# Métodos de Krylov II: MINRES, FOM, GMRES.

Alan Reyes-Figueroa
Métodos Numéricos II (Aula 11) 11.agosto.2021

# Métodos de Krylov

El método de Arnoldi y el método de Lanczos proporcionan una técnica económica para calcular vectores base ortogonales para el subespacio de Krylov $\mathcal{K}_k(A, \mathbf{r}^{(0)})$.

Vimos que la aproximación $\mathbf{x}^{(k)}$ se escribe como

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k \mathbf{y}^{(k)},$$

donde $\mathbf{y}^{(k)}$ se determina de forma que o bien se minimiza

$$f(\mathbf{x}^{(k)}) = ||\mathbf{x}^{(k)} - \mathbf{x}||_A^2 = (\mathbf{x}^{(k)} - \mathbf{x})^T A (\mathbf{x}^{(k)} - \mathbf{x}), \tag{1}$$

respecto de la norma inducida por $A$ ($A$ simétrica y positiva definida) o bien minimizan

$$g(\mathbf{x}^{(k)}) = ||A(\mathbf{x}^{(k)} - \mathbf{x})||_2^2 = \mathbf{r}^{(k)T} \mathbf{r}^{(k)}, \tag{2}$$

esto es, se minimiza la norma del residuo.

Veremos ahora algunos detalles desde este punto de vista de minimizar residuos.

# Métodos de Krylov

Si se considera la minimización del error en la $A$-norma:

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)} \qquad f(\mathbf{x}^{(k)}) = (\mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)} - \mathbf{x})^T A(\mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)} - \mathbf{x}).$$

Calculando la derivada respecto de $\mathbf{y}^{(k)}$, resulta

$$\nabla_{\mathbf{y}^{(k)}} f(\mathbf{x}^{(k)}) = 2Q_k^T A(\mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)} - \mathbf{x}) = \mathbf{0},$$

de modo que $Q_k^T A Q_k \mathbf{y}^{(k)} = Q_k^T A(\mathbf{x} - \mathbf{x}^{(0)}) = Q_k^T A\mathbf{r}^{(0)}$.

Como $Q_k^T A Q_k = H_k$, y $\mathbf{r}^{(0)} = ||\mathbf{r}^{(0)}||\mathbf{q}_1$, se tiene que

$$H_k\mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||\mathbf{e}_1,$$

donde $\mathbf{e}_1$ es el primer vector de la base canónica de $\mathbb{R}^n$.

A partir de lo anterior, se deduce que los residuos son orgonales a los vectores base

$$\mathbf{r}^{(k)} = \mathbf{r}^{(0)} - AQ_k\mathbf{y}^{(k)} \quad \Longrightarrow \quad Q_k^T\mathbf{r}^{(k)} = Q_k^T\mathbf{r}^{(0)} - Q_k^T A Q_k\mathbf{y}^{(k)} = \mathbf{0}.$$

Esta condición es equivalente a minimizar $f(\mathbf{x}^{(k)})$ cuando $A$ es simétrica y positiva definida. En este caso se obtiene el llamado **método del gradiente conjugado**.

El gradiente conjugado minimiza la $A$-norma del error.
Otra forma de construir una aproximación óptima $\mathbf{x}^{(k)}$ es minimizar el residuo (2)

$$g(\mathbf{x}^{(k)}) = ||A(\mathbf{x}^{(k)} - \mathbf{x})||_2^2 = \mathbf{r}^{(k)T}\mathbf{r}^{(k)},$$

sobre todos los $\mathbf{x}^{(k)} \in \{\mathbf{x}^{(0)}\} \cup \mathcal{K}(A, \mathbf{b})$.

Definiendo, como en el método de Lanczos

$$H_k = \begin{pmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_k \\ & & \beta_k & \alpha_k \end{pmatrix}.$$

Tenemos que $AQ_k = Q_{k+1}H_k$.

