Поглавље 7

Правоугаона матрична једначина

Правоугаона матрична једначина гласи

$$Ax = b,$$

где је A правоугаона матрица, а x и b одговарајући вектори, која се назива правоугаона матрична једначина. У овом поглављу је делимично коришћена следећа књига [1].

7.1 Представљање правоугаоног система линеарних једначина у матричном облику

Систем од m линеарних једначина са n непознатих може се представити у матричном облику на следећи начин

$$Ax = b,$$

где је

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{и} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Помоћу матрица A и b формирамо нову матрицу B, на следећи начин

$$B = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix}.$$

Матрица B је формата $m \times (n+1)$ и назива се проширена матрица датог система линеарних једначина. При овако уведеним ознакама формулисаћемо критеријум решивости датог система линеарних једначина:

"систем $Ax = b$ је решив ако и само ако је $\mathrm{r}(A) = \mathrm{r}(B)$".

Недовољно одређени систем линеарних једначина је систем

$$Ax = b,$$

где је $A \in \mathbb{R}^{m\times n}$, при чему је број једначина мањи од броја непознатих, т.ј. $m < n$.

Инверз матрице A је дефинисан само када је матрица A квадратна и њена детерминанта различита од нуле. Уколико је матрица A правоугаона, рецимо, $A \in \mathbb{R}^{m\times n}$, $m \neq n$, онда, увек постоји и једниствен је Мур-Пенроузов инверз матрице A у ознаци $A^\dagger \in \mathbb{R}^{n\times m}$, који се одређује као јединствено решење система матричних једначина

$$\begin{aligned} &(1)\ AXA = A; \qquad && (3)\ (AX)^{\mathrm{T}} = AX;\\ &(2)\ XAX = X; \qquad && (4)\ (XA)^{\mathrm{T}} = XA. \end{aligned}$$

Мур-Пенроузов инверз $A^\dagger$ можемо искористити да одредимо решење система линеарних једначина $Ax = b$, где је матрица A правоугаона, односно $A \in \mathbb{R}^{m\times n}$, $m \neq n$.

Када је матрица A потпуног ранга врста, т.ј. када је $A \in \mathbb{R}^{m\times n}$, при чему је $r(A) = m$, тада за сваки вектор $b \in \mathbb{R}^m$ постоји бар једно решење датог система. У многим применама довољно је пронаћи само једно решење.

7.2 Проблем апроксимације помоћу норме

Проблем апроксимације помоћу норме у свом најједноставнијем облику гласи

$$minimize\|Ax - b\|,$$

где је $A \in \mathbb{R}^{m\times n}$ дата матрица, $b \in \mathbb{R}$ је дати вектор, а $x \in \mathbb{R}^n$ је непознати вектор и $\|\cdot\|$ је норма на $\mathbb{R}^m$, т.ј. функција $\|\cdot\| : \mathbb{R}^m \to \mathbb{R}^+$, која задовољава следеће услове:

(i) $\|x\| \geqslant 0$;

(ii) $\|x\| = 0 \Leftrightarrow x = 0$;

(iii) $\|x + y\| \leqslant \|x\| + \|y\|$.

Решење проблема апроксимације помоћу норме се понекад назива апроксимативно решење проблема

$$Ax \approx b$$

у односу на норму $\|\cdot\|$. Вектор $r = Ax - b$ се назива резидуал проблема, а његове компоненте се називају резидуали придружени x.

Проблем апроксимације помоћу норме $minimize\|Ax-b\|$ је конвексан проблем, који је увек решив, т.ј. увек постоји најмање једно оптимално решење таквог проблема. Његова оптимална вредност је нула ако и само ако $b \in \mathcal{R}(A)$, где је $\mathcal{R}(A)$ скуп слика матрице A. Међутим, проблем је интересантнији и користан када $b \notin \mathcal{R}(A)$. Можемо претпоставити, без губитка општости да су колоне матрице A линеарно независне, а посебно да је $m \geqslant n$. У случају када је $m = n$, оптимална тачка је $A^{-1}b$, тако да можемо претпоставити да је $m > n$.

Интерпретација апроксимације Изражавајући Ax у облику

$$Ax = x_1a_1 + \ldots + x_na_n,$$

где су $a_1, \ldots, a_n \in \mathbb{R}^m$ колоне матрице A, можемо видети да је циљ апроксимационог проблема помоћу норме да се "фитује" или апроксимира вектор b линеарном комбинацијом колона матрице A, колико год је могуће близу, при чему се девијација мери помоћу норме $\|\cdot\|$.

Апроксимациони проблем се такође назива и регресиони проблем. Због тога се вектори $a_1, \ldots, a_n$ називају регресори, а вектор $x_1a_1 + \ldots + x_na_n$, где је $x = (x_1, \ldots, x_n)$ оптимално решење проблема, се назива регресија од b (на регресоре).

Интерпретација процене Блиско повезана интерпретација проблема апроксимације помоћу норме се јавља у проблему процене параметарског вектора на базу несавршених линеарних векторских мерења. Посматрајмо модел линеарног мерења

$$y = Ax + v,$$

где је $y \in \mathbb{R}^m$ вектор мерења, $x \in \mathbb{R}^n$ је вектор параметара који се процењују и $v \in \mathbb{R}^m$ је грешка мерења, која је непозната, али се претпоставља да је мала у односу на норму $\|\cdot\|$. Проблем процене је да се смислено предвиди којем x одговара дато y.

