1 February 25th, 2021

1.1 Gauss-Seidel Iteration

Recall that in Jacobi, $\xi_i^{(k+1)}$ are updated parallelly in Algorithm ??.

Gauss-Seidel is just modifying Jacobi to be updated "successively". When we modify $\xi_i^{(k+1)}$ we can use $\xi_1^{(k+1)}, \dots \xi_{i-1}^{(k+1)}$. This gives us Algorithm 1.

Algorithm 1 Gauss-Seidel Iteration

```
1: for k = 0, 1, 2, ... do
2:     for i = 1, ..., n do
3:         ξ_i^(k+1) = (β_i − Σ_{j>i} a_ij ξ_j^(k) − Σ_{j<i} a_ij ξ_j^(k+1))/a_ii
4:     end for
5: end for
```

In the matrix reformulation, recalling that $A = D - E - F$, where E is lower triangular and F is upper triangular, we have:

$$\begin{aligned} & x_{k+1} = D^{-1}(b + Fx_k + Ex_{k+1}) \\ \iff & (D-E)x_{k+1} = b + Fx_k \\ \iff & x_{k+1} = (D-E)^{-1}Fx_k + (D-E)^{-1}b \\ \iff & x_{k+1} = \underbrace{(I-(D-E)^{-1}A)}_{G}\, x_x + \underbrace{(D-E)^{-1}}_{f}\, b. \end{aligned}$$

which is in the stationary iteration form.

Remark 1.1 — In the algorithm above, we are updating $\xi_i^{(k+1)}$ in regular numerical order (from 1 to n). However, we can use other ordering, such as going from n to 1 in reverse order. Gauss-Seidel is sensitive to the ordering of updating of unknowns.

1.1.1 Convergence of Gauss-Seidel

Since Gauss-Seidel is a stationary iteration, we can use the same framework as the Jacobi iteration, with:

$$(x_k - x_*) = G^k(x_0 - x_*)$$

Theorem 1.2

Gauss-Seidel converges to x_* for any x_0 if and only if $\rho(G) < 1$.

The convergence speed is $\|G\|$, as we have:

$$\|x_{k+1} - x_*\| \le \|G\|\|x_k - x_*\|$$

In particular, we have:

$$\|G\| = \begin{cases} \rho(G) & \text{if } G \text{ is symmetric and } \|\cdot\| = \|\cdot\|_2 \\ \rho(G) + \epsilon & \text{otherwise} \end{cases}$$

with ϵ arbitrarily small.

Theorem 1.3

Let $A \in \mathbb{R}^{n\times n}$ be SPD. Then $\rho(G) < 1$.

Proof. Recall $G = I - (D-E)^{-1}A$. Let (λ, u) be an eigenpair of G (note that λ, u may be complex). We have:

$$\begin{aligned} Gu &= \lambda u \\ (I - (D-E)^{-1}Au) &= \lambda u \\ (D-E)^{-1}Fu &= \lambda u \\ Fu &= \lambda(D-E)^{-1}u. \end{aligned}$$

Since A is symmetric, we have:

$$A = A^T \implies F = E^T.$$

Thus, we have:

$$\begin{aligned} \lambda(D-E)u &= E^T u \\ &\lambda u^*(D-E)u = u^*E^T u. \end{aligned}$$

by left-multiplying both side by $u^* = \overline{u}^T$

Let us set $u^*E^Tu = \alpha + i\beta \in \mathbb{C}$ and $u^*Du = \delta \in \mathbb{R}$, since D is real and symmetric. Then, we have:

$$u^*E^Tu = (u^*E^Tu)^T = u^TE(u^*)^T = \overline{\overline{u^*}\overline{E}\overline{u}} = \overline{u^*Eu} = \alpha - i\beta$$

Since A is SPD, D is also SPD, giving us $\lambda = u^*Du > 0$.

Thus

$$\lambda(\delta - (\alpha + i\beta)) = \alpha - i\beta \implies \lambda = \frac{\alpha - i\beta}{(\delta - \alpha) - i\beta}$$

Giving us:

$$|\lambda|^2 = \frac{(\alpha^2 + \beta^2)}{(\delta-\alpha)^2 + \beta^2}$$

Because:

$$(\delta - \alpha)^2 + \beta^2 = \delta^2 + \alpha^2 + \beta^2 - 2\alpha\delta = (\alpha^2 + \beta^2) + \underbrace{\delta}_{>0}(\delta - 2\alpha)$$

we need to check the sign of $\delta - 2\alpha$.

Since A is SPD, $u^*Au > 0 \implies u^*(D - E - E^T)u = \delta - (\alpha + i\beta) - (\alpha - i\beta) = \delta - 2\alpha > 0$. As such:

$$(\delta - \alpha)^2 + \beta^2 > \alpha^2 + \beta^2 \implies |\lambda|^2 < 1$$

□

Remark 1.4 — This is more general than Jacobi, which is convergent if and only if $2D - A$ is SPD.

