

Gradiente Conjugado Precondicionado - CGNR y CGNE

M. Sepúlveda

8 de Junio, 2021



Precondicionamiento

- La velocidad de convergencia del Método de GC (y Gradiente) está sujeto al condicionamiento de la matriz; con una matriz mal condicionada el método converge muy lentamente.



Precondicionamiento

- La velocidad de convergencia del Método de GC (y Gradiente) está sujeto al condicionamiento de la matriz; con una matriz mal condicionada el método converge muy lentamente.
- La idea (general) de Precondicionamiento es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ tal que \mathbf{B} sea "*fácil*" de invertir, y $\text{cond}(\mathbf{B}^{-1}\mathbf{A}) \ll \text{cond}(\mathbf{A})$ de modo que el sistema "mal condicionado" $\mathbf{A}\mathbf{x} = \mathbf{b}$ es reemplazado por un sistema bien condicionado equivalente:

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \mathbf{B}^{-1}\mathbf{b}.$$



Precondicionamiento

- La velocidad de convergencia del Método de GC (y Gradiente) está sujeto al condicionamiento de la matriz; con una matriz mal condicionada el método converge muy lentamente.
- La idea (general) de Precondicionamiento es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ tal que \mathbf{B} sea "fácil" de invertir, y $\text{cond}(\mathbf{B}^{-1}\mathbf{A}) \ll \text{cond}(\mathbf{A})$ de modo que el sistema "mal condicionado" $\mathbf{A}\mathbf{x} = \mathbf{b}$ es reemplazado por un sistema bien condicionado equivalente:

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \mathbf{B}^{-1}\mathbf{b}.$$

- El problema principal es que si \mathbf{A} es simétrica definida positiva, es "muy difícil" (y limitante) asegurar que $\mathbf{B}^{-1}\mathbf{A}$ también lo sea.



Precondicionamiento

- La velocidad de convergencia del Método de GC (y Gradiente) está sujeto al condicionamiento de la matriz; con una matriz mal condicionada el método converge muy lentamente.
- La idea (general) de Precondicionamiento es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ tal que \mathbf{B} sea "fácil" de invertir, y $\text{cond}(\mathbf{B}^{-1}\mathbf{A}) \ll \text{cond}(\mathbf{A})$ de modo que el sistema "mal condicionado" $\mathbf{A}\mathbf{x} = \mathbf{b}$ es reemplazado por un sistema bien condicionado equivalente:

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{x} = \mathbf{B}^{-1}\mathbf{b}.$$

- El problema principal es que si \mathbf{A} es **simétrica definida positiva**, es "*muy difícil*" (y limitante) asegurar que $\mathbf{B}^{-1}\mathbf{A}$ también lo sea.
- $\mathbf{B}^{-1}\mathbf{A}$ no necesariamente es simétrica definida positiva, pero si puede ser similar a una matriz que si lo sea.



Precondicionamiento para Gradiente Conjugado

- La idea de la técnica de precondicionamiento para Gradiente Conjugado es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ simétrica definida positiva, fácil de invertir tal que $\mathbf{B}^{-1}\mathbf{A}$ sea una buena aproximación de la identidad.



Precondicionamiento para Gradiente Conjugado

- La idea de la técnica de precondicionamiento para Gradiente Conjugado es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ simétrica definida positiva, fácil de invertir tal que $\mathbf{B}^{-1}\mathbf{A}$ sea una buena aproximación de la identidad.
- Se define entonces la matriz simétrica definida positiva y similar a $\mathbf{B}^{-1}\mathbf{A}$:

$$\mathbf{A}' = \mathbf{B}^{1/2}(\mathbf{B}^{-1}\mathbf{A})\mathbf{B}^{-1/2} = \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2},$$

de modo que $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$.



Precondicionamiento para Gradiente Conjugado

- La idea de la técnica de precondicionamiento para Gradiente Conjugado es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ simétrica definida positiva, fácil de invertir tal que $\mathbf{B}^{-1}\mathbf{A}$ sea una buena aproximación de la identidad.
- Se define entonces la matriz simétrica definida positiva y similar a $\mathbf{B}^{-1}\mathbf{A}$:

$$\mathbf{A}' = \mathbf{B}^{1/2}(\mathbf{B}^{-1}\mathbf{A})\mathbf{B}^{-1/2} = \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2},$$

de modo que $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$.

