

Cap. 5 Descompunerea valorilor singulare

5.1 Formularea problemei

$$A \in \mathcal{R}^{m \times n} \quad p = \min\{m, n\}$$

$$r - \text{rangul matricii } A \Rightarrow r \leq p$$

}

$r < p$ - matrice deficientă de rang

$r = p$ - matrice de rang complet

Teoremă:

Oricare ar fi matricea $A \in \mathcal{R}^{m \times n}$, există două matrici ortogonale $\tilde{U} \in \mathcal{R}^{m \times m}$ și $\tilde{V} \in \mathcal{R}^{n \times n}$, astfel încât:

$$\tilde{U}^T \cdot A \cdot \tilde{V} = \Sigma \quad A = \tilde{U} \cdot \Sigma \cdot \tilde{V}^T \quad (1)$$

unde este o matrice pseudo-diagonală, elementele nenule ale acesteia satisfacând relația:

$$\Sigma = \text{diag}\{\sigma_1, \dots, \sigma_p\}, \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$$

$\sigma_1, \dots, \sigma_p \in \mathcal{R}^+$ - valori singulare ale matricii A

$\Sigma \in \mathcal{R}^{m \times n}$ - formă canonică diagonală a matricii A

👉 relația (1) – **descompunerea valorilor singulare**

Demonstrația teoremei → algoritmul de descompunere a valorilor singulare

Observație:

- orice matrice este ortogonal echivalentă bilateral cu o matrice (pseudo)diagonală

👉 detaliind structura matricei $\Sigma \in \mathbb{R}^{m \times n}$, pot apare următoarele situații:

① $m \geq n$ $\text{rang}(A) = r = p = n$ \longrightarrow $\Sigma = \begin{bmatrix} \Sigma_1 \\ \boxtimes \\ \mathbf{0}_{(m-n) \times n} \end{bmatrix}, \quad \Sigma_1 = \text{diag}\{\sigma_1, \boxtimes, \sigma_n\}$

$\text{rang}(A) = r < p$ \longrightarrow $\Sigma = \begin{bmatrix} \Sigma_1 & \boxtimes & \mathbf{0}_{r \times (n-r)} \\ \boxtimes & \boxtimes & \boxtimes \\ \mathbf{0}_{(m-r) \times r} & \boxtimes & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix}, \quad \Sigma_1 \in \mathbb{R}^{r \times r}, \quad \Sigma_1 = \text{diag}\{\sigma_1, \boxtimes, \sigma_r\}$

② $m < n$ $\text{rang}(A) = r = p = m$ \longrightarrow $\Sigma = [\Sigma_1 \quad \boxtimes \quad \mathbf{0}_{m \times (n-m)}], \quad \Sigma_1 = \text{diag}\{\sigma_1, \boxtimes, \sigma_m\}$

$\text{rang}(A) = r < p$ \longrightarrow $\Sigma = \begin{bmatrix} \Sigma_1 & \boxtimes & \mathbf{0}_{r \times (n-r)} \\ \boxtimes & \boxtimes & \boxtimes \\ \mathbf{0}_{(m-r) \times r} & \boxtimes & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix}, \quad \Sigma_1 \in \mathbb{R}^{r \times r}, \quad \Sigma_1 = \text{diag}\{\sigma_1, \boxtimes, \sigma_r\}$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0, \quad \sigma_{r+1} = \dots = \sigma_p = 0$$

☞ rangul matricei A este egal cu rangul matricei Σ , care este egal cu r

⇓

rangul acestor matrici este egal cu numărul valorilor singulare nenule

☞ relația (1) poate fi scrisă sub forma: $A \cdot \tilde{V} = \tilde{U} \cdot \Sigma$

$\underline{\tilde{v}}_i$ - vectorii coloană ai \tilde{V}

$\underline{\tilde{u}}_i$ - vectorii coloană ai \tilde{U}



$$A \cdot \underline{\tilde{v}}_i = \sigma_i \cdot \underline{\tilde{u}}_i \Leftrightarrow \underline{\tilde{u}}_i^T \cdot A = \sigma_i \cdot \underline{\tilde{v}}_i^T$$

$i = 1, \dots, p$

Definiție:

Coloanele matricei \tilde{V} se numesc vectori singulari la dreapta ai matricei A, iar coloanele

matricei \tilde{U} se numesc vectori singulari la stânga ai matricei A.