Example 1.5

Recall that the 1D discrete Laplacian is SPD, meaning that Gauss-Seidel converges. Similarly for the 2D case.

Example 1.6

Gauss-Seidel converges for irriducible diagonally dominant matrices. This is used for laplacian on irregular domains.

1.2 Acceleration of G-S / and Jacobi (SOR)

Here we try to accelerate Gauss-Seidel and Jacobi by considering momentum acceleration. We can consider the iteration of x_k to x_{k+1} as a movement by the Gauss-Seidel iteration. The idea of SOR is to consider the momentum of the movement until some stage.

For Gauss-Seidel, the components are:

- $x_k : \xi_i^{(k)}$
- $x_{k+1} : \xi_i^{(k+1)} = (\beta_i - \sum_{j>i} a_{ij}\xi_j^{(k)} - \sum_{j<i} a_{ij}\xi_j^{(k+1)})/a_{ii}$

To consider the momentum, we would have:

$$\begin{aligned}\xi_i^{(k+1)} &= \xi_i^{(k)} + \omega(\xi_i^{(k+1)} - \xi_i^{(k)}) \\ &= \xi_i^{(k)} + \omega\left((\beta_i - \sum_{j>i} a_{ij}\xi_j^{(k)} - \sum_{j<i} a_{ij}\xi_j^{(k+1)})/a_{ii} - \xi_i^{(k)}\right).\end{aligned}$$

for some $\omega > 1$, giving us the Algorithm 2.

Algorithm 2 SOR Iteration

1: **for** $k = 0, 1, 2, \ldots$ **do**
2:     **for** $i = 1, \ldots, n$ **do**
3:         $\xi_i^{(k+1)} = \xi_i^{(k)} + \omega\left[(\beta_i - \sum_{j>i} a_{ij}\xi_j^{(k)} - \sum_{j<i} a_{ij}\xi_j^{(k+1)})/a_{ii} - \xi_i^{(k)}\right]$
4:     **end for**
5: **end for**

In the matrix reformulation, we have:

$$x_{k+1} = x_k + \omega\left[D^{-1}(b + Fx_k + Ex_{k+1}) - x_k\right]$$

Expressing this in its stationary form, we have:

$$x_{k+1} = \underbrace{(D - \omega E)^{-1}((1-\omega)D + \omega F)}_{G_\omega} x_k + \underbrace{\omega(D - \omega E)^{-1}}_{f_\omega} b$$

Now the question is finding a suitable ω.

1.2.1 Convergence of SOR

Theorem 1.7

If $\omega \notin (0,2)$, then $\rho(G_\omega) > 1$.

Proof. We have:

$$\det(G_\omega) = \det((D-\omega E)^{-1}) \cdot \det((1-\omega)D + \omega F)$$

- Since $D - \omega E$ is lower triangular, $(D-\omega E)^{-1}$ is also lower triangular. Thus, $\det((D-\omega E)^{-1})$ is the product of its diagonals, which are the inverse of the diagonal of $D - \omega E$ which are $\frac{1}{a_{ii}}$. As such, we have:

$$\det((D-\omega E)^{-1}) = \prod_{i=1}^{n} a_{ii}^{-1}$$

- Since $(1-\omega)D + \omega F$ is upper triangular, we have:

$$\det((1-\omega)D + \omega F) = \prod_{i=1}^{n} (1-\omega)\ a_{ii}$$

Thus, we have:

$$\det(G_\omega) = \prod_{i=1}^{n} (1-\omega)\ a_{ii} \cdot a_{ii}^{-1} = (1-\omega)^n = \prod_{i=1}^{n} \lambda_i$$

This gives us:

$$\begin{aligned} &|1-\omega|^n = |\lambda_1||\lambda_2| \dots |\lambda_n| \leq (\rho(G_\omega))^n \\ \implies &|1-\omega| \leq \rho(G_\omega). \end{aligned}$$

□

Corollary 1.8

In order for SOR to converge, we must have $\omega < 2$.

- If $\omega = 1$, we have Gauss Seidel
- If $\omega \in (0,1)$, we have successive under relaxation
- If $\omega \in (1,2)$ we have successive over relaxation (SOR).

Theorem 1.9

If A is SPD, then $\rho(G_\omega) < 1$ for all $\omega \in (0,2)$.

Proof. Similar to Gauss-Seidel. □

How do we pick the optimal ω? Roughly speaking, with the optimal ω:

- $\rho(G_{\omega_{\text{opt}}}) \sim \rho(G_1)^2$. Where G_1 is the iteration matrix in G-S.

Remark 1.10 — This means that SOR is roughly two times faster than G-S, since it is the square.

- $\rho(G_1) \sim (\rho(G_{\text{Jacobi}}))^2$.