- Si \mathbf{A} es simétrica definida positiva, entonces existe $\mathbf{C} := \mathbf{B}^{1/2}$ tal que $\mathbf{C}^2 = \mathbf{B}$.



Precondicionamiento para Gradiente Conjugado

- La idea de la técnica de precondicionamiento para Gradiente Conjugado es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ simétrica definida positiva, fácil de invertir tal que $\mathbf{B}^{-1}\mathbf{A}$ sea una buena aproximación de la identidad.
- Se define entonces la matriz simétrica definida positiva y similar a $\mathbf{B}^{-1}\mathbf{A}$:

$$\mathbf{A}' = \mathbf{B}^{1/2}(\mathbf{B}^{-1}\mathbf{A})\mathbf{B}^{-1/2} = \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2},$$

de modo que $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$.

- Si \mathbf{A} es simétrica definida positiva, entonces existe $\mathbf{C} := \mathbf{B}^{1/2}$ tal que $\mathbf{C}^2 = \mathbf{B}$.
- En la práctica más que \mathbf{B} sea fácil de invertir, nos interesa que $\mathbf{B}\mathbf{q} = \mathbf{r}$ sea fácil de resolver, lo cual es el caso con la descomposición de Choleski $\mathbf{B} = \mathbf{L}\mathbf{L}^T$.



Precondicionamiento para Gradiente Conjugado

- La idea de la técnica de precondicionamiento para Gradiente Conjugado es encontrar una matriz $\mathbf{B} \approx \mathbf{A}$ simétrica definida positiva, fácil de invertir tal que $\mathbf{B}^{-1}\mathbf{A}$ sea una buena aproximación de la identidad.
- Se define entonces la matriz simétrica definida positiva y similar a $\mathbf{B}^{-1}\mathbf{A}$:

$$\mathbf{A}' = \mathbf{B}^{1/2}(\mathbf{B}^{-1}\mathbf{A})\mathbf{B}^{-1/2} = \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2},$$

de modo que $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$.

- Si \mathbf{A} es simétrica definida positiva, entonces existe $\mathbf{C} := \mathbf{B}^{1/2}$ tal que $\mathbf{C}^2 = \mathbf{B}$.
- En la práctica más que \mathbf{B} sea fácil de invertir, nos interesa que $\mathbf{B}\mathbf{q} = \mathbf{r}$ sea fácil de resolver, lo cual es el caso con la descomposición de Choleski $\mathbf{B} = \mathbf{L}\mathbf{L}^T$.
- Una vez elegido \mathbf{B} , el vector $\mathbf{x}' := \mathbf{B}^{1/2}\bar{\mathbf{x}}$ es solución de

$$\mathbf{A}'\mathbf{x}' = \mathbf{b}',$$

con $\mathbf{b}' := \mathbf{B}^{-1/2}\mathbf{b}$, el cual es equivalente a $\mathbf{A}\mathbf{x} = \mathbf{b}$.



Método de Gradiente Conjugado para $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$

- Aplicamos ahora Gradiente conjugado al sistema bien condicionado $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$, usando $\mathbf{x}'_0 = \mathbf{B}^{1/2}\mathbf{x}_0$ como vector inicial.



Método de Gradiente Conjugado para $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$

- Aplicamos ahora Gradiente conjugado al sistema bien condicionado $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$, usando $\mathbf{x}'_0 = \mathbf{B}^{1/2}\mathbf{x}_0$ como vector inicial.
- Como $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$, la secuencia \mathbf{x}'_k generada por el método de GC convergerá muy rápido a $\bar{\mathbf{x}}'$.