👉 **proprietăți ale decomunerii valorilor**

singulare:

P1. valorile singulare ale matricei A sunt egale cu rădăcina pătratică pozitivă a valorilor

proprii ale matricei $A^T \cdot A$:

$$\sigma_i(A) = +\sqrt{\lambda_i(A^T \cdot A)}, \quad i = 1, \dots, n$$

P2. dacă $A \in \mathbb{R}^{n \times n}$ este o matrice simetrică și pozitiv semidefinită, atunci valorile proprii ale matricei A sunt reale, pozitive, iar valorile singulare ale matricei A sunt egale cu valorile ei

proprii: $\lambda_i(A) = \sigma_i(A), \quad i = 1, \dots, n$

P3. scriind matricile \tilde{U} și \tilde{V} sub forma:

$$\tilde{U} = [\underline{\tilde{u}}_1 \quad \underline{\tilde{u}}_r \quad \underline{\tilde{u}}_{r+1} \quad \underline{\tilde{u}}_m] = [\tilde{U}_1 \quad \tilde{U}_2]$$

$$\tilde{V} = [\underline{\tilde{v}}_1 \quad \underline{\tilde{v}}_r \quad \underline{\tilde{v}}_{r+1} \quad \underline{\tilde{v}}_n] = [\tilde{V}_1 \quad \tilde{V}_2]$$



$$A = \tilde{U} \cdot \Sigma \cdot \tilde{V}^T = [\tilde{U}_1 \quad \tilde{U}_2] \cdot \begin{bmatrix} \Sigma_1 & \mathbb{0}_{r \times (n-r)} \\ \mathbb{0}_{(m-r) \times r} & \mathbb{0}_{(m-r) \times (n-r)} \end{bmatrix} \cdot \begin{bmatrix} \tilde{V}_1^T \\ \tilde{V}_2^T \end{bmatrix}$$



$$A = \tilde{U}_1 \cdot \Sigma_1 \cdot \tilde{V}_1^T = \sum_{i=1}^r \sigma_i \cdot \underline{\tilde{u}}_i \cdot \underline{\tilde{v}}_i^T$$

P4. primele r coloane ale matricei \tilde{U} , unde $r = \text{rang}(A)$, formează o *bază ortogonală* pentru *subspațiul imagine al matricei A*:

$$\text{Im}(A) = \{ \underline{y} / \underline{y} = A \cdot \underline{x}, \forall \underline{x} \in \mathbb{R}^{n \times 1} \},$$

$$\underline{y} = A \cdot \underline{x} = \sum_{i=1}^r \tilde{u}_i \cdot (\sigma_i \cdot \tilde{v}_i^T \cdot \underline{x}) = \sum_{i=1}^r \alpha_i \cdot \tilde{u}_i$$

P5. ultimele $n - r$ coloane ale matricei \tilde{V} , formează o *bază ortogonală* pentru *subspațiul nul al matricei A*:

$$N(A) = \{ \underline{x} / A \cdot \underline{x} = \underline{0}_m, \underline{x} \in \mathbb{R}^{n \times 1} \}$$

$$A \cdot \tilde{V} = \tilde{U} \cdot \Sigma \Rightarrow A \cdot [\tilde{V}_1 \quad \tilde{V}_2] = [\tilde{U}_1 \quad \tilde{U}_2] \cdot \Sigma$$



$$A \cdot \tilde{V}_1 = \tilde{U}_1 \cdot \Sigma_1, A \cdot \tilde{V}_2 = \mathbf{0}_{m \times (n-r)}$$

5.2 Algoritmul SVD

👉 Algoritmul SVD (SVD – abr. eng. “Singular Value Decomposition”) → *construirea unui șir de matrici ortogonal echivalente bilateral, convergent către forma canonică pseudo-diagonală*



se bazează pe faptul
că:

$$\sigma_i(A) = +\sqrt{\lambda_i(A^T \cdot A)}$$

👉 Principiul de calcul → a aplica “mască” (direct matricii A) algoritmul QR



se aplică matricii A, transformările corespunzătoare celor care s-ar aplica matricii B în cadrul algoritmului QR de aducere la forma canonică Schur

$$A \leftrightarrow B = A^T \cdot A, \quad B = B^T, \quad B \in \mathcal{R}^{n \times n}$$

