Análisis de convergencia del método del subespacio iterativo

Pedro Massey

Centro de Matemática de La Plata, FCE - UNLP &
Instituto Argentino de Matemática “Alberto Calderón” - CONICET

Reunión anual de la UMA- Neuquén - Septiembre 2022

Aproximaciones de rango bajo

Problema: sea $A \in \mathbb{C}^{m\times n}$ y $1 \le h \le p := \min\{m, n\}$: *calcular* $\hat{A} \in \mathbb{C}^{m\times n}$ tal que $\operatorname{rank}(\hat{A}) \le h$ y $\|A - \hat{A}\|$ sea *lo más chica posible*.

Sea $A = U\Sigma V^*$ una DVS: $U \in \mathbb{C}^{m\times m}$ y $V \in \mathbb{C}^{n\times n}$ unitarias, $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{C}^{m\times n}$ con $\sigma_1 \ge \ldots \ge \sigma_p$ (valores singulares).

Si $\Sigma_h := \operatorname{diag}(\sigma_1, \ldots, \sigma_h, \underbrace{0, \ldots, 0}_{p-h}) \in \mathbb{C}^{m\times n}$:

$$\|A - \underbrace{U\Sigma_h V^*}_{A_h}\| = \sigma_{h+1} \le \|A - \hat{A}\| \qquad (\, A_h \text{ es } \textit{una} \text{ solución óptima}) \,.$$

Observación: si $u_j = C_j(U)$, $1 \le j \le m$:

$$\text{si } \ \mathcal{S} := \operatorname{Span}\{u_1, \ldots, u_h\} \implies A_h = U\Sigma_h V^* = P_{\mathcal{S}}\, A \,.$$

Pero las DVS's así como $\mathcal{S}$ son *costosas de calcular* si p es grande.

Alternativa: aproximar $P_{\mathcal{S}}$ (aproximar un *subespacio dominante* $\mathcal{S}$) !

Un ejemplo ($h = 1$)

Sea $A \in \mathbb{C}^{m\times n}$ y $A = U\Sigma V^*$ una DVS; Si $q \geq 0$ entonces

$$A^*A = (U\Sigma V^*)^*(U\Sigma V^*) = V\,|\Sigma|^2\,V^* \implies$$

$$A(A^*A)^q = U\Sigma V^*(V\,|\Sigma|^{2q}\,V^*) = U\,\Sigma\,|\Sigma|^{2q}\,V^*$$

Si $x \in \mathbb{C}^n$ y $v_j = C_j(V)$, $1 \leq j \leq n$ ($p = \min\{m, n\}$):

$$\begin{aligned}
\underbrace{A(A^*A)^q x}_{(computable)} &= \sum_{j=1}^{p} \sigma_j^{2q+1}\,\langle x, v_j\rangle\, u_j \\
&= \sigma_1^{2q+1}\left(\langle x, v_1\rangle\, u_1 + \sum_{j=2}^{p}\left(\frac{\sigma_j}{\sigma_1}\right)^{2q+1}\langle x, v_j\rangle\, u_j\right)
\end{aligned}$$

$$\text{Si}\quad \langle x, v_1\rangle = \underbrace{v_1^* x \neq 0}_{\text{compatibilidad}} \quad \text{y} \quad \underbrace{\sigma_1 > \sigma_2}_{\text{separación}} \implies \mathbb{C}\cdot A(A^*A)^q x \xrightarrow[q\to\infty]{} \underbrace{\mathbb{C}\cdot u_1}_{\mathcal{S}}$$

Método del subespacio iterativo (simplificado)

Sea $A \in \mathbb{C}^{m\times n}$, sea $X \in \mathbb{C}^{n\times h}$ y sea $q \geq 0$:

Paso 1. Se calcula $M_q = A(A^*A)^q X \in \mathbb{C}^{m\times h}$

Paso 2. Se calcula una b.o. $\{w_1, \ldots, w_t\}$ de $R(M_q) \subset \mathbb{C}^m$ $(t \leq h)$.

Paso 3. Se construye $Q \in \mathbb{C}^{m\times t}$, $C_j(Q) = w_j$, $1 \leq j \leq t$ $(QQ^* = P_{R(M_q)})$.

