf{x} = \mathbf{b}$ with non-singular A and non-zero $\mathbf{b}$.
- Let $\mathcal{F}(\cdot)$ be an in-exact equation solver, so $\mathcal{F}(\mathbf{b})$ is approximate solution.
- Assume $\mathcal{F}(\cdot)$ is accurate enough that there exists a $\rho < 1$ so

$$\frac{\|\mathbf{b} - A\mathcal{F}(\mathbf{b})\|}{\|\mathbf{b}\|} \le \rho \quad \text{for any } \mathbf{b} \neq \mathbf{0}.$$

Iterative Refinement (I)

- Let $A\mathbf{x} = \mathbf{b}$ with non-singular A and non-zero $\mathbf{b}$.
- Let $\mathcal{F}(\cdot)$ be an in-exact equation solver, so $\mathcal{F}(\mathbf{b})$ is approximate solution.
- Assume $\mathcal{F}(\cdot)$ is accurate enough that there exists a $\rho < 1$ so

$$\frac{\|\mathbf{b} - A\,\mathcal{F}(\mathbf{b})\|}{\|\mathbf{b}\|} \leq \rho \quad \text{for any } \mathbf{b} \neq \mathbf{0}.$$

In practice,

- $\mathcal{F}(\cdot)$ could be from an (in-exact) LU factorization,

$$\mathcal{F}(\mathbf{b}) = U^{-1}\left(L^{-1}\,\mathbf{b}\right).$$

- Inaccuracies in LU factorization could be due to rounding-error,

$$A \approx L\,U.$$

Ex: $A = \texttt{randn}(n, n)$, $\mathbf{b} = \texttt{randn}(n, 1)$, $n = 3000$

- **LU factorize** A to get L, U, (LU without pivoting)
- $\mathbf{x}_0 = U^{-1}\left(L^{-1}\,\mathbf{b}\right)$,
- $\mathbf{r}^{(0)} = \mathbf{b} - A\,\mathbf{x}_0$,
- $\Delta\mathbf{x}_1 = U^{-1}\left(L^{-1}\,\mathbf{r}_0\right)$,
- $\mathbf{r}_1 = \mathbf{r}_0 - A\,\Delta\mathbf{x}_1$,
- $\mathbf{x} = \mathbf{x}_0 + \Delta\mathbf{x}_1$
- $\texttt{disp}\left(\texttt{norm}(\mathbf{r}_0),\ \texttt{norm}(\mathbf{r}_1)\right)$

2.6606e-07 1.0996e-16

Iterative Refinement (II)

Given a tolerance $\tau > 0$ and $\mathbf{x}^{(0)}$

- Initialize $\mathbf{r}^{(0)} = \mathbf{b} - A\,\mathbf{x}^{(0)}$.
- for $k = 0, 1, \cdots$
  - Compute

$$\begin{aligned}
\Delta\mathbf{x}^{(k)} &= \mathcal{F}\left(\mathbf{r}^{(k)}\right), \\
\mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}, \\
\mathbf{r}^{(k+1)} &= \mathbf{r}^{(k)} - A\,\Delta\mathbf{x}^{(k)}.
\end{aligned}$$

  - If $\left\|\mathbf{r}^{(k+1)}\right\| \leq \tau \left\|\mathbf{b}\right\|$ **stop**.

Iterative Refinement (II)

Given a tolerance $\tau > 0$ and $\mathbf{x}^{(0)}$

- Initialize $\mathbf{r}^{(0)} = \mathbf{b} - A\,\mathbf{x}^{(0)}$.
- for $k = 0, 1, \cdots$
  - Compute

$$\begin{aligned}
\Delta\mathbf{x}^{(k)} &= \mathcal{F}\left(\mathbf{r}^{(k)}\right), \\
\mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}, \\
\mathbf{r}^{(k+1)} &= \mathbf{r}^{(k)} - A\,\Delta\mathbf{x}^{(k)}.
\end{aligned}$$

  - If $\left\|\mathbf{r}^{(k+1)}\right\| \leq \tau\,\|\mathbf{b}\|$ **stop**.