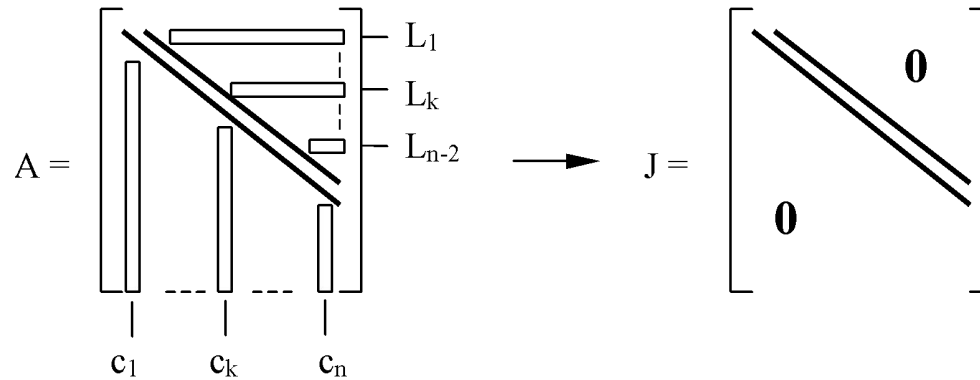
👉 se consideră $m > n$ (fără a restrânge generalitatea)

👉 Algoritmul parcurge două faze de lucru:

Faza 1: pregătitoare, de *aducere a matricii A la forma superior bidiagonală*

Faza 2: procedură iterativă, de *construcție a unui șir de matrici ortogonal echivalente bilateral, convergent către forma canonică pseudo-diagonală a matricii A*

□ Faza 1



$U_k \in \mathbb{R}^{m \times m}, k = 1, \dots, n$ - matrici Householder de ordin m care acțiunează pe coloane

$V_{k+1} \in \mathbb{R}^{n \times n}, k = 1, \dots, n - 2$ - matrici Householder de ordin n care acțiunează pe linii

👉 Tabloul general al transformărilor:

$$U_n \cdot \square \cdot U_k \cdot \square \cdot U_1 \cdot A \cdot V_2 \cdot \square \cdot V_{k+1} \cdot \square \cdot V_{n-1} = J$$

$$U = U_n \cdot \square \cdot U_k \cdot \square \cdot U_1, \quad V = V_2 \cdot \square \cdot V_{k+1} \cdot \square \cdot V_{n-1}$$

$$U \cdot A \cdot V = J$$



Observație:

$$B = A^T \cdot A \xrightarrow{\text{algoritm QR}} H - \text{matrice tridiagonală}$$

$$H = J^T \cdot J$$

5.3 Aplicații ale descompunerii valorilor singulare

- în urma aplicării SVD: $\hat{\Sigma} \cong \Sigma$

- algoritmul SVD – stabil numeric: $\hat{A} = A + E \Rightarrow \hat{\Sigma}_A \equiv \Sigma_{A+E}, \|E\|_2 \ll \|A\|_2$

$\hat{\sigma}_i, i = 1, \dots, p$ - valori singulare calculate

$\sigma_i, i = 1, \dots, p$ - valori singulare exacte

$$|\hat{\sigma}_i - \sigma_i| < \tau'$$

5.3.1 Calculul rangului unei matrici

$$\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \dots \geq \hat{\sigma}_p \geq 0$$

Definiție:

Se numește rang efectiv sau numeric, notat cu \hat{r} , al matricii A , acel index pentru care are loc relația de ordine:

$$\frac{\hat{\sigma}_{\hat{r}}}{\hat{\sigma}_1} \geq \tau' > \frac{\hat{\sigma}_{\hat{r}+1}}{\hat{\sigma}_1} \Rightarrow \hat{\sigma}_{\hat{r}} \geq \hat{\sigma}_1 \cdot \tau' > \hat{\sigma}_{\hat{r}+1}$$

$$\tau' = \max\{m, n\} \cdot \varepsilon_m$$

$$\tau = \hat{\sigma}_1 \cdot \tau' = \max\{m, n\} \cdot \varepsilon_m \cdot \hat{\sigma}_1$$

$$\hat{\sigma}_1 \geq \dots \geq \hat{\sigma}_{\hat{r}} \geq \tau > \hat{\sigma}_{\hat{r}+1} \geq \dots \geq \hat{\sigma}_q \geq 0$$