Teorema (A. Saibaba, SIMAX (2019))

Sea $A = U\Sigma V^*$ una DVS. Sean $V_h \in \mathbb{C}^{n\times h}$ y $V_{h,\perp} \in \mathbb{C}^{n\times(n-h)}$, $V = [V_h \mid V_{h,\perp}]$. Sea $\mathcal{S} = Span\{u_1, \ldots, u_h\}$. Supongamos que

$$\underbrace{V_h^* X \in \mathcal{G}\ell(h)}_{compatibilidad} \quad y \quad \underbrace{\sigma_h > \sigma_{h+1}}_{separación} \implies \quad \mathcal{S} \text{ no depende de la DVS,}$$

1. $\|P_{\mathcal{S}} - QQ^*\| \leq \left(\frac{\sigma_{h+1}}{\sigma_h}\right)^{2q} \left\|\left(V_{h,\perp}^* X\right)\left(V_h^* X\right)^\dagger\right\|$
2. $\|A - QQ^*A\| \leq \sigma_{h+1} + \frac{\left(\frac{\sigma_{h+1}}{\sigma_h}\right)^{2q}}{1-\left(\frac{\sigma_{h+1}}{\sigma_h}\right)} \left\|\left(V_{h,\perp}^* X\right)\left(V_h^* X\right)^\dagger\right\|$

Método del subespacio iterativo: si $\sigma_h = \sigma_{h+1}$?

Sea $A \in \mathbb{C}^{m\times n}$ tal que $\sigma_h = \sigma_{h+1}$. Sea $A = U\Sigma V^*$ una DVS, $u_j = C_j(U)$ y $\mathcal{S} = \text{Span}\{u_1, \ldots, u_h\}$.

$\mathcal{S}$ depende de la DVS considerada!

$$j := \max\{0 \le \ell < h \,:\, \sigma_\ell > \sigma_h\}\,,\ k := \max\{h \le \ell \le \text{rank}(A)\,:\, \sigma_\ell = \sigma_h\}$$

$$\sigma_j > \sigma_{j+1} = \sigma_h = \sigma_k > \sigma_{k+1}\,.$$

Entonces existe $A = U'\Sigma(V')^*$ DVS tal que $\tilde{\mathcal{S}} = \text{Span}\{u'_1, \ldots, u'_h\}$ sii

$$\tilde{\mathcal{S}} = \text{Span}\{u_1, \ldots, u_j\} \oplus \mathcal{T} \quad \text{con} \quad \mathcal{T} \subset \text{Span}\{u_{j+1}, \ldots, u_k\}\,,\ \dim \mathcal{T} = h - j$$

y decimos que $\tilde{\mathcal{S}}$ es un *h-subespacio dominante a derecha*; en este caso

$P_{\tilde{\mathcal{S}}} A$ es una aproximación óptima de A de rank a lo sumo h.

Resultados cuando $\sigma_h = \sigma_{h+1}$

Paso 1. Se calcula $M_q = A(A^*A)^q X \in \mathbb{C}^{m\times h}$

Paso 2. Se calcula una b.o. $\{w_1, \ldots, w_t\}$ de $R(M_q)$ $(t \leq h)$.

Paso 3. Se construye $Q \in \mathbb{C}^{m\times t}$, $C_j(Q) = w_j$, $1 \leq j \leq t$ $(QQ^* = P_{R(M_q)})$.

$j = \max\{0 \leq \ell < h \,:\, \sigma_\ell > \sigma_h\} < h < k = \max\{h \leq \ell \leq \operatorname{rank}(A) \,:\, \sigma_\ell = \sigma_h\}$

$$\sigma_j > \sigma_{j+1} = \ldots = \sigma_h = \ldots = \sigma_k > \sigma_{k+1}$$

Teorema (M, 2022)

Supongamos que $1 \leq j, k < rank(A)$ *y que existe una DVS,* $A = U\Sigma V^*$ *tal que* $V_h^* X \in \mathcal{G}\ell(h)$. *Entonces existe* $\tilde{\mathcal{S}} \subset \mathbb{C}^m$ *un* h*-subespacio dominante a derecha tal que*

$$\|P_{\tilde{\mathcal{S}}} - QQ^*\| \leq 4\left(\frac{\sigma_h}{\sigma_j}\right)^{2q} \|X_{2,j}X_{1,j}^\dagger\| + \left(\frac{\sigma_{k+1}}{\sigma_h}\right)^{2q+1} \|X_{2,k}X_{1,k}^\dagger\|$$

donde $V_j \in \mathbb{C}^{n\times j}$, $V_k \in \mathbb{C}^{n\times k}$: $V = [V_j \,|\, V_{j,\perp}] = [V_k \,|\, V_{k,\perp}]$ *y*

$$X_{1,\ell} = V_\ell^* X\,,\; X_{2,\ell} = V_{\ell,\perp}^* X \quad con \quad \ell = j, k\,.$$

