

Aproximările unei norme

Cozma Laura-Elena

Manolache Andrei

May 2021

1 Problema aproximării normei

Problema aproximării normei este o problemă fără constrângeri de forma:

$$\text{minimize } \|Ax - b\|$$

unde $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $x \in \mathbb{R}^n$ și $\|\cdot\|$ este o normă pe \mathbb{R}^m . Aceasta este o problemă convexă, existând mereu cel puțin o soluție optimă.

1.1 Interpretarea aproximării

Relația poate fi scrisă ca:

$$Ax = x_1 a_1 + \dots + x_n a_n$$

unde a_i este coloana i a matricei A . Astfel, scopul problemei va fi de a aproxima cu o deviație măsurată de norma $\|\cdot\|$ aproximarea lui b cu o combinație liniară a coloanelor lui A . Se va nota cu $r = Ax - b$ reziduul problemei, fiind interpretat ca deviația rezultatului actual Ax de la rezultatul target b .

1.2 Interpretarea geometrică

Din punct de vedere geometric, se va considera un subspațiu $A \subseteq \mathbb{R}^m$ și punctul $b \in \mathbb{R}^m$, iar problema se va reduce la găsirea proiecției punctului b în subspațiul A în norma $\|\cdot\|$, adică soluția optimă a problemei

$$\begin{aligned} &\text{minimize } \|u - b\| \\ &\text{subject to } u \in A \end{aligned} \tag{1}$$

1.3 Tipuri de aproximări

- Aproximarea celor mai mici pătrate

În cazul în care se ridică la pătrat funcția obiectiv, iar norma aleasă este l_2 , se obține problema aproximării celor mai mici pătrate

$$\text{minimize } \|Ax - b\|_2^2 = r_1^2 + r_2^2 + \dots + r_m^2$$

- Aproximarea Chebyshev

Folosind norma l_∞ , problema aproximării Chebyshev se bazează pe minimizarea valorii absolute a reziduului maxim.

- Aproximarea sumei a valorilor absolute ale reziduului

Se folosește norma l_1 și se obține

$$\text{minimize } \|Ax - b\|_1 = |r_1| + \dots + |r_m|$$

2 Aproximarea folosind funcția de penalitate

Problema aproximării funcției de penalitate are forma:

$$\begin{aligned} &\text{minimize } \phi(r_1) + \dots + \phi(r_m) \\ &\text{subject to } r = Ax - b \end{aligned} \tag{2}$$

Expresia a fost construită pornind de la aproximarea lui Ax cu b , x fiind rezultatul ales, iar $r \in R^m$ reziduul obținut, adică $r = Ax - b$. Fiecărei componente $r_i, i \in \{1, \dots, m\}$ din vectorul r îi asociem o penalizare $\phi(r_i)$ dată de funcția de penalitate $\phi : R \rightarrow R$, penalizarea totală constând în suma penalizărilor fiecărei componente a reziduului. Astfel, problema de minimizare se reduce la a minimiza penalizarea totală.

Exemple de funcții de penalizare:

- $\phi(u) = |u|^p$, $p \geq 1$, problema aproximării cu această funcție de penalitate fiind echivalentă cu problema aproximării cu norma lp . În continuare se va pune accentul pe cazurile când $p = 1$, $\phi_1(u) = |u|$, respectiv $p = 2$, $\phi_2(u) = u^2$.
- Funcția de penalizare *deadzone-linear* cu o limită $a > 0$

$$\phi(u) = \begin{cases} 0 & |u| \leq a \\ |u| - a & |u| > a \end{cases}$$

- Funcția *log barrier* cu o limită $a > 0$

$$\phi(u) = \begin{cases} -a^2 \log(1 - (u/a)^2) & |u| < a \\ \infty & |u| \geq a \end{cases}$$

În continuare se vor compara aproximările folosind norma l_1 , norma l_2 , penalizarea deadzone-linear și penalizarea log barrier. Se va alege o matrice $A \in R^{m \times n}$ și un vector b^m și se vor calcula soluțiile $Ax \approx b$, asupra cărora se aplică funcțiile de penalizare enumerate mai sus. Analizând distribuția amplitudinii a reziduului optim asociată celor patru funcții, se observă că:

1. Penalizarea cu norma l_1

S-a observat că impactul funcției de penalizare este mare pe valorile mici ale reziduului și mic pentru valorile mari. Cum $\phi(u) = u$ iar accentul este pus pe valorile mici, atunci norma l_1 va produce valori optime ale reziduului foarte apropiate de 0 sau chiar 0. Astfel, utilizarea aproximării folosind norma l_1 va duce la găsirea unui număr mare de ecuații $a_i^T x = b_i$ care sunt satisfăcute exact.