5.3.2 Rezolvarea sistemelor de ecuații algebrice liniare în sensul celor mai mici pătrate generalizate

$$A \cdot \underline{x} = \underline{b}, \quad A \in \mathcal{R}^{m \times n}, \quad \underline{b} \in \mathcal{R}^{m \times 1}, \quad \underline{x} \in \mathcal{R}^{n \times 1} \quad (3)$$

- dacă A – de rang complet

→ există o pseudosoluție unică în sensul celor mai mici pătrate ($m \geq n$)

↘ $m < n$

→ există o infinitate de pseudosoluții în sensul celor mai mici pătrate

- dacă A - deficientă de rang

👉 descompunerea valorilor singulare oferă posibilitatea determinării unei pseudosoluții unice, în sensul celor mai mici pătrate (generalizate), \underline{x}_{SVD}^* , indiferent de rangul matricei A.

👉 proprietăți:

- minimizează norma euclidiană a rezidului:

$$\underline{r} = \underline{b} - A \cdot \underline{x} \Rightarrow \underline{x}^* = \arg \min_{\underline{x} \in \mathcal{R}^{n \times 1}} \{ \|\underline{r}\|_2^2 \}$$

- este de normă euclidiană minimă:

$$\|\underline{x}_{SVD}^*\|_2 = \min_{\underline{x}^* \in \mathcal{R}^{n \times 1}} \{ \|\underline{x}^*\|_2 \}$$

Teoremă:

Oricare ar fi matricea $A \in \mathcal{R}^{m \times n}$ și vectorul $\underline{b} \in \mathcal{R}^{m \times 1}$, problema (3) are o soluție unică, aceasta fiind pseudosoluția normală obținută pe baza descompunerii valorilor singulare:

$$\underline{x}_{\text{SVD}}^* = A^+ \cdot \underline{b}, \quad A^+ \in \mathcal{R}^{n \times m}$$

A^+ - pseudosoluție Moore-Penrose $A^+ = \tilde{V}_1 \cdot \Sigma_1^{-1} \cdot \tilde{U}_1^T = \sum_{i=1}^r \frac{\tilde{v}_i \cdot \tilde{u}_i^T}{\sigma_i}$

$$\underline{x}_{\text{SVD}}^* = \left(\sum_{i=1}^r \frac{\tilde{v}_i \cdot \tilde{u}_i^T}{\sigma_i} \right) \cdot \underline{b}$$

Demonstrație:

$$\left. \begin{array}{l} \tilde{U}^T \cdot A \cdot \tilde{V} \cdot \tilde{V}^T \cdot \underline{x} = \tilde{U}^T \cdot \underline{b} \\ \tilde{U}^T \cdot A \cdot \tilde{V} = \Sigma \end{array} \right\} \Rightarrow \Sigma \cdot \underline{y} = \underline{d} \quad \text{unde} \quad \begin{cases} \underline{y} = \tilde{V}^T \cdot \underline{x} \\ \underline{d} = \tilde{U}^T \cdot \underline{b} \end{cases}$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & \boxtimes & \mathbf{0}_{r \times (n-r)} \\ \boxtimes & \boxtimes & \boxtimes \\ \mathbf{0}_{(m-r) \times r} & \boxtimes & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix}, \quad \Sigma_1 \in \mathcal{R}^{r \times r}, \quad \Sigma_1 = \text{diag}\{\sigma_1, \boxtimes, \sigma_r\}$$

$$\underline{y} = \begin{bmatrix} \underline{y}_1 \\ \boxtimes \\ \underline{y}_2 \end{bmatrix},$$

$$\underline{y}_1 = \tilde{V}_1^T \cdot \underline{x} \in \mathfrak{R}^{r \times 1},$$

$$\underline{y}_2 = \tilde{V}_2^T \cdot \underline{x} \in \mathfrak{R}^{(n-r) \times 1}$$