Convergence Proof:

$$\left\|\mathbf{r}^{(k+1)}\right\| \leq \rho\,\left\|\mathbf{r}^{(k)}\right\| \leq \rho^2\,\left\|\mathbf{r}^{(k-1)}\right\| \leq \cdots \leq \rho^{k+1}\,\left\|\mathbf{r}^{(0)}\right\|.$$

Perturbation Theory

Thm: Let $\mathbf{x}$ and $\widehat{\mathbf{x}}$ be solutions to

$$A\,\mathbf{x} = \mathbf{b} \quad \text{and} \quad (A + \Delta A)\,\widehat{\mathbf{x}} = \mathbf{b} + \Delta\mathbf{b}$$

with PERTURBATIONS ΔA and $\Delta\mathbf{b}$. Then

$$\frac{\|\widehat{\mathbf{x}} - \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\kappa\,(A)}{1 - \kappa\,(A)\,\frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \right).$$

with $\kappa\,(A) = \|A\|\,\left\|A^{-1}\right\|$ $\square$.

§7.6 The Conjugate Gradient Method (CG) for $A\mathbf{x} = \mathbf{b}$

Assumption: A is SYMMETRIC POSITIVE DEFINITE (SPD)

- $A^T = A$,
- $\mathbf{x}^T A\mathbf{x} \geq 0$ for any $\mathbf{x}$,
- $\mathbf{x}^T A\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

§7.6 The Conjugate Gradient Method (CG) for $A\mathbf{x} = \mathbf{b}$

Assumption: A is SYMMETRIC POSITIVE DEFINITE (SPD)

- $A^T = A$,
- $\mathbf{x}^T A\mathbf{x} \geq 0$ for any $\mathbf{x}$,
- $\mathbf{x}^T A\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

Thm: The vector $\mathbf{x}^*$ solves the SPD equations $A\mathbf{x} = \mathbf{b}$ if and only if it minimizes function

$$g(\mathbf{x}) \stackrel{def}{=} \mathbf{x}^T A\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}.$$

§7.6 The Conjugate Gradient Method (CG) for $A\mathbf{x} = \mathbf{b}$

Assumption: A is SYMMETRIC POSITIVE DEFINITE (SPD)

- $A^T = A$,
- $\mathbf{x}^T A\mathbf{x} \geq 0$ for any $\mathbf{x}$,
- $\mathbf{x}^T A\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

Thm: The vector $\mathbf{x}^*$ solves the SPD equations $A\mathbf{x} = \mathbf{b}$ if and only if it minimizes function

$$g(\mathbf{x}) \stackrel{def}{=} \mathbf{x}^T A\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}.$$

Proof: Let $A\mathbf{x}^* = \mathbf{b}$. Then

$$\begin{aligned} g(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x} - 2\,\mathbf{x}^T A\,\mathbf{x}^* \\ &= (\mathbf{x} - \mathbf{x}^*)^T A\,(\mathbf{x} - \mathbf{x}^*) - (\mathbf{x}^*)^T A\,(\mathbf{x}^*) \\ &= (\mathbf{x} - \mathbf{x}^*)^T A\,(\mathbf{x} - \mathbf{x}^*) + g(\mathbf{x}^*)\,. \end{aligned}$$

Thus, $g(\mathbf{x}) \geq g(\mathbf{x}^*)$ for all $\mathbf{x}$; and $g(\mathbf{x}) = g(\mathbf{x}^*)$ iff $\mathbf{x} = \mathbf{x}^*$. □

CG for $A\mathbf{x} = \mathbf{b}$

Thm: The vector $\mathbf{x}^*$ solves the SPD equations $A\mathbf{x} = \mathbf{b}$ if and only if it minimizes function

$$g(\mathbf{x}) \stackrel{def}{=} \mathbf{x}^T A\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}.$$

The CG Idea: Starting from an initial vector $\mathbf{x}^{(0)}$, *quickly* compute new vectors $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(k)} \cdots$, with

$$g\left(\mathbf{x}^{(0)}\right) > g\left(\mathbf{x}^{(1)}\right) > g\left(\mathbf{x}^{(2)}\right) > \cdots > g\left(\mathbf{x}^{(k)}\right) > \cdots$$

so that the sequence $\{\mathbf{x}^{(k)}\}$ will converge to $\mathbf{x}^*$.

