Taller de módelos y métodos
Aplicación de precondicionadores

Ariel Sanhueza - *asanhuez@alumnos.inf.utfsm.cl* - *201173005-4*
Axel Símonsen - *axel.simonsen@alumnos.usm.cl* - *201173007-0*

Valparaíso, 4 de diciembre de 2015

Resumen

Al resolver el sistema de ecuaciones de la forma $A\vec{x} = \vec{b}$, dependemos en gran medida de la calidad de la matriz. Cuando se da el caso de que la matriz está mal condicionada, el error en la resolución del sistema de ecuaciones crece. Cuando esto ocurre, necesitamos encontrar un modo de mejorar nuestra matriz con el objetivo de disminuir su número de condición. A esta solución se le llama *Precondicionador*, el cual es una matriz aplicada a la matriz A, que mejora su número de condición. Para este caso éste precondicionador es de la forma f(M), donde f es una función dada y M una matriz obtenida mediante un shift a la matriz A.

1. Desarrollo del problema

Sea un sistema de ecuaciones de la forma $Ax = b$, con ~~$A \in \mathcal{C}_{nxn}, x, b \in \mathcal{C}_{nx1}$~~$A \in \mathbb{C}^{nxn}, x, b \in \mathbb{C}^{n}$. Sea $\mathcal{K}(A)$ el número de condición de la matriz A, donde $\mathcal{K}(A) >> 1$, es decir, la matriz está mal condicionada, lo que significa que el sistema es más sensible a los errores.

El parámetro presentado para medir la ~~eficiencia de GMRes~~ convergencia de GMRes para matrices positivas definidas $(X^T AX > 0)$ es un valor ρ~~definido~~ , el cual depende de la matriz sobre la que se calcule. Se define como:

$$\rho = \frac{\lambda_M - \lambda_m + 2\delta}{\lambda_M + \lambda_m + 2\sqrt{\lambda_M \lambda_m + \delta^2}}$$

Donde:
λ_M = mayor ~~$Re(\lambda)$~~Re(λ)
λ_m = menor ~~$Re(\lambda)$~~Re(λ)
δ = mayor ~~$Im(\lambda)$~~Im(λ)

Aquí la idea es minimizar el error representado como $\|\vec{b} - A\vec{x_n}\| \leq \mathcal{C}\rho^n$.

El problema mencionado actualmente en [?] , propone encontrar una matriz M, ~~la cual~~ con la cual, mediante una función propuesta (actualmente $f(x) = \frac{1}{x}$), obtendremos el precondicionador $f(M)$. Esta matriz no es más que un corrimiento en α de la diagonal y sus valores propios. Ésta matriz M se define como:

$$M = \alpha I + A \tag{1}$$

Para el problema planteado en ~~el paper, se buscan los α tal que la siguiente expresión sea minimizada dado que se requiere minimizar la tasa de convergencia para el problema~~ [?] , se propone la función antes mencionada $f(x) = \frac{1}{x}$, mediante la cual crearemos el precondicionador $f(M)$, creando un nuevo sistema de ecuaciones de la forma $f(M)A\vec{x} = f(M)\vec{b}$, donde $f(M)$ es el precondicionador actualmente propuesto, el cual se explicará mas adelante:

$$\cancel{\rho_{f(M)A}(\alpha) = \frac{\lambda_M(\lambda_M+\alpha)^{-1} - \lambda_m(\lambda_m+\alpha)^{-1} + 2\delta}{\lambda_M(\lambda_M+\alpha)^{-1} + \lambda_m(\lambda_m+\alpha)^{-1} + 2\alpha + 2\sqrt{(\lambda_M(\lambda_M+\alpha)^{-1}+\alpha)(\lambda_m(\lambda_m+\alpha)^{-1}+\alpha) + \delta^2}}}$$

~~La idea es encontrar un α que minimice esta expresión, tal que $\rho_{f(M)A}(\alpha) \leq \rho$, con el objetivo de disminuir la tasa de convergencia de GMRes (método utilizado en el *paper* para resolver el sistema). Sin embargo, el parámetro no puede ser muy pequeño, pues se asemeja a no hacer nada, y no puede ser muy grande, dado que no crearíamos un buen precondicionador.~~

Se propone actualmente una matriz M, cuya inversa sea cercana a la inversa de A, con el objetivo de llegar a algo similar a la identidad, sin tener que calcular la inversa de A, cuyo costo computacional es alto.

El problema presentado original propone como precondicionador la evaluación de la matriz M en la función $f(x) = \frac{1}{x}$. Esta función matricial es definida por medio de la fórmula integral de Cauchy:

$$f(M) = \frac{1}{2\pi i}\oint_C f(z)(zI - M)^{-1}dz$$

donde $f(z)$ es una función analítica (~~en el problema presentado,~~ vale decir que sea representable como una serie de potencias convergente) y $f(x) = x^{-1}$, C una elipse que encierre todos los valores propios de la matriz M. Esta definición nos permite saber el valor de una función matricial sin la necesidad de evaluar *directamente* la matriz en la función y se reduce finalmente al cálculo de una integral de línea cerrada sobre la curva C.

