Convergence of the basic iterative methods

We already studied the convergence of Jacobi and Gauss-Seidel iterations for strictly diagonally dominant matrices $A$ in $Ax = b$. This class is not exotic, but much larger in the class of SPD matrices, so we consider $A$ an SPD.

In fact, we can study convergence of much more general class of iterative method based on the following more general splitting of $A$

1. $A$ is an SPD matrix, $(Ax,x) > 0$, $x \neq 0$ even when $x \in \mathbb{C}^n$.
2. $A = D - C - C^T$, where
   (a) $D$ is an SPD matrix
   (b) $\alpha D - C$, $\alpha > \frac{1}{2}$ is invertible
Examples of splittings

I. The simplest case is when

\[ D = \text{diagonal of } A \]

obviously all \( a_{ii} > 0 \) \( \Rightarrow \) \( D \) is an SPD

\[ C = \text{strictly lower triangular part of } A \]

obviously this is the classical SOR, GS

II. Let \( A \) be written in the block form

\[ A = D - C - CT \]

where \( D \) is block diagonal matrix

III. Any other splitting that satisfies the above requirements

Example

\[ X = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \]

\[ x_j = (x_{j1}, x_{j2}, \ldots, x_{jm}) \]

\[ x_m \]

\[ x_1 \]

\[ -x_1 \]
\[ D_1 = \begin{bmatrix} 4 + h^2 & -1 & 0 & 0 & 0 \\ -1 & 4 + h^2 & -1 \\ 0 & 0 & -1 & 4 + h^2 \end{bmatrix} \text{ n x n matrix} \]

\[ C_1 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \text{ n x n matrix} \]

\[ A = \begin{bmatrix} D_1 & 0 & 0 & 0 \\ 0 & D_1 & 0 & 0 \\ 0 & 0 & \ldots & D_1 \end{bmatrix} \]

\[ A \mathbf{x} = \mathbf{A} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = -n \text{ dimensional vector} \]

\[ D_1^{-1} \text{ is essentially inverting the above tridiagonal matrix} \]
Possible splits:

\[
\begin{array}{cccc}
  & y_1 & y_2 & y_3 & y_4 \\
1 & 1 & 0 & 0 & 0 \\
2 & 0 & 1 & 0 & 0 \\
3 & 0 & 0 & 1 & 0 \\
4 & 0 & 0 & 0 & 1 \\
\end{array}
\]

Reordering of the unknowns in some applications.

\[
\begin{bmatrix}
  2 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 2 \\
\end{bmatrix} \begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3 \\
  x_4 \\
\end{bmatrix} = b
\]

\[
\begin{bmatrix}
  2 & 0 & -1 & 0 \\
 0 & 2 & -1 & -1 \\
-1 & -1 & 2 & 0 \\
0 & -1 & 0 & 2 \\
\end{bmatrix} \begin{bmatrix}
  y_1 \\
  y_2 \\
  y_3 \\
  y_4 \\
\end{bmatrix}
\begin{cases}
  \text{red} \\
  \text{black}
\end{cases}
\]

\[
A = \begin{bmatrix} D_1 & C^T \\ C & D_1 \end{bmatrix}
\]

\[
D_2 = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}
\]

Red black ordering of the nodes/unknowns.
Convergence of SOR (and GS) for $A$ an SPD

Consider $A$ an SPD matrix, i.e., $(Ax,x)>0$, $x \neq 0$

$$A = D - C - C^T \quad (C - \text{lower triangular})$$

$$Q = \alpha D - C$$

More general splitting is when $C$ is such

that $D$, SPD, $4D - C - C^T = A$

**Theorem:** If $A$ is an SPD matrix, $Q$-SPD and $\alpha > \frac{1}{2}$ then the SOR iteration with

$Q = \alpha D - C$ converges for any starting vector $x^{(0)}$

$$Q (x^{(k+1)} - x^{(k)}) = b - A x^{(k)}$$

Equivalent

$$x^{(k+1)} = (I - Q^{-1}A) x^{(k)} + Q^{-1}b$$

**Remark:** SOR $\alpha = \frac{1}{2}$, $D$-diagonal $C$-lower

$G$-S $\alpha = 1$

**Proof:** We write the error transfer matrix

$G = I - Q^{-1}A$ and we need to prove that $p(G) < 1$. 
Let \((\lambda, x)\) be an eigenpair of \(Q\), i.e.
\[ Qx = \lambda x \]
\(\Rightarrow\) define \(y = (I - Q)x\)