- **The CG Idea:** Starting from an initial vector $\mathbf{x}^{(0)}$, *quickly* compute new vectors $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(k)} \cdots$, with

$$g\left(\mathbf{x}^{(0)}\right) > g\left(\mathbf{x}^{(1)}\right) > g\left(\mathbf{x}^{(2)}\right) > \cdots > g\left(\mathbf{x}^{(k)}\right) > \cdots$$

  so that the sequence $\{\mathbf{x}^{(k)}\}$ will converge to $\mathbf{x}^*$.

SEARCH DIRECTION and LINE SEARCH

- **The CG Idea:** Starting from an initial vector $\mathbf{x}^{(0)}$, *quickly* compute new vectors $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(k)} \cdots$, with

$$g\left(\mathbf{x}^{(0)}\right) > g\left(\mathbf{x}^{(1)}\right) > g\left(\mathbf{x}^{(2)}\right) > \cdots > g\left(\mathbf{x}^{(k)}\right) > \cdots$$

so that the sequence $\{\mathbf{x}^{(k)}\}$ will converge to $\mathbf{x}^*$.

- **Descent Method:** Assume a SEARCH DIRECTION $\mathbf{v}^{(k)}$ at iteration $\mathbf{x}^{(k-1)}$, next iteration with STEP-SIZE t_k

$$\mathbf{x}^{(k)} \stackrel{def}{=} \mathbf{x}^{(k-1)} + t_k \mathbf{v}^{(k)} \quad \text{minimizes} \quad g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right).$$

SEARCH DIRECTION and LINE SEARCH

- **The CG Idea:** Starting from an initial vector $\mathbf{x}^{(0)}$, *quickly* compute new vectors $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(k)} \cdots$, with

$$g\left(\mathbf{x}^{(0)}\right) > g\left(\mathbf{x}^{(1)}\right) > g\left(\mathbf{x}^{(2)}\right) > \cdots > g\left(\mathbf{x}^{(k)}\right) > \cdots$$

  so that the sequence $\{\mathbf{x}^{(k)}\}$ will converge to $\mathbf{x}^*$.
- **Descent Method:** Assume a SEARCH DIRECTION $\mathbf{v}^{(k)}$ at iteration $\mathbf{x}^{(k-1)}$, next iteration with STEP-SIZE t_k

$$\mathbf{x}^{(k)} \stackrel{def}{=} \mathbf{x}^{(k-1)} + t_k \, \mathbf{v}^{(k)} \quad \text{minimizes} \quad g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right).$$

- **Optimality Condition:**

$$\begin{aligned} 0 &= \frac{d}{d\,t} g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) = \left(\mathbf{v}^{(k)}\right)^T \nabla g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) \\ &= \left(\mathbf{v}^{(k)}\right)^T \left(2A\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) - 2\,\mathbf{b}\right), \end{aligned}$$

SEARCH DIRECTION and LINE SEARCH

- **The CG Idea:** Starting from an initial vector $\mathbf{x}^{(0)}$, *quickly* compute new vectors $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(k)} \cdots$, with

$$g\left(\mathbf{x}^{(0)}\right) > g\left(\mathbf{x}^{(1)}\right) > g\left(\mathbf{x}^{(2)}\right) > \cdots > g\left(\mathbf{x}^{(k)}\right) > \cdots$$

  so that the sequence $\{\mathbf{x}^{(k)}\}$ will converge to $\mathbf{x}^*$.
- **Descent Method:** Assume a SEARCH DIRECTION $\mathbf{v}^{(k)}$ at iteration $\mathbf{x}^{(k-1)}$, next iteration with STEP-SIZE t_k

$$\mathbf{x}^{(k)} \stackrel{def}{=} \mathbf{x}^{(k-1)} + t_k \mathbf{v}^{(k)} \quad \text{minimizes} \quad g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right).$$

- **Optimality Condition:**

$$\begin{aligned} 0 &= \frac{d}{d\,t} g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) = \left(\mathbf{v}^{(k)}\right)^T \nabla g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) \\ &= \left(\mathbf{v}^{(k)}\right)^T \left(2A\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) - 2\,\mathbf{b}\right), \\ t_k &= \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}. \end{aligned}$$

SEARCH DIRECTION choices

For a small step-size t:

$$g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right) \approx g\left(\mathbf{x}^{(k-1)}\right) + t\left(\mathbf{v}^{(k)}\right)^T \nabla g\left(\mathbf{x}^{(k-1)}\right).$$