Ahora, la minimización se reduce a encontrar $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)}$ tal que $||\mathbf{r}^{(k)}||_2$ sea mínima. Como

$$\mathbf{r}^{(k)} = b - A\mathbf{x}^{(k)} = \mathbf{r}^{(0)} - AQ_k\mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||\mathbf{q}_1 - AQ_k\mathbf{y}^{(k)},$$

entonces se debe minimizar

$$||\mathbf{r}^{(k)}||_2 = ||\,||\mathbf{r}^{(0)}||\mathbf{q}_1 - AQ_k\mathbf{y}^{(k)}|| = ||\,||\mathbf{r}^{(0)}||Q_{k+1}\mathbf{e}_1 - Q_{k+1}H_k\mathbf{y}^{(k)}|| = ||\,||\mathbf{r}^{(0)}||\mathbf{e}_1 - H_k\mathbf{y}^{(k)}||.$$

Resolviendo el sistema sobredeterminado $H_k\mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||\mathbf{e}_1$ se obtienen las iteraciones

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)}$$

que minimizan el residuo. El algoritmo resultante se llama **MINRES** (o residuo conjugado).

**Algoritmo:** (MINRES).
*Inputs*: $A \in \mathbb{R}^{n\times n}$, simétrica y positiva definida.
*Outputs*: Secuencia de residuos $\{\mathbf{r}^{(k)}\}$, $H \in \mathbb{R}^{n\times n}$ en la forma de Hessemberg.
`Initialize` $H = \mathbf{0}$, choose $\mathbf{r}^0 = \mathbf{b} - A\mathbf{x}^{(0)}$, $\mathbf{p}_0 = \mathbf{r}^{(0)}$.
`For` $k = 1, 2, \ldots$:

$$\alpha_k = \frac{(\mathbf{r}^k)^T A\mathbf{r}^k}{(A\mathbf{p}_k)^T(A\mathbf{p}_k)}.$$
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{p}_k,$$
$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A\mathbf{p}_k,$$
$$\beta_k = \frac{(\mathbf{r}^{(k+1)})^T A\mathbf{r}^{(k+1)}}{(\mathbf{r}^{(k)})^T A\mathbf{r}^{(k)}},$$
$$H_{k+1,k} = ||\mathbf{v}||_2,$$
$$\mathbf{p}_{k+1} = \mathbf{r}^{(k+1)} + \beta_k \mathbf{r}^{(k)},$$
$$A\mathbf{p}_{k+1} = A\mathbf{r}^{(k+1)} + \beta_k A\mathbf{r}^{(k)}.$$

# Métodos de Krylov

Al igual que el método de Lanczos, el método de Arnoldi nos da una base ortogonal para el subespacio de Krylov $\mathcal{K}(A, \mathbf{r}^{(0)})$, esta vez para $A$ no simétrica.

Al igual que antes, se tiene la aproximación

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k \mathbf{y}^{(k)},$$

donde $\mathbf{y}^{(k)}$ es tal que se minimiza el error (1)

$$f(\mathbf{x}^{(k)}) = ||\mathbf{x}^{(k)} - \mathbf{x}||_A^2 = (\mathbf{x}^{(k)} - \mathbf{x})^T A(\mathbf{x}^{(k)} - \mathbf{x}),$$

respecto de la norma inducida por $A$, o bien se minimiza el residuo (2)

$$g(\mathbf{x}^{(k)}) = ||A(\mathbf{x}^{(k)} - \mathbf{x})||_2^2 = \mathbf{r}^{(k)T}\mathbf{r}^{(k)}.$$

Cuando $A$ no es simétrica y positiva definida, la $A$-norma no está bien definida. Imponiendo ahora que los residuos sean ortogonales a $Q_k$, se tiene

$$Q_k^T(\mathbf{r}^{(0)} - AQ_k\mathbf{y}^{(k)}) = \mathbf{0} \quad \Longrightarrow \quad ||\mathbf{r}^{(0)}||\mathbf{e}_1 - H_k\mathbf{y}^{(k)} = \mathbf{0}.$$

Resolviendo el sistema $\mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||H_k^{-1}\mathbf{e}_1$, $\mathbf{x}^{(k)} = Q_k\mathbf{y}^{(k)}$, se obtiene el llamado **método FOM** (*Full Orthogonalization Method*).

# Métodos de Krylov

Observaciones:

- FOM es equivalente al gradiente conjugado si $A$ es simétrica y positiva definida. Pero tiene inconvenientes:
- FOM no trata de forma eficiente la memoria: $Q_k$ se tiene que guardar completa. En cada iteración, un nuevo vector base se tiene que calcular y guardar.
- FOM no tiene una propiedad de optimalidad.
- FOM no es robusto ya que $H_k$ puede ser no singular.