or \((1)\)
\[ y = x - Qx = (1 - \lambda)x = x - (I - Q)x = Q^T Ax \]

\[ \Rightarrow \quad Qy = Ax \neq (\lambda D - C)y = Ax \]

and also

\(2)\)
\[ Q - A = \lambda D - C - (D - C - C^T) = \lambda D - D + C^T \]

\[(a)\]
\[ (\lambda D - C)y = Ax \]
\[ (\lambda D - D + C^T)y = (Q - A)y = Ax - Ay = A(x - y) = A(x - Qx) \]

\[(b)\]
\[ (\lambda D - D + C^T)y = AGx \]

(a) take inner product with \(y\) to get
\[ ((\lambda D - C)y, y) = (Ax, y) \]

(b) take inner product with \(y\) to get
\[ ((\lambda D - D + C^T)y, y) = (AGx, y) \]
\[ \alpha(Dy,y) - (C'y,y) = (Ax, y) \]
\[ \alpha(Dy,y) - (D'y,y) + (C'y,y) = (y, A'Ax) \]

Now add these two to get
\[ (2\alpha-1)(D'y,y) = (Ax, y) + (y, A'Ax) \]
\[ (2\alpha-1)(Dy,y) = (Ax, (1-x)x) + (y, (1-x)x, A'Ax) \]
\[ (2\alpha-1)(D'y,y) = (1-x)(Ax, x) + (1-x)(x, A'Ax) \]
\[ y = (1-x)x \]
\[ (2\alpha-1)(1-x)^2(D'y,x) = (1-121^2)(Ax, x) \]
\[ 0 \quad \text{real} > 0 \quad \text{real} > 0 \]

Is it possible \( \alpha = 1 \) ?
\[ Gx = x \quad \Rightarrow \quad (I - G)x = 0 \]
\[ (I - \Xi + QA)x = 0 \quad QA \text{ Singularity} \]

Thus \( 1 - |x|^2 > 0 \) impossible
\[ (Ax, x) > 0 \quad \Rightarrow \quad |x|^2 > 0 \]
\[ (D'y, x) > 0 \quad |x| < 1 \]

\[ p(G) < 1 \]

Converges
There is a way to optimize the choice of the parameter in SOR. For the case of the matrix

\[
A = \begin{bmatrix}
2 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 2
\end{bmatrix}
\]

\[
\omega_{\text{opt}} = 2 \left(1 - \frac{\lambda}{n+1}\right)
\]

for this choice \( \rho(G_{\text{SOR}}) = 1 - \frac{2\lambda}{n+1} \)

Please, run your programs of PA*2 with this choice of \( \omega \) to see the difference in the number of iterations.

The same example Jacobi has

\[
\rho(G_{\text{Jacobi}}) = 1 - \frac{\pi^2}{(n+1)^2}
\]

much closer to 1

\[
\rho(G) = \rho(G_{\text{Jacobi}})
\]

much slower to converge.

I would like you to experiment with these two methods for this matrix to see the difference in the number of iterations when you run it for \( n = 20, 40, 80 \).
Quitting

Any matrix

\[ \| A \|_2 = \frac{\| Ax \|_2}{\| x \|_2} = ? \]

Any symmetric \( A \) is diagonalizable \( A \) is symmetric \( A = A^T \)

\[ \| A \|_2 = \rho(A) = \max_{\lambda \in \text{spec}(A)} | \lambda | \]

\[ \| A \|_2 = \frac{\| A x, A x \|}{\| x \|_2^2} = \frac{(A^T A x, x)}{(x, x)} = \rho(A^T A) \]

\[ \| A \|_2 = \rho(A A^T)^{1/2} \]

A strictly diagonally dominant Jacobi converges
A strictly column-wise diagonally dominant does Jacobi converge ?

A SPD \( \rho(\text{rich}) = | 1 - \epsilon \lambda | = \frac{\lambda - \lambda_0}{\lambda_0 + \lambda} \)

Richardson \( \epsilon = \lambda_1 \leq \ldots \leq \lambda_n = \lambda_n \)