- STEEPEST DESCENT: Greatest decrease in the value of $g\left(\mathbf{x}^{(k-1)} + t\mathbf{v}^{(k)}\right)$:

$$\mathbf{v}^{(k)} = -\nabla g\left(\mathbf{x}^{(k-1)}\right).$$

- *A*-**orthogonal** DIRECTIONS: non-zero vectors $\{\mathbf{v}^{(i)}\}_{i=1}^n$

$$\left(\mathbf{v}^{(i)}\right)^T \left(A\,\mathbf{v}^{(j)}\right) = 0 \quad \text{for all } i \neq j.$$

  A-**orthogonal** vectors associated with the positive definite matrix A is linearly independent.

$A-$orthogonality Craft

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}_1 + \cdots + \tau_k\mathbf{v}_k\right) = g\left(\mathbf{x}_0 + t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k\right).$$

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}_1 + \cdots + \tau_k \mathbf{v}_k\right) = g\left(\mathbf{x}_0 + t_1 \mathbf{v}_1 + \cdots + t_k \mathbf{v}_k\right).$$

Magic (I): $\mathbf{min}_{\tau_1} g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}^{(1)}\right) = g\left(\mathbf{x}_0 + t_1 \mathbf{v}^{(1)}\right).$

$$\mathbf{min}_{\tau_1,\tau_2} g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}^{(1)} + \tau_2 \mathbf{v}^{(2)}\right) = g\left(\mathbf{x}_0 + t_1 \mathbf{v}^{(1)} + t_2 \mathbf{v}^{(2)}\right).$$

$$\mathbf{min}_{\mathbf{x}} g(\mathbf{x}) =$$

$$\mathbf{min}_{\tau_1,\cdots,\tau_n} g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}^{(1)} + \cdots + \tau_n \mathbf{v}^{(n)}\right) = g\left(\mathbf{x}_0 + t_1 \mathbf{v}^{(1)} + \cdots + t_n \mathbf{v}^{(n)}\right).$$

Thus $\mathbf{x} = \mathbf{x}_0 + t_1 \mathbf{v}^{(1)} + \cdots + t_n \mathbf{v}^{(n)}$ is solution to $A\mathbf{x} = \mathbf{b}$.

$A-$orthogonality Craft

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}_1 + \cdots + \tau_k\mathbf{v}_k\right) = g\left(\mathbf{x}_0 + t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k\right).$$

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}_1 + \cdots + \tau_k \mathbf{v}_k\right) = g\left(\mathbf{x}_0 + t_1 \mathbf{v}_1 + \cdots + t_k \mathbf{v}_k\right).$$

Proof (I): Let $\mathbf{t} = (\tau_1, \cdots, \tau_k)$. Then

$$\begin{aligned} & g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}_1 + \cdots + \tau_k \mathbf{v}_k\right) = g\left(\mathbf{x}_0\right) \\ & + \mathbf{t}^T \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T A \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right) \mathbf{t} - 2\mathbf{t}^T \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T \mathbf{r}^{(0)}, \\ & \nabla_{\mathbf{t}}\, g = 2\left(\left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T A \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right) \mathbf{t} - \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T \mathbf{r}^{(0)}\right) \end{aligned}$$

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}_1 + \cdots + \tau_k\mathbf{v}_k\right) = g\left(\mathbf{x}_0 + t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k\right).$$

Proof (I): Let $\mathbf{t} = (\tau_1, \cdots, \tau_k)$. Then

$$g\left(\mathbf{x}_0 + \tau_1\mathbf{v}_1 + \cdots + \tau_k\mathbf{v}_k\right) = g\left(\mathbf{x}_0\right)$$
$$+\mathbf{t}^T \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T A \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)\mathbf{t} - 2\mathbf{t}^T \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T \mathbf{r}^{(0)},$$
$$\nabla_{\mathbf{t}}\, g = 2\left(\left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T A \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)\mathbf{t} - \left(\mathbf{v}_1, \cdots, \mathbf{v}_k\right)^T \mathbf{r}^{(0)}\right)$$

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}_1 + \cdots + \tau_k\mathbf{v}_k\right) \iff \nabla_{\mathbf{t}}\, g = \mathbf{0}.$$

$A-$orthogonality Craft

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}^{(1)} + \cdots + \tau_k\mathbf{v}^{(k)}\right) = g\left(\mathbf{x}_0 + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}\right).$$