El método FOM para obtener $\mathbf{x}^{(k)}$ es parte de una familia de técnicas para extraer una solución aproximada a partir de un espacio de búsqueda $Q_k$ haciendo que el residuo sea ortogonal a un espacio test $W_k$, así: Dada $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)}$, se busca $\mathbf{y}^{(k)}$ tal que

$$W_k^T(\mathbf{r}^{(0)} - AQ_k\mathbf{y}^{(k)}) = \mathbf{0}. \tag{3}$$

Estas son las condiciones de Petrov-Galerkin.



UNIVERSIDAD DEL VALLE DE GUATEMALA

# GMRES

Un segundo método para obtener un mínimo del residuo es el siguiente. Observe que el problema de minimizar

$$g(\mathbf{x}^{(k)}) = ||A(\mathbf{x}^{(k)} - \mathbf{x})||_2^2 = (\mathbf{r}^{(k)})^T\mathbf{r}^{(k)},$$

está siempre bien definido incluso si $A$ es no simétrica.

El problema es encontrar $\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)}$ tal que $||\mathbf{r}^{(k)}||_2^2$ es mínimo. Como $\mathbf{r}^{(k)} = b - A\mathbf{x}^{(k)} = \mathbf{r}^{(0)} - AQ_k\mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||\mathbf{q}_1 - AQ_k\mathbf{y}^{(k)}$, tenemos

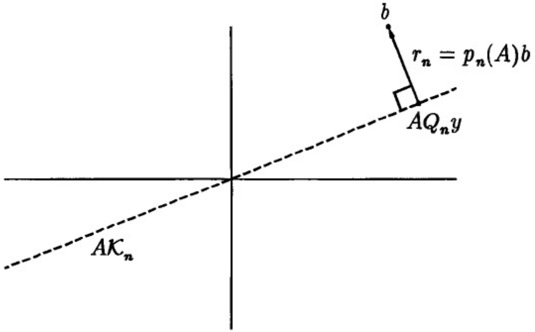
$$||\mathbf{r}^{(k)}||_2 = \Big|\Big|\,||\mathbf{r}^{(0)}||\mathbf{q}_1 - AQ_k\mathbf{y}^{(k)}\Big|\Big| = \Big|\Big|\,||\mathbf{r}^{(0)}||Q_{k+1}\mathbf{e}_1 - Q_{k+1}H_k\mathbf{y}^{(k)}\Big|\Big| = \Big|\Big|\,||\mathbf{r}^{(0)}||\mathbf{e}_1 - H_k\mathbf{y}^{(k)}\Big|\Big|.$$

Al resolver el sistema sobredeterminado $H_k\mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||\mathbf{e}_1$, se obtienen las iteraciones

$$\mathbf{x}^{(k)} = \mathbf{x}^{(0)} + Q_k\mathbf{y}^{(k)}$$

que minimizan el residuo.

El algoritmo resultatnte se llama **GMRES** (*General Minimal Residual Method*). GMRES, es uno de los métodos más populares para resolver sistemas no simétricos.




GMRES minimiza la norma del residuo $||\mathbf{r}^{(n)}||$.

Observaciones:

- El método GMRES es equivalente al MINRES, cuando $A$ es simétrica.
- GMRES no tiene limitada la memoria: $Q_k$ se ha de guardar completamente. En cada iteración se ha de calcular y guardar un nuevo vector de la base. La ortogonalización de un nuevo vector se hace cada vez más cara al aumentar $k$.
- GMRES minimiza la norma del residuo.
- GMRES es robusto: el sistema $H_k \mathbf{y}^{(k)} = ||\mathbf{r}^{(0)}||\mathbf{e}_1$ tiene siempre una solución de mínimos cuadrados.

**Algoritmo:** (GMRES)
$\mathbf{q}_1 = \mathbf{b}/||\mathbf{b}||$,
For $k = 1, 2, 3, \ldots$
    (step $k$ of Arnoldi iteration)
    Find $\mathbf{y}^{(k)}$ to minimize $||H_k \mathbf{y}^{(k)} - ||\mathbf{b}||\mathbf{e}_1|| = ||\mathbf{r}^{(k)}||$,
    $\mathbf{x}^{(k)} = Q_k \mathbf{y}^{(k)}$.

# Métodos de Krylov

Otros métodos:

- Bi-Lanczos,
- BICG (Bi-graiente conjugado),
- QMR
- Algoritmo GC cuadrado
- Método Bi-CGSTAB