A pesar de la definición previa, nosotros no queremos computar explícitamente $f(M)$, si no que $f(M)\vec{v}$, lo que queda finalmente:

$$f(M)\vec{v} = \frac{1}{2\pi i}\oint_C f(z)(zI - M)^{-1}\vec{v}dz$$

$$f(M)\vec{v} = \frac{1}{2\pi i}\oint_C f(z)(zI - \alpha * I - A)^{-1}\vec{v}dz$$

$$f(M)\vec{v} = \frac{1}{2\pi i}\oint_C f(z)((z-\alpha)I - A)^{-1}\vec{v}dz$$

que se reduce finalmente a resolver el sistema de ecuaciones $(Iz-M)\vec{h} = \vec{v}$ por medio de GMres para obtener $\vec{h} = (zI - M)^{-1}\vec{v}$. La integral finalmente se discretiza para poder generar una aproximación de su valor (por ejemplo, utilizando el método del trapezoide).

Nuestro trabajo es ~~buscar otro precondicionador , diferente al propuesto, que permita mejorar aun más~~ implementar el precondicionador que se propone, el cual permita mejorar el cálculo del sistema de ecuaciones. ~~Se~~ Si se consiguen buenos resultados en poco tiempo, se buscará otra función analítica $f(x)$ que no solo mejore el número de condición de la matriz A, si no que también sea mejor que el precondicionador propuesto.

2. Plan de trabajo:

Se propone como primer objetivo ~~encontrar una función que mejore la ya propuesta en el paper, y a partir de esta construir un algoritmo que dada~~ implementar el algoritmo GMRes, con el uso del precondicionador utilizando la función dada para reparar una matriz mal condicionada, ~~encuentre~~ encontrando el parámetro α para construir M y ~~construya~~ construyendo la matriz $f(M)$, con la cual se pueda mejorar la resolución del sistema $A\vec{x} = \vec{b}$, transformándolo en el sistema $f(M)A\vec{x} = f(M)\vec{b}$. Para medir que tan bueno es nuestro precondicionador, utilizaremos los valores propios de la matriz original, aplicándoles la función seleccionada, y con ellos calculando el número de condición, el cual servirá como parámetro de medida de la eficacia del precondicionador utilizado, junto con el calculo del ~~parametro~~ parámetro ρ, para el precondicionador seleccionado por la matriz ingresada. Todo lo anterior estará apoyado en los papers sugeridos e investigación independiente.

3. Avances al 4/12:

- Precondicionador por la izquierda (Sin Integral de contorno)
- Precondicionador por la derecha (Sin Integral de contorno)
- Calculo de la integral de contorno
- Avances en precondicionador de contorno

3.1. Precondicionador por la izquierda (sin integral de contorno)

Para poder evaluar las diferencias de rendimiento, se implementaron precondicionadores sin utilizar la integral de contorno. Esto significa que se aplica como precondicionador la matriz M^{-1} (no calculándolo directamente). Entonces nuestro sistema de ecuaciones pasaría a:

$$\begin{aligned} A\vec{x} &= \vec{b} \\ M^{-1}A\vec{x} &= M^{-1}\vec{b} \\ \widetilde{A}\vec{x} &= \widetilde{b} \end{aligned}$$

Definiendo así $\widetilde{A} = M^{-1}A$ y $\widetilde{b} = M^{-1}\vec{b}$. Claramente no debemos calcular M^{-1}, por lo que utilizaremos GMRes para solucionar estos problemas. Para $\widetilde{A} = M^{-1}A$ basta que cuando se multiplique por un vector $\vec{q}$, tenemos:

$$\begin{aligned} \widetilde{A} &= M^{-1}A \\ \widetilde{A}\vec{q} &= M^{-1}(A\vec{q}) \\ \widetilde{A}\vec{q} &= M^{-1}\widetilde{q} \end{aligned}$$

Entonces $\widetilde{A}\vec{q}$ es igual a $\vec{h} = M^{-1}\widetilde{q}$, con $\widetilde{q} = A\vec{q}$ que no es más que resolver el sistema de ecuaciones lineales $M\vec{h} = \widetilde{q}$. Para calcular $\widetilde{b} = M^{-1}\vec{b}$, al igual que antes, hay que resolver el sistema de ecuaciones lineales $M\widetilde{b} = \vec{b}$.