Método de Gradiente Conjugado para $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$

- Aplicamos ahora Gradiente conjugado al sistema bien condicionado $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$, usando $\mathbf{x}'_0 = \mathbf{B}^{1/2}\mathbf{x}_0$ como vector inicial.
- Como $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$, la secuencia \mathbf{x}'_k generada por el método de GC convergerá muy rápido a $\bar{\mathbf{x}}'$.
- Pero en lugar de calcular la matriz \mathbf{A}' y el vector \mathbf{x}'_k explícitamente, generamos la sucesión $\mathbf{x}_k := \mathbf{B}^{-1/2}\mathbf{x}'_k$ directamente como sigue:



Método de Gradiente Conjugado para $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$

- Aplicamos ahora Gradiente conjugado al sistema bien condicionado $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$, usando $\mathbf{x}'_0 = \mathbf{B}^{1/2}\mathbf{x}_0$ como vector inicial.
- Como $\text{cond}(\mathbf{A}') \ll \text{cond}(\mathbf{A})$, la secuencia \mathbf{x}'_k generada por el método de GC convergerá muy rápido a $\bar{\mathbf{x}}'$.
- Pero en lugar de calcular la matriz \mathbf{A}' y el vector \mathbf{x}'_k explícitamente, generamos la sucesión $\mathbf{x}_k := \mathbf{B}^{-1/2}\mathbf{x}'_k$ directamente como sigue:
- Usando las reglas de transformación:

$$\begin{aligned}\mathbf{A}' &= \mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}, & \mathbf{b}' &= \mathbf{B}^{-1/2}\mathbf{b} \\ \mathbf{x}_k &:= \mathbf{B}^{-1/2}\mathbf{x}'_k, \\ \mathbf{r}'_k &= \mathbf{b}' - \mathbf{A}'\mathbf{x}'_k = \mathbf{B}^{-1/2}\mathbf{r}_k \\ \mathbf{p}'_k &= \mathbf{B}^{1/2}\mathbf{p}_k,\end{aligned}$$

obtenemos el Algoritmo de Gradiente Conjugado Precondicionado.



Algoritmo de Gradiente Conjugado Precondicionado

- **Algoritmo:**

Dado el vector inicial x_0 ,

$$r_0 = \mathbf{b} - \mathbf{A}x_0,$$

$$q_0 = \mathbf{B}^{-1}r_0; \quad p_0 = q_0,$$

para $k = 0, 1, 2, \dots$

$$\alpha_k = \frac{r_k^T q_k}{p_k^T \mathbf{A} p_k},$$

$$x_{k+1} = x_k + \alpha_k p_k,$$

$$r_{k+1} = r_k - \alpha_k \mathbf{A} p_k,$$

$$q_{k+1} = \mathbf{B}^{-1}r_{k+1},$$

$$\beta_k = \frac{r_{k+1}^T q_{k+1}}{r_k^T q_k},$$

$$p_{k+1} = q_{k+1} + \beta_k p_k,$$

hasta que se satisfaga un criterio de detención,
o bien hasta que $p_{k+1} = 0$.



Características del Algoritmo

- Esencialmente la única diferencia comparada con el Algoritmo GC es que en cada paso se resuelve adicionalmente un sistema lineal $\mathbf{B}q = \mathbf{r}$ ($\mathbf{Ly} = \mathbf{r}$ y $\mathbf{L}^T q = \mathbf{y}$, con $\mathbf{B} = \mathbf{LL}^T$).



Características del Algoritmo

- Esencialmente la única diferencia comparada con el Algoritmo GC es que en cada paso se resuelve adicionalmente un sistema lineal $\mathbf{B}\mathbf{q} = \mathbf{r}$ ($\mathbf{Ly} = \mathbf{r}$ y $\mathbf{L}^T\mathbf{q} = \mathbf{y}$, con $\mathbf{B} = \mathbf{LL}^T$).
- Al reemplazar \mathbf{x}_k , \mathbf{q}_k , \mathbf{p}_k , y \mathbf{r}_k , por \mathbf{x}'_k , \mathbf{p}'_k , y \mathbf{r}'_k , se obtiene exactamente la descripción del Algoritmo de GC para $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$.



Características del Algoritmo

- Esencialmente la única diferencia comparada con el Algoritmo GC es que en cada paso se resuelve adicionalmente un sistema lineal $\mathbf{B}\mathbf{q} = \mathbf{r}$ ($\mathbf{Ly} = \mathbf{r}$ y $\mathbf{L}^T\mathbf{q} = \mathbf{y}$, con $\mathbf{B} = \mathbf{LL}^T$).
- Al reemplazar \mathbf{x}_k , \mathbf{q}_k , \mathbf{p}_k , y \mathbf{r}_k , por \mathbf{x}'_k , \mathbf{p}'_k , y \mathbf{r}'_k , se obtiene exactamente la descripción del Algoritmo de GC para $\mathbf{A}'\mathbf{x}' = \mathbf{b}'$.
- No es necesario calcular $\mathbf{B}^{1/2}$.