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}^{(1)} + \cdots + \tau_k\mathbf{v}^{(k)}\right) = g\left(\mathbf{x}_0 + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}\right).$$

Proof (II): Since vectors $\{\mathbf{v}^{(k)}\}$ are A-orthogonal

$$\nabla_{\mathbf{t}}\, g = 2\left(\mathbf{diag}\left(\left(\mathbf{v}^{(1)}\right)^T A\,\mathbf{v}^{(1)}, \cdots, \left(\mathbf{v}^{(k)}\right)^T A\,\mathbf{v}^{(k)}\right)\mathbf{t} - \left(\mathbf{v}^{(1)}, \cdots, \mathbf{v}^{(k)}\right)^T \mathbf{r}^{(0)}\right)$$

$$\nabla_{\mathbf{t}}\, g = \mathbf{0} \Longleftrightarrow \mathbf{t} = \begin{pmatrix} \frac{\left(\mathbf{v}^{(1)}\right)^T\left(\mathbf{r}^{(0)}\right)}{\left(\mathbf{v}^{(1)}\right)^T\left(A\,\mathbf{v}^{(1)}\right)} \\ \vdots \\ \frac{\left(\mathbf{v}^{(k)}\right)^T\left(\mathbf{r}^{(0)}\right)}{\left(\mathbf{v}^{(k)}\right)^T\left(A\,\mathbf{v}^{(k)}\right)} \end{pmatrix}.$$

$A-$orthogonality Craft

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g(\mathbf{x}) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T \mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1, \cdots, \tau_k} g\left(\mathbf{x}_0 + \tau_1 \mathbf{v}^{(1)} + \cdots + \tau_k \mathbf{v}^{(k)}\right) = g\left(\mathbf{x}_0 + t_1 \mathbf{v}^{(1)} + \cdots + t_k \mathbf{v}^{(k)}\right).$$

$A-$orthogonality Craft

Thm: Let non-zero vectors $\{\mathbf{v}^{(k)}\}$ be A-orthogonal with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{r}^{(k-1)} \stackrel{def}{=} \mathbf{b} - A\,\mathbf{x}^{(k-1)} \quad \text{(residual)}.$$

Then for $g\left(\mathbf{x}\right) = \mathbf{x}^T A\,\mathbf{x} - 2\,\mathbf{x}^T\mathbf{b}$ and for $k = 1, \cdots, n$,

$$\mathbf{min}_{\tau_1,\cdots,\tau_k} g\left(\mathbf{x}_0 + \tau_1\mathbf{v}^{(1)} + \cdots + \tau_k\mathbf{v}^{(k)}\right) = g\left(\mathbf{x}_0 + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}\right).$$

Proof (III): Since

$$\left(\mathbf{v}^{(k)}\right)^T\left(\mathbf{r}^{(k-1)}\right) = \left(\mathbf{v}^{(k)}\right)^T\left(\mathbf{r}^{(0)} - \sum_{j=1}^{k-1} t_j\,A\,\mathbf{v}^{(j)}\right) = \left(\mathbf{v}^{(k)}\right)^T\left(\mathbf{r}^{(0)}\right), \quad \text{so}$$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(0)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)} = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}.$$

$A-$orthogonality vectors (I)

$A-$orthogonality vectors (I)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)}\,\mathbf{v}^{(j)}.$$

Assume that $\{\mathbf{v}^{(k)}\}$ are non-zero. Then they are $A-$orthogonal.

$A-$orthogonality vectors (I)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \mathbf{v}^{(j)}.$$

Assume that $\{\mathbf{v}^{(k)}\}$ are non-zero. Then they are $A-$orthogonal.
Induction Proof: For all $1 \leq i < k$,

$$\begin{aligned} \left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(i)}\right) &= -\left(\mathbf{r}^{(k-1)}\right)^T \left(A\,\mathbf{v}^{(i)}\right) \\ &\quad + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(i)}\right) \\ &= -\left(\mathbf{r}^{(k-1)}\right)^T \left(A\,\mathbf{v}^{(i)}\right) + \left(\mathbf{v}^{(i)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right) = 0. \end{aligned}$$

$A-$orthogonality vectors (II)