2. Penalizarea cu norma l_2

Aceasta stabilește o diferențiere mai mică pentru reziduuri mici, întrucât $\phi(u) = u^2$ este foarte mic pentru valori u mici, și o diferențiere mare pentru valorile mari. Se observă o frecvență mare a valorilor subunitare și foarte puține valori mari. Spre deosebire de norma l_1 , unde pentru valori subunitare avem $\phi_1(u) \gg \phi_2(u)$, deci accentul va fi pus pe valorile mici, aproximarea cu norma l_2 va face o diferențiere mai puțin semnificativă pentru valorile mici. Pentru valori mai mari ca 1, $\phi_2(u) \gg \phi_1(u)$, astfel că se va face o diferențiere mai accentuată a funcției de penalizare, iar valorile acesteia fiind mari, în distribuția amplitudinii reziduului optim vor apărea puține reziduuri cu valori mari.

3. Penalizarea cu deadzone-linear

Aceasta asociază un cost 0 valorilor mai mici decât a , iar accentul asupra restului valorilor este unul scăzut.

4. Penalizarea cu log barrier

Accentul între valorile mici este scăzut, fiind similară cu distribuția normei l_2 , iar pentru valorile mai mari ca 1, costul este infinit.

2.1 Senzitivitatea la erori mari

Un outlier este o valoare $y_i = a_i^T x + v_i$, unde zgomotul v_i este mare, care poate apărea în urma unor date incorecte sau al calculelor defectuoase. Această valoare va avea asociat un reziduu care va avea valori mari pe unele componente. Pentru a nu impacta rezultatele, scopul este eliminarea acestora din estimare sau asocierea unor greutăți mici.

Prima modalitate care ar putea fi abordată este asocierea aproximării cu norma l_2 pentru valori mai mici decât un prag M , iar pentru valorile de peste pragul M se asociază o valoare M^2 constantă. Astfel, valorile mai mari ca M vor fi ignorate. Însă această abordare duce către o funcție de penalizare care nu este convexă, deci problema de optimizare nu va fi convexă. Raportându-ne la funcții de penalitate convexe, cele care au o sensibilitate mai mică la valorile mici, precum aproximarea folosind norma l_1 , se numesc funcții robuste.

3 Aproximări cu constrângeri

În problema minimizării funcției obiectiv, se pot introduce anumite restricții privind domeniul soluției.

3.1 Constrângerea cu valori pozitive

$$\begin{aligned} & \text{minimize } \|Ax - b\| \\ & \text{subject to } x \succeq 0 \end{aligned} \tag{3}$$

O astfel de estimare apare atunci când vectorul conține doar valori pozitive, de exemplu când componentele sale $x_1, x_2 \dots x_n$ reprezintă intensități, semnale etc. Putem interpreta problema ca pe o aproximare a vectorului b printr-o combinație liniară a coloanelor din matricea A .

3.2 Constrângerea cu mărginirea domeniului

$$\begin{aligned} & \text{minimize } \|Ax - b\| \\ & \text{subject to } l \preceq x \preceq u \end{aligned} \tag{4}$$

O astfel de estimare apare atunci când cunoaștem domeniul de apartenență ale componentelor vectorului x .

3.3 Distribuția de probabilitate

$$\begin{aligned} & \text{minimize } \|Ax - b\| \\ & \text{subject to } x \succeq 0, 1^T x = 1 \end{aligned} \tag{5}$$

Apare atunci când componentele vectorului reprezintă densități de probabilitate într-un punct (i.e. suma lor este 1 și sunt pozitive). Poate fi interpretată și ca aproximarea lui b cu combinație convexă a coloanelor matricei A .

3.4 Constrângerea de tip bilă

$$\begin{aligned} & \text{minimize } \|Ax - b\| \\ & \text{subject to } \|x - x_0\| \leq d \end{aligned} \tag{6}$$

O astfel de constrângere poate apărea atunci când se alege x_0 o aproximare a soluției, iar d reprezintă deviația maximă de la soluția aleasă și care minimizează funcția obiectiv.