Elección del Precondicionador B : (SSOR)

- Se puede considerar como matriz $\mathbf{B} \approx \mathbf{A}$ a la matriz \mathbf{M} del método de SSOR (Sobre-relajación sucesiva simétrica) en la descomposición $\mathbf{A} = \mathbf{M} - \mathbf{N}$:

$$\mathbf{B} = \frac{1}{2-\omega} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E} \right) \left(\frac{1}{\omega} \mathbf{D} \right)^{-1} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E}^T \right)$$

para un valor $\omega \in (0, 2)$ adecuado.



Elección del Precondicionador B : (SSOR)

- Se puede considerar como matriz $\mathbf{B} \approx \mathbf{A}$ a la matriz \mathbf{M} del método de SSOR (Sobre-relajación sucesiva simétrica) en la descomposición $\mathbf{A} = \mathbf{M} - \mathbf{N}$:

$$\mathbf{B} = \frac{1}{2-\omega} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E} \right) \left(\frac{1}{\omega} \mathbf{D} \right)^{-1} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E}^T \right)$$

para un valor $\omega \in (0, 2)$ adecuado.

- \mathbf{D} y \mathbf{E} corresponden a la descomposición estandar $\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{E}^T$.



Elección del Precondicionador B : (SSOR)

- Se puede considerar como matriz $\mathbf{B} \approx \mathbf{A}$ a la matriz \mathbf{M} del método de SSOR (Sobre-relajación sucesiva simétrica) en la descomposición $\mathbf{A} = \mathbf{M} - \mathbf{N}$:

$$\mathbf{B} = \frac{1}{2-\omega} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E} \right) \left(\frac{1}{\omega} \mathbf{D} \right)^{-1} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E}^T \right)$$

para un valor $\omega \in (0, 2)$ adecuado.

- \mathbf{D} y \mathbf{E} corresponden a la descomposición estandar $\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{E}^T$.
- Si \mathbf{A} es simétrica definida positiva, entonces \mathbf{B} también y $\mathbf{L} = (1/\omega)\mathbf{D} - \mathbf{E}$.



Elección del Precondicionador B : (SSOR)

- Se puede considerar como matriz $\mathbf{B} \approx \mathbf{A}$ a la matriz \mathbf{M} del método de SSOR (Sobre-relajación sucesiva simétrica) en la descomposición $\mathbf{A} = \mathbf{M} - \mathbf{N}$:

$$\mathbf{B} = \frac{1}{2-\omega} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E} \right) \left(\frac{1}{\omega} \mathbf{D} \right)^{-1} \left(\frac{1}{\omega} \mathbf{D} - \mathbf{E}^T \right)$$

para un valor $\omega \in (0, 2)$ adecuado.

- \mathbf{D} y \mathbf{E} corresponden a la descomposición estandar $\mathbf{A} = \mathbf{D} - \mathbf{E} - \mathbf{E}^T$.
- Si \mathbf{A} es simétrica definida positiva, entonces \mathbf{B} también y $\mathbf{L} = (1/\omega)\mathbf{D} - \mathbf{E}$.
- Los coeficientes no nulos de \mathbf{L} ocupan la misma posición que los coeficientes no nulos de \mathbf{A} .



Otro Precondicionador B : Cholesky incompleto

- Consiste en utilizar $\mathbf{B} = \mathbf{H}\mathbf{H}^T$, donde $\mathbf{H} = (h_{ij})$ se obtiene aplicando el algoritmo del método de Cholesky, pero sólo para las entradas h_{ij} tales que $a_{ij} \neq 0$. Para las restantes se toma $h_{ij} = 0$.