$A-$orthogonality vectors (II)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(j)}\right)^T \left(\mathbf{r}^{(k)}\right) = 0, \quad j = 1, \cdots, k; \quad \left(\mathbf{r}^{(j)}\right)^T \left(\mathbf{r}^{(k)}\right) = 0, \;\; j = 1, \cdots, k-1.$$

$A-$orthogonality vectors (II)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)}\,\mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(j)}\right)^T \left(\mathbf{r}^{(k)}\right) = 0, \quad j = 1, \cdots, k; \quad \left(\mathbf{r}^{(j)}\right)^T \left(\mathbf{r}^{(k)}\right) = 0, \; j = 1, \cdots, k-1.$$

Proof: Due to optimality property of $\mathbf{x}^{(k)}$, for all τ and for $1 \leq j \leq k$,

$$\begin{aligned} g\left(\mathbf{x}^{(k)}\right) &\leq g\left(\mathbf{x}^{(k)} + \tau\,\mathbf{v}^{(j)}\right) \\ &= g\left(\mathbf{x}^{(k)}\right) - 2\tau\left(\mathbf{r}^{(k)}\right)^T \mathbf{v}^{(j)} + \tau^2 \left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right). \end{aligned}$$

This is true only when $\left(\mathbf{r}^{(k)}\right)^T \mathbf{v}^{(j)} = 0$.

$A-$orthogonality vectors (II)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(j)}\right)^T \left(\mathbf{r}^{(k)}\right) = 0, \quad j = 1, \cdots, k; \quad \left(\mathbf{r}^{(j)}\right)^T \left(\mathbf{r}^{(k)}\right) = 0, \ \ j = 1, \cdots, k-1.$$

Proof: Due to optimality property of $\mathbf{x}^{(k)}$, for all τ and for $1 \leq j \leq k$,

$$\begin{aligned} g\left(\mathbf{x}^{(k)}\right) &\leq g\left(\mathbf{x}^{(k)} + \tau\,\mathbf{v}^{(j)}\right) \\ &= g\left(\mathbf{x}^{(k)}\right) - 2\tau\left(\mathbf{r}^{(k)}\right)^T \mathbf{v}^{(j)} + \tau^2\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right). \end{aligned}$$

This is true only when $\left(\mathbf{r}^{(k)}\right)^T \mathbf{v}^{(j)} = 0$.

Residual vector orthogonality: $\mathbf{r}^{(j)} =$ linear combination of $\mathbf{v}^{(1)}, \cdots, \mathbf{v}^{(j+1)}$.

$A-$orthogonality vectors (III)

$A-$orthogonality vectors (III)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)}\right) = -\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right), \quad j = 1, \cdots, k-1.$$

$A-$orthogonality vectors (III)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)}\right) = -\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right), \quad j = 1, \cdots, k-1.$$

Proof (I): For $j = k - 1$,

$$\begin{aligned}
\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right) &= \left(-\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)} \mathbf{v}^{(j)}\right)^T \left(\mathbf{r}^{(k-1)}\right) \\
&= -\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right).
\end{aligned}$$

$A-$orthogonality vectors (III)

$A-$orthogonality vectors (III)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)}\,\mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)}\right) = -\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right), \quad j = 1, \cdots, k-1.$$

$A-$orthogonality vectors (III)

Thm: Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + \sum_{j=1}^{k-1} \frac{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(j)}\right)^T \left(A\,\mathbf{v}^{(j)}\right)}\,\mathbf{v}^{(j)}.$$

Let $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + t_1\mathbf{v}^{(1)} + \cdots + t_k\mathbf{v}^{(k)}$ and $\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{x}^{(k)}$. Then

$$\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)}\right) = -\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right), \quad j = 1, \cdots, k-1.$$

Proof (II): For $j < k-1$

$$\begin{aligned}
\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)}\right) &= \left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right) + \left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)} - \mathbf{r}^{(k-1)}\right) \\
&= \left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right) + \left(\mathbf{v}^{(k)}\right)^T \left(\sum_{i=j+1}^{k-1} t_i A\,\mathbf{v}^{(i)}\right) \\
&= -\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right).
\end{aligned}$$

$A-$orthogonality: A Gift from Math God

$A-$orthogonality: A Gift from Math God

Set $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$, and for $k = 2, \cdots, n$, write

$$\mathbf{v}^{(k)} = \sum_{j=0}^{k-1} \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(j)}\right)}{\left(\mathbf{r}^{(j)}\right)^T \left(\mathbf{r}^{(j)}\right)} \mathbf{r}^{(j)}.$$