Resultados cuando $\sigma_h = \sigma_{h+1}$ (II)

$A \in \mathbb{C}^{m\times n}$, $X \in \mathbb{C}^{n\times h}$; Q se construye con q iteraciones del método del subespacio iterativo $(QQ^* = P_{R(A(A^*A)^q X)})$.

$$j = \max\{0 \le \ell < h \,:\, \sigma_\ell > \sigma_h\} < h < k = \max\{h \le \ell \le \operatorname{rank}(A) \,:\, \sigma_\ell = \sigma_h\}$$

Teorema (M, 2022)

Supongamos que $1 \le j < k < rank(A)$ *y que existe una DVS,* $A = U\Sigma V^*$ *tal que* $V_h^* X \in \mathcal{G}\ell(h)$*. Sea* $q \ge 2$ *tal que*

$$4\left(\frac{\sigma_h}{\sigma_j}\right)^{2(q-1)} \|X_{2,j}X_{1,j}^\dagger\| + \left(\frac{\sigma_{k+1}}{\sigma_h}\right)^{2q-1} \|X_{2,k}X_{1,k}^\dagger\| \le \frac{1}{\sqrt{2}}$$

$$\implies \|A - QQ^*A\| \le \sigma_{h+1} + \sqrt{2}\,\sigma_{h+1}\,\delta_F \quad \textit{donde} \quad \delta_F :=$$

$$4\left(\frac{\sigma_h}{\sigma_j}\right)^{2q-1} \frac{\|\Sigma_{j,\perp}\|_F}{\sigma_j} \left\|X_{2,j}X_{1,j}^\dagger\right\| + \left(\frac{\sigma_{k+1}}{\sigma_h}\right)^{2q+1} \frac{\|\Sigma_{k,\perp}\|_F}{\sigma_k} \left\|X_{2,k}X_{1,k}^\dagger\right\|$$

donde $\|\Sigma_{\ell,\perp}\|_F^2 = \|diag(\sigma_{\ell+1},\ldots,\sigma_p)\|_F^2 = \sum_{i=\ell+1}^p \sigma_i^2$, $\ell = j, k$.

Observaciones finales

Las técnicas desarrolladas permiten:

- Considerar $X \in \mathbb{C}^{n\times r}$, para $h \leq r \leq p$ ($r - h$ es la redundancia).
- Considerar el caso $1 \leq j < h < k + 1$, de forma que

$$\sigma_j > \sigma_{j+1} \geq \ldots \geq \sigma_h \geq \ldots \geq \sigma_k > \sigma_{k+1} \, , \; \frac{\sigma_{j+1}}{\sigma_k} \approx 1 \, .$$

  Estos casos corresponden a valores singulares *agrupados* (clusters).
- Obtener estimaciones de los primeros h valores singulares de A en términos de los valores singulares de QQ^*A.
- Deducir estimaciones en el caso en que $X \in \mathbb{C}^{n\times r}$ está generada aleatoriamente (Gaussiana estándar).

Algunas referencias relacionadas

- P. Drineas, I.C.F. Ipsen, E.M. Kontopoulou, M. Magdon-Ismail, Structural convergence results for approximation of dominant subspaces from block Krylov spaces. SIAM J. Matrix Anal. Appl. 39 (2018), no. 2, 567-586.
- M. Gu, Subspace iteration randomization and singular value problems. SIAM J. Sci. Comput. 37 (2015), no. 3, A1139-A1173.
- P. Massey, Dominant subspace and low-rank approximations from block Krylov subspaces without a gap, 2021, enviado.
- P. Massey, Admissible subspaces and low rank approximations from the Subspace Iteration method, 2022, enviado.
- A.K. Saibaba, Randomized subspace iteration: analysis of canonical angles and unitarily invariant norms. SIAM J. Matrix Anal. Appl. 40 (2019), no. 1, 23-48.

Muchas gracias!!