3.2. Precondicionador por la derecha (Sin integral de contorno)

Dado el cálculo del sistema de ecuaciones aplicando el precondicionador por la izquierda, para poseer una comparación se implementó el precondicionador multiplicando al sistema por la derecha con el objetivo de comprobar como cambia el sistema de ecuaciones sobre el cual estamos trabajando, mediante el siguiente proceso, recalcando que todo esto se realiza sin el calculo de $M^{-}1$:

$$
\begin{aligned}
A\vec{x} &= \vec{b} \\
AM^{-1}M\vec{x} &= \vec{b} \\
\widetilde{A}\widetilde{x} &= \vec{b}
\end{aligned}
$$

Definiendo así $\widetilde{A} = AM^{-1}$ y $\widetilde{x} = M\vec{x}$. Como ya sabemos, no debemos calcular M^{-1}, por lo que utilizaremos GMRes para llegar a $\widetilde{A}$. Para $\widetilde{A} = AM^{-1}$ al multiplicar por un vector $\vec{q}$ (el cual puede ser el vector inicial del GMRes), tenemos:

$$
\begin{aligned}
\widetilde{A} &= AM^{-1} \\
\widetilde{A}\vec{q} &= AM^{-1}\vec{q} \\
\widetilde{A}\vec{q} &= A\widetilde{q} \\
\widetilde{A}\vec{q} &= \vec{h}
\end{aligned}
$$

Aquí finalmente, iterativamente se calcula el producto $\widetilde{A}\vec{q}$, el cual es utilizado en el algoritmo GMRes. Por lo tanto, tenemos una versión de GMRes com precondicionador por la derecha, pero en una versión mas primitiva sin el uso de la integral de contorno.

3.3. Precondicionador de Contorno

Ahora, a diferencia de la implementación sin integral de contorno del precondicionador por la izquierda, es que el cálculo de este es distinto: acá buscamos calcular $f(M)$, donde M es definido al igual que antes y f es una función analítica dentro de un contorno Γ. En el paper leído, proponen una aproximación de la función $f(x) = \dfrac{1}{x}$, que queda definido como $f(z) = \dfrac{1-(\frac{\alpha}{z})^{N_f+1}}{z-\alpha}$. Así, por la *Cauchy's integral formula*, la definición de $f(M)$ queda:

$$
\begin{aligned}
f(M) &= \frac{1}{2\pi i}\oint_{\Gamma} f(z)(zI-M)^{-1}dz \\
f(M)\vec{q} &= \frac{1}{2\pi i}\oint_{\Gamma} f(z)(zI-M)^{-1}\vec{q}dz
\end{aligned}
$$

El vector $\vec{q}$ que multiplica a $f(A)$ es usado porque no necesitamos $f(A)$, solo necesitamos el resultado multiplicado por ese vector. Así, encontrar el valor de $\vec{h} = (zI-M)^{-1}\vec{q}$ es resolver el sistema de ecuaciones lineales $(zI-M)\vec{h} = \vec{q}$. Gracias a las propiedades de la integral de Cauchy, podemos cambiar el contorno de integración siempre que encierre al punto al que estamos evaluando (para el caso de la integral con complejos) pero para el caso de nuestra función matricial el nuevo contorno será un círculo de radio R suficientemente grande para encerrar todos los valores propios de A y un centro C. Parametrizamos la integral por $z(t) = C + Re^{it} = C + R(\cos(t) + i\sin(t))$ y tenemos que $\frac{dz}{dt} = iRe^{it}$ con $t \in [0, 2\pi]$. Así, nuestra integral de contorno queda como:

$$f(M)\vec{q} = \frac{1}{2\pi i}\oint_{\Gamma} f(z)(zI - M)^{-1}\vec{q}dz$$

$$f(M)\vec{q} = \frac{1}{2\pi i}\int_0^{2\pi} f(z(t))(z(t)I - M)^{-1}\vec{q}\frac{dz}{dt}dt$$

$$f(M)\vec{q} = \frac{1}{2\pi i}\int_0^{2\pi} f(z(t))(z(t)I - M)^{-1}\vec{q}iRe^{it}dt$$

$$f(M)\vec{q} = \frac{1}{2\pi}\int_0^{2\pi} f(z(t))(z(t)I - M)^{-1}\vec{q}Re^{it}dt$$

La integral final puede ser calculada con algún método de integración numérica, en nuestro caso, usaremos la regla del trapezoide.

3.4. Acotaciones:

Para el calculo con el precondicionador de contorno, se ha podido notar que el algoritmo es extremadamente sensible en cuanto los siguientes parámetros, en comparación a las versiones sin la integral de contorno, lo cual no debiera tener sentido.

Parámetros:

1. Precisión de GMRes, dentro de la integración numérica
2. Alpha

Variaciones en estos parámetros, hacen que el método de soluciones correctas o erróneas para distintos sistemas, por lo que parte del trabajo futuro es analizar los porqués de estas variaciones.