$$\underline{d} = \begin{bmatrix} \underline{d}_1 \\ \boxtimes \\ \underline{d}_2 \end{bmatrix},$$

$$\underline{d}_1 = \tilde{U}_1^T \cdot \underline{b} \in \mathfrak{R}^{r \times 1},$$

$$\underline{d}_2 = \tilde{U}_2^T \cdot \underline{b} \in \mathfrak{R}^{(m-r) \times 1}$$



$$\begin{bmatrix} \Sigma_1 & \boxtimes & \mathbf{0}_{r \times (n-r)} \\ \boxtimes & \boxtimes & \boxtimes \\ \mathbf{0}_{(m-r) \times r} & \boxtimes & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix} \cdot \begin{bmatrix} \underline{y}_1 \\ \boxtimes \\ \underline{y}_2 \end{bmatrix} = \begin{bmatrix} \underline{d}_1 \\ \boxtimes \\ \underline{d}_2 \end{bmatrix} \Rightarrow \begin{bmatrix} \Sigma_1 \cdot \underline{y}_1 \\ \boxtimes \\ \mathbf{0}_{(m-r) \times 1} \end{bmatrix} = \begin{bmatrix} \underline{d}_1 \\ \boxtimes \\ \underline{d}_2 \end{bmatrix}, \quad \forall \underline{y}_2 \in \mathfrak{R}^{(n-r) \times 1}$$

👉 minimizare normă euclidiană a reziduului:

$$\|\underline{r}\|_2^2 = \|\tilde{U}^T \cdot \underline{r}\|_2^2 = \|\tilde{U}^T \cdot (\underline{b} - \mathbf{A} \cdot \underline{x})\|_2^2 = \|\tilde{U}^T \cdot \underline{b} - \tilde{U}^T \cdot \mathbf{A} \cdot \tilde{V} \cdot \tilde{V}^T \cdot \underline{x}\|_2^2 =$$

$$= \|\underline{d} - \Sigma \cdot \underline{y}\|_2^2 = \|\underline{d}_1 - \Sigma_1 \cdot \underline{y}_1\|_2^2 + \|\underline{d}_2\|_2^2 = \min$$

- în general $\underline{d}_2 \neq \underline{0}_{m-r} \Leftrightarrow \|\underline{d}_2\|_2^2 \neq 0 \rightarrow$ singura minimizare posibilă:

$$\|\underline{d}_1 - \Sigma_1 \cdot \underline{y}_1^*\|_2^2 = 0 \Rightarrow \underline{d}_1 - \Sigma_1 \cdot \underline{y}_1^* = \underline{0}_p \Rightarrow \Sigma_1 \cdot \underline{y}_1^* = \underline{d}_1$$



$$\tilde{V}^T \cdot \underline{x} = \underline{y} \Rightarrow \underline{x}^* = \tilde{V} \cdot \underline{y}^* = \tilde{V} \cdot \begin{bmatrix} \underline{y}_1^* \\ \emptyset \\ \underline{y}_2 \end{bmatrix}, \quad \forall \underline{y}_2 \in \mathcal{R}^{(n-r) \times 1}$$

 **pseudosoluție de normă minimă:**

$$\|\underline{x}^*\|_2^2 = \min$$

$$\|\underline{x}^*\|_2^2 = \|\tilde{V}^T \cdot \underline{x}^*\|_2^2 = \|\underline{y}^*\|_2^2 = \|\underline{y}_1^*\|_2^2 + \|\underline{y}_2\|_2^2 = \min$$

- în general $\underline{y}_1^* = \Sigma_1^{-1} \cdot \tilde{U}^T \cdot \underline{b} \neq \underline{0}_p$



$$\|\underline{y}_2^*\|_2^2 = 0 \Leftrightarrow \underline{y}_2^* = \underline{0}_{n-p}$$



$$\underline{x}_{SVD}^* = \tilde{V} \cdot \underline{y}_{SVD}^* = [\tilde{V}_1 \quad \emptyset \quad \tilde{V}_2] \cdot \begin{bmatrix} \Sigma_1^{-1} \cdot \tilde{U}_1^T \cdot \underline{b} \\ \emptyset \\ \underline{0}_{n-r} \end{bmatrix} = \tilde{V}_1 \cdot \Sigma_1^{-1} \cdot \tilde{U}_1^T \cdot \underline{b} = \sum_{i=1}^r \left(\frac{\tilde{v}_i \cdot \tilde{u}_i^T}{\sigma_i} \right) \cdot \underline{b}$$