Otro Precondicionador B : Cholesky incompleto

- Consiste en utilizar $\mathbf{B} = \mathbf{H}\mathbf{H}^T$, donde $\mathbf{H} = (h_{ij})$ se obtiene aplicando el algoritmo del método de Cholesky, pero sólo para las entradas h_{ij} tales que $a_{ij} \neq 0$. Para las restantes se toma $h_{ij} = 0$.
- **Algoritmo de la descomposición incompleta de Cholesky.**

Para $j = 1, \dots, n$

$$h_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} h_{jk}^2}$$

para $i = j + 1, \dots, n$

si $a_{ij} = 0$, entonces $h_{ij} = 0$,

$$\text{si no, } h_{ij} = \frac{1}{h_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} h_{ik}h_{jk} \right)$$



Otro Precondicionador B : Cholesky incompleto

- Consiste en utilizar $\mathbf{B} = \mathbf{H}\mathbf{H}^T$, donde $\mathbf{H} = (h_{ij})$ se obtiene aplicando el algoritmo del método de Cholesky, pero sólo para las entradas h_{ij} tales que $a_{ij} \neq 0$. Para las restantes se toma $h_{ij} = 0$.
- **Algoritmo de la descomposición incompleta de Cholesky.**

Para $j = 1, \dots, n$

$$h_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} h_{jk}^2}$$

para $i = j + 1, \dots, n$

si $a_{ij} = 0$, entonces $h_{ij} = 0$,

si no, $h_{ij} = \frac{1}{h_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} h_{ik}h_{jk} \right)$

- Si $h_{jj} \neq 0$, $j = 1, \dots, n$, entonces \mathbf{H} resulta no singular y $\mathbf{B} = \mathbf{H}\mathbf{H}^T$ puede usarse como precondicionante.



Otro Precondicionador B : Cholesky incompleto

- Consiste en utilizar $\mathbf{B} = \mathbf{H}\mathbf{H}^T$, donde $\mathbf{H} = (h_{ij})$ se obtiene aplicando el algoritmo del método de Cholesky, pero sólo para las entradas h_{ij} tales que $a_{ij} \neq 0$. Para las restantes se toma $h_{ij} = 0$.
- **Algoritmo de la descomposición incompleta de Cholesky.**

Para $j = 1, \dots, n$

$$h_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} h_{jk}^2}$$

para $i = j + 1, \dots, n$

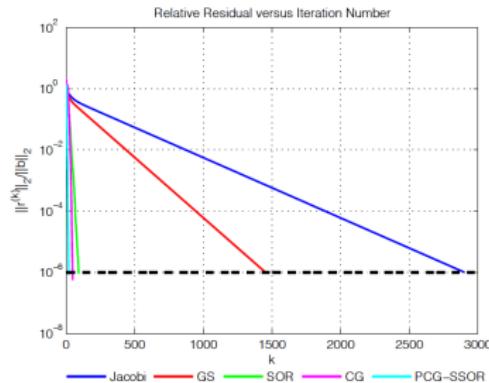
si $a_{ij} = 0$, entonces $h_{ij} = 0$,

$$\text{si no, } h_{ij} = \frac{1}{h_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} h_{ik}h_{jk} \right)$$

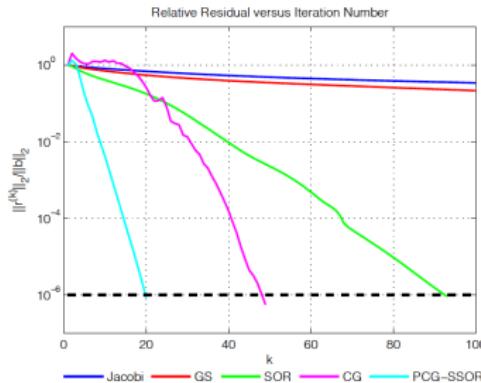
- Si $h_{jj} \neq 0$, $j = 1, \dots, n$, entonces \mathbf{H} resulta no singular y $\mathbf{B} = \mathbf{H}\mathbf{H}^T$ puede usarse como precondicionante.
- Si \mathbf{A} no tuviera entradas nulas, la descomposición de Cholesky sería completa y, en tal caso, $\mathbf{B} = \mathbf{H}\mathbf{H}^T = \mathbf{A}$. Para matrices dispersas, esto no es cierto, pero usualmente $\mathbf{B} \approx \mathbf{A}$.



Comparación gráfica de algunos métodos iterativos¹



(a)



(b)

Figure 5.2: The relative residual versus iteration number k for $N = 32$ for all iterative methods where “GS” represents Gauss-Seidel, “SOR” represents successive overrelaxation, “CG” represents conjugate gradient, and “PCG-SSOR” represents conjugate gradient with symmetric successive overrelaxation preconditioning. Figure (a) shows the graph to the maximum number of iterations required for all methods and (b) shows the plot to $k = 100$. The dashed black line represents the desired tolerance, 10^{-6} . These figures were produced in Matlab.

