Reasons of Gauss-Seidel Methods I

- An optimization problem

$$\min_x \frac{1}{2}x^T Ax - b^T x$$

is the same as solving

$$Ax - b = 0$$

if A is symmetric positive definite

Reasons of Gauss-Seidel Methods II

- If

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

then

$$\nabla f(x) \equiv \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = Ax - b$$

Reasons of Gauss-Seidel Methods III

- Derivation:

$$
\begin{aligned}
x^T A x &= \sum_{i=1}^{n} x_i (Ax)_i \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i A_{ij} x_j \\
&= x_1 A_{11} x_1 + \cdots + x_1 A_{1n} x_n \\
&\quad + x_2 A_{21} x_1 + \cdots + x_n A_{n1} x_1 + \cdots
\end{aligned}
$$

Reasons of Gauss-Seidel Methods IV

Therefore

$$\begin{aligned}\frac{\partial x^T A x}{\partial x_1} &= 2A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n \\ &\quad + x_2A_{21} + \cdots + x_nA_{n1} \\ &= 2(A_{11}x_1 + \cdots + A_{1n}x_n)\end{aligned}$$

- Coordinate-wise minimization is a classic way for solving optimization problems

Reasons of Gauss-Seidel Methods V

Sequentially update one variable at a time

$$
\begin{gathered}
\min_{x_1} f(x_1, x_2^k, \ldots, x_n^k) \\
\min_{x_2} f(x_1^{k+1}, x_2, \ldots, x_n^k) \\
\min_{x_3} f(x_1^{k+1}, x_2^{k+1}, x_3, \ldots, x_n^k) \\
\vdots
\end{gathered}
$$

Reasons of Gauss-Seidel Methods VI

- Assume x_i is updated

$$\min_{d} \frac{1}{2}(x + de_i)^T A(x + de_i) - b^T(x + de_i)$$

where

$$e_i = [\underbrace{0, \ldots, 0}_{i-1}, 1, 0, \ldots, 0]^T$$

Reasons of Gauss-Seidel Methods VII

Next,

$$\min_d \frac{1}{2}d^2 A_{ii} + d(Ax)_i - b_i d$$

$$A_{ii}d + (Ax)_i - b_i = 0$$

$$d = \frac{b_i - (Ax)_i}{A_{ii}}$$

$$x_i + d \leftarrow \frac{b_i - \sum_{j:j\neq i} A_{ij}x_j}{A_{ii}}$$

We then get the Gauss-Seidel update