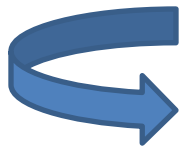
Cap. 6 Ecuații și sisteme de ecuații neliniare

6.1 Formularea problemei

□ Fie $\underline{f}(\underline{x}) = \underline{0}_n$, $\underline{f}: D_1 \rightarrow D_2$, $D_1, D_2 \subseteq \mathfrak{R}^{n \times 1}$ ($\mathbf{C}^{n \times 1}$) (1)

$$\underline{x} = [x_1 \quad \& \quad x_n]^T \quad \underline{f}(\underline{x}) = [f_1(x_1, \dots, x_n) \quad \& \quad f_n(x_1, \dots, x_n)]^T$$

$f_i(\underline{x})$, $i = 1, \dots, n$ - compunere de funcții elementare de tip polinomial, trigonometric sau transcendent (exponențiale, logaritmi) în argumentele x_i , $i = 1, \dots, n$



(1) - **ecuație** ($n = 1$) / **sistem de ecuații** ($n \geq 2$) **neliniare**

□ Fie $\underline{\alpha} = [\alpha_1 \quad \& \quad \alpha_n]^T$, $\underline{f}(\underline{\alpha}) = \underline{0}_n$

□ Metodele numerice pentru rezolvarea problemei (1) □ metode iterative

construiesc un șir de vectori convergent la soluția $\underline{\alpha}$:

$$\{\underline{x}^{[k]}\}_{k \geq 0}, \lim_{k \rightarrow \infty} \underline{x}^{[k]} = \underline{\alpha}$$

- punct de start: $\underline{x}^{[0]} = [x_1^{[0]} \quad \& \quad x_n^{[0]}]^T \in V_{\underline{\alpha}} = \{\underline{x} \in D_1 / \|\underline{x} - \underline{\alpha}\|_p \leq h, h > 0\} \subset D_1$

Definiție:

Se numește formulă de iterare sau șir de iterare (de rang k și ordin m) o relație de recurență de forma:

$$\underline{x}^{[k]} = \underline{g}(k, \underline{x}^{[k-1]}, \underline{x}^{[k-2]}, \dots, \underline{x}^{[k-m]}), \quad \underline{x}^{[0]}, \dots, \underline{x}^{[m-1]} \in V_\alpha, \quad 1 \leq m \leq k$$

👉 formulă de iterare staționară de ordinul m → \underline{g} nu depinde de

k

👉 uzuale sunt metodele iterative care folosesc formule staționare de ordin $m = 1$:

$$\underline{x}^{[k]} = \underline{g}(\underline{x}^{[k-1]}) \tag{2}$$

Propoziție:

O metodă iterativă de tipul (2) se spune că este convergentă pe domeniul V_α , dacă oricare ar fi

$\underline{x}, \underline{y} \in V_\alpha$, există $\lambda \in (0,1)$ astfel încât, într-o normă vectorială oarecare p , aplicația \underline{g} îndeplinește relația:

$$\|\underline{g}(\underline{x}) - \underline{g}(\underline{y})\|_p \leq \lambda \cdot \|\underline{x} - \underline{y}\|_p$$

Altfel spus, aplicația este o contracție pe domeniul V_α .

- soluție generală de alegere a funcției de iterare, g :

se rescrie ecuația (1) $\longrightarrow \underline{x} = \underline{g}(\underline{x})$ astfel încât $\underline{f}(\underline{\alpha}) = \underline{0} \Leftrightarrow \underline{\alpha} = \underline{g}(\underline{\alpha})$

- critériu general de stop al metodei iterative:

$$\|\underline{x}^{[s]} - \underline{x}^{[s-1]}\|_p \leq \varepsilon_{\underline{x}} \quad \text{sau} \quad \|\underline{f}(\underline{x}^{[s]})\|_p \leq \varepsilon_{\underline{f}}$$