¹Sarah Swatski, S. Khuvis, M. Gobbert, *A Comparison of Solving the Poisson Equation Using Several Numerical Methods in Matlab and Octave on the Cluster maya*, Technical Report 2011
DOI:10.13016/M27S7HW7N



El Método CGNR

- Si A es no simétrica definida positiva, se puede aplicar el método de GC al sistema equivalente de ecuaciones normales:

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}.$$



El Método CGNR

- Si A es no simétrica definida positiva, se puede aplicar el método de GC al sistema equivalente de ecuaciones normales:

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}.$$

- Este método conocido como CGNR minimiza:

$$\begin{aligned}\|\bar{\mathbf{x}} - \mathbf{x}\|_{\mathbf{A}^T \mathbf{A}} &= (\bar{\mathbf{x}} - \mathbf{x})^T \mathbf{A}^T \mathbf{A} (\bar{\mathbf{x}} - \mathbf{x}) \\ &= (\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x})^T (\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x}) \\ &= (\mathbf{b} - \mathbf{A}\mathbf{x})^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = \|\mathbf{r}\|_2^2.\end{aligned}$$



El Método CGNR

- Si A es no simétrica definida positiva, se puede aplicar el método de GC al sistema equivalente de ecuaciones normales:

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}.$$

- Este método conocido como CGNR minimiza:

$$\begin{aligned}\|\bar{\mathbf{x}} - \mathbf{x}\|_{\mathbf{A}^T \mathbf{A}} &= (\bar{\mathbf{x}} - \mathbf{x})^T \mathbf{A}^T \mathbf{A} (\bar{\mathbf{x}} - \mathbf{x}) \\ &= (\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x})^T (\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x}) \\ &= (\mathbf{b} - \mathbf{A}\mathbf{x})^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = \|\mathbf{r}\|_2^2.\end{aligned}$$

- Es decir GCNR escoge \mathbf{x}_k en $\mathbf{x}_0 + K_k(\mathbf{A}^T \mathbf{A}, \mathbf{r}_0)$ de modo que se minimiza el residual $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$.



El Método CGNR

- Si A es no simétrica definida positiva, se puede aplicar el método de GC al sistema equivalente de ecuaciones normales:

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}.$$

- Este método conocido como CGNR minimiza:

$$\begin{aligned}\|\bar{\mathbf{x}} - \mathbf{x}\|_{\mathbf{A}^T \mathbf{A}} &= (\bar{\mathbf{x}} - \mathbf{x})^T \mathbf{A}^T \mathbf{A} (\bar{\mathbf{x}} - \mathbf{x}) \\ &= (\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x})^T (\mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x}) \\ &= (\mathbf{b} - \mathbf{A}\mathbf{x})^T (\mathbf{b} - \mathbf{A}\mathbf{x}) = \|\mathbf{r}\|_2^2.\end{aligned}$$

- Es decir GCNR escoge \mathbf{x}_k en $\mathbf{x}_0 + K_k(\mathbf{A}^T \mathbf{A}, \mathbf{r}_0)$ de modo que se minimiza el residual $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$.
- De aquí el nombre CGNR que significa CG aplicado a ecuaciones Normales para minimizar el Residual.



El Método CGNR

- Si A es no simétrica definida positiva, se puede aplicar el método de GC al sistema equivalente de ecuaciones normales:

$$A^T A x = A^T b.$$

- Este método conocido como CGNR minimiza:

$$\begin{aligned}\|\bar{x} - x\|_{A^T A} &= (\bar{x} - x)^T A^T A (\bar{x} - x) \\ &= (A \bar{x} - Ax)^T (A \bar{x} - Ax) \\ &= (b - Ax)^T (b - Ax) = \|r\|_2^2.\end{aligned}$$

- Es decir GCNR escoge x_k en $x_0 + K_k(A^T A, r_0)$ de modo que se minimiza el residual $\|b - Ax\|_2$.
- De aquí el nombre CGNR que significa CG aplicado a ecuaciones Normales para minimizar el Residual.
- Si A es rectangular también sirve para resolver problemas de mínimos cuadrados.