Then

$$\begin{aligned}
\mathbf{v}^{(k)} &= -\sum_{j=0}^{k-1} \frac{\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{r}^{(j)}\right)^T \left(\mathbf{r}^{(j)}\right)} \mathbf{r}^{(j)} \\
&= -\mathbf{r}^{(k-1)} - \frac{\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{r}^{(k-2)}\right)^T \left(\mathbf{r}^{(k-2)}\right)} \sum_{j=0}^{k-2} \frac{\left(\mathbf{r}^{(k-2)}\right)^T \left(\mathbf{r}^{(k-2)}\right)}{\left(\mathbf{r}^{(j)}\right)^T \left(\mathbf{r}^{(j)}\right)} \mathbf{r}^{(j)} \\
&= -\mathbf{r}^{(k-1)} + s_{k-1}\, \mathbf{v}^{(k-1)},
\end{aligned}$$

with $s_{k-1} = \frac{\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{r}^{(k-2)}\right)^T \left(\mathbf{r}^{(k-2)}\right)}$.

Thm: Let $\{\mathbf{v}^{(i)}\}_{i=1}^{n}$ be A-**orthogonal** with $\mathbf{v}^{(1)} = -\mathbf{r}^{(0)}$ and for $k = 1, \cdots, n$

$$t_k = \frac{\left(\mathbf{v}^{(k)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)} = -\frac{\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}, \quad \mathbf{x}^{(k)} \stackrel{def}{=} \mathbf{x}^{(k-1)} + t_k \mathbf{v}^{(k)}.$$

Then $A\,\mathbf{x}^{(n)} = b$ in *exact arithmetic*.

Conjugate Gradient Algorithm

Thm: For $k = 1, \cdots, n$, define,

$$\mathbf{v}^{(k)} = -\mathbf{r}^{(k-1)} + s_{k-1}\,\mathbf{v}^{(k-1)} \text{ with } s_{k-1} = \frac{\left(\mathbf{r}^{(k-1)}\right)^T\left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{r}^{(k-2)}\right)^T\left(\mathbf{r}^{(k-2)}\right)},$$

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + t_k\,\mathbf{v}^{(k)} \quad \text{with} \quad t_k = -\frac{\left(\mathbf{r}^{(k-1)}\right)^T\left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T\left(A\,\mathbf{v}^{(k)}\right)}.$$

Then vectors $\{\mathbf{v}^{(k)}\}$ are A-orthogonal and $A\,\mathbf{x}^{(n)} = b$ in *exact arithmetic*.

The CG Algorithm: C is for *Craft*, G is for *Gift*.

Algorithm 1 Conjugate Gradient Algorithm

Input: Symmetric positive definite $A \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$,
initial guess $\mathbf{x}^{(0)} \in \mathbb{R}^n$, and tolerance $\tau > 0$.

Output: Approximate solution $\mathbf{x}$.

Algorithm:

Initialize: $\mathbf{r}^{(0)} = \mathbf{b} - A\,\mathbf{x}^{(0)}$, $\mathbf{v}^{(0)} = -\mathbf{r}^{(0)}$, $k = 1$

while $\left\| \mathbf{r}^{(k-1)} \right\|_2 \geq \tau$ **do**

$$t_k = -\frac{\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}{\left(\mathbf{v}^{(k)}\right)^T \left(A\,\mathbf{v}^{(k)}\right)}.$$

$$\mathbf{x}^{(k)} = \mathbf{x}^{(k-1)} + t_k\,\mathbf{v}^{(k)}$$

$$\mathbf{r}^{(k)} = \mathbf{r}^{(k-1)} - t_k\,A\,\mathbf{v}^{(k)}.$$

$$s_k = \frac{\left(\mathbf{r}^{(k)}\right)^T \left(\mathbf{r}^{(k)}\right)}{\left(\mathbf{r}^{(k-1)}\right)^T \left(\mathbf{r}^{(k-1)}\right)}.$$

$$\mathbf{v}^{(k+1)} = -\mathbf{r}^{(k)} + s_k\,\mathbf{v}^{(k)} \ .$$

$$k = k + 1.$$

end while