Algoritmo del Método CGNR

Dado el vector inicial x_0 ,

$$s_0 = b - Ax_0,$$

$$p_0 = r_0 = A^T s_0,$$

para $k = 0, 1, 2, \dots$

$$q_k = Ap_k,$$

$$\alpha_k = \frac{r_k^T r_k}{q_k^T q_k},$$

$$x_{k+1} = x_k + \alpha_k p_k,$$

$$s_{k+1} = s_k - \alpha_k q_k,$$

$$r_{k+1} = A^T s_{k+1},$$

$$\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k},$$

$$p_{k+1} = r_{k+1} + \beta_k p_k,$$

hasta que se satisfaga un criterio de detención,
o bien hasta que $p_{k+1} = 0$.

- **Algoritmo:**



El Método CGNE

- La segunda forma de abordar problemas lineales no singulares no simétricos es la conocida como CGNE, que significa CG aplicado a ecuaciones Normales para minimizar el Error.



El Método CGNE

- La segunda forma de abordar problemas lineales no singulares no simétricos es la conocida como CGNE, que significa CG aplicado a ecuaciones Normales para minimizar el Error.
- En este caso el sistema que se resuelve por CG es el sistema

$$\mathbf{A}\mathbf{A}^T \mathbf{y} = \mathbf{b}$$

y luego se hace $\mathbf{x} = \mathbf{A}^T \mathbf{y}$.



El Método CGNE

- La segunda forma de abordar problemas lineales no singulares no simétricos es la conocida como CGNE, que significa CG aplicado a ecuaciones Normales para minimizar el Error.
- En este caso el sistema que se resuelve por CG es el sistema

$$\mathbf{A}\mathbf{A}^T \mathbf{y} = \mathbf{b}$$

y luego se hace $\mathbf{x} = \mathbf{A}^T \mathbf{y}$.

- La naturaleza del nombre de la estrategia se debe a que en este caso la propiedad de minimización es de la forma

$$\begin{aligned}\|\bar{\mathbf{y}} - \mathbf{y}\|_{\mathbf{A}\mathbf{A}^T} &= (\bar{\mathbf{y}} - \mathbf{y})^T \mathbf{A}\mathbf{A}^T (\bar{\mathbf{y}} - \mathbf{y}) \\ &= (\mathbf{A}^T \bar{\mathbf{y}} - \mathbf{A}^T \mathbf{y})^T (\mathbf{A}^T \bar{\mathbf{y}} - \mathbf{A}^T \mathbf{y}) \\ &= (\bar{\mathbf{x}} - \mathbf{x})^T (\bar{\mathbf{x}} - \mathbf{x}) = \|\mathbf{e}\|_2^2.\end{aligned}$$



El Método CGNE

- La segunda forma de abordar problemas lineales no singulares no simétricos es la conocida como CGNE, que significa CG aplicado a ecuaciones Normales para minimizar el Error.
- En este caso el sistema que se resuelve por CG es el sistema

$$\mathbf{A}\mathbf{A}^T \mathbf{y} = \mathbf{b}$$

y luego se hace $\mathbf{x} = \mathbf{A}^T \mathbf{y}$.

- La naturaleza del nombre de la estrategia se debe a que en este caso la propiedad de minimización es de la forma

$$\begin{aligned}\|\bar{\mathbf{y}} - \mathbf{y}\|_{\mathbf{A}\mathbf{A}^T} &= (\bar{\mathbf{y}} - \mathbf{y})^T \mathbf{A}\mathbf{A}^T (\bar{\mathbf{y}} - \mathbf{y}) \\ &= (\mathbf{A}^T \bar{\mathbf{y}} - \mathbf{A}^T \mathbf{y})^T (\mathbf{A}^T \bar{\mathbf{y}} - \mathbf{A}^T \mathbf{y}) \\ &= (\bar{\mathbf{x}} - \mathbf{x})^T (\bar{\mathbf{x}} - \mathbf{x}) = \|\mathbf{e}\|_2^2.\end{aligned}$$

- Tanto el método CGNR como CGNE tienen condicionamiento del orden de $\text{cond}_2(\mathbf{A})^2$, con lo cual convergen muy lentamente.