Ако претпоставимо да x има вредност $\hat{x}$, онда ми имплицитно предвиђамо да v има вредност $y - A\hat{x}$. Претпоставимо да су мање вредности за v (у односу на норму $\|\cdot\|$) вероватније него веће вредности, највероватнија претпоставка за x је

$$\hat{x} = argmin_z\|Az - y\|.$$

Проблем апроксимације методом најмањих квадрата Најчешћи проблем апроксимације помоћу норме укључује Еуклидову или l_2-норму. Квадрирањем функције циља добијамо еквивалентан проблем, који се назива проблем апроксимације методом најмањих квадрата:

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

где је функција циља сума квадрата резидуала $r_1, r_2, \ldots, r_m$.

Функција коју је неопходно минимизирати код проблема апроксимације методом најмањих квадрата је

$$f(x) = \|Ax - b\|_2^2,$$

где је $\|\cdot\|_2$ Еуклидова или l_2-норма, која је дефинисана на следећи начин

$$\begin{aligned}
f(x) =&((Ax)_1 - b_1)^2 + ((Ax)_2 - b_2)^2 + \ldots + ((Ax)_m - b_m)^2\\
=&(Ax)_1^2 - 2(Ax)_1 b_1 + b_1^2\\
&+ (Ax)_2^2 - 2(Ax)_2 b_2 + b_2^2\\
&+ \ldots\\
&+ (Ax)_m^2 - 2(Ax)_m b_m + b_m^2\\
=&(Ax)_1^2 + (Ax)_2^2 + \ldots + (Ax)_m^2 - 2(Ax)_1 b_1 - 2(Ax)_2 b_2 - \ldots - 2(Ax)_m b_m\\
&+ b_1^2 + b_2^2 + \ldots + b_m^2\\
=&(Ax)_1^2 + (Ax)_2^2 + \ldots + (Ax)_m^2 - 2b_1(Ax)_1 - 2b_2(Ax)_2 - \ldots - 2b_m(Ax)_m\\
&+ b_1^2 + b_2^2 + \ldots + b_m^2\\
=&(Ax)^{\mathrm{T}} Ax - 2b^{\mathrm{T}} Ax + b^{\mathrm{T}} b\\
=&x^{\mathrm{T}} A^{\mathrm{T}} Ax - 2b^{\mathrm{T}} Ax + b^{\mathrm{T}} b.
\end{aligned}$$

Дакле, функција коју је неопходно минимизирати је дата у облику

$$f(x) = x^{\mathrm{T}} A^{\mathrm{T}} Ax - 2b^{\mathrm{T}} Ax + b^{\mathrm{T}} b,$$

а градијент функције $f(x)$ је

$$(\nabla f)(x) = 2A^{\mathrm{T}} Ax - 2b^{\mathrm{T}} A.$$

Функција $f(x)$ достиже минимум у тачки x_0 онда и само онда, када је градијент функције $f(x)$ у тачки x_0 једнак нула вектору:

$$(\nabla f)(x_0) = 2A^{\mathrm{T}} Ax_0 - 2A^{\mathrm{T}} b = 0,$$

т.ј. ако и само ако x_0 задовољава нормалну једначину

$$A^{\mathrm{T}} Ax_0 = A^{\mathrm{T}} b,$$

која увек има решење. Под претпоставком да су колоне матрице A линеарно независне, односно да је за матрицу $A \in \mathbb{R}^{m\times n}$, $r(A) = n$, решење проблема апроксимације методом најмањих квадрата јединствено и гласи

$$x_0 = (A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}} b.$$

Образложење зашто матрица $A^{\mathrm{T}} A \in \mathbb{R}^{n\times n}$, при чему је $r(A) = n$ мора бити регуларна је приказано у наставку.

Регуларност матрице $A^{\mathrm{T}}A$ Ако је $A \in \mathbb{R}^{m\times n}$, онда матрице A и AA^{T} имају по m врста. Ранг било које матрице која има m врста једнак је броју m умањеном за број независних линеарних релација међу њеним редовима.

Да бисмо доказали да је $r(A) = r(AA^{\mathrm{T}})$ довољно је доказати да свака линеарна релација између врста матрице A важи за одговарајуће врсте матрице AA^{T} и обратно.

Било која нетривијална релација између врста матрице H је еквивалентна егзистенцији ненула вектор врсте x^{T} такве да је $x^{\mathrm{T}}H = 0$. Очигледно је

$$x^{\mathrm{T}}H = 0 \Longrightarrow x^{\mathrm{T}}AA^{\mathrm{T}} = 0$$

и обратно

$$x^{\mathrm{T}}AA^{\mathrm{T}} = 0 = 0 \Longrightarrow 0 = x^{\mathrm{T}}AA^{\mathrm{T}}x = (A^{\mathrm{T}}x)^{\mathrm{T}}A^{\mathrm{T}}x \Longrightarrow A^{\mathrm{T}}x = 0 \Longrightarrow x^{\mathrm{T}}A = 0.$$

Овде смо искористили чињеницу да за било коју вектор колону y реалних бројева израз $y^{\mathrm{T}}y$ је сума квадрата тих реалних бројева, а таква сума је једнака нули једино када је сваки њен сабирак једнак нули.

Примењујући наведени резултат на матрицу A^{T}, добијамо да је

$$r(A^{\mathrm{T}}A) = r(A^{\mathrm{T}}),$$

где можемо искористити да је $r(A^{\mathrm{T}}) = r(A)$, па добијамо да је $r(A^{\mathrm{T}}A) = r(A)$, односно да је матрица $A^{\mathrm{T}}A$ инвертибилна у случају када је ранг матрице $A \in \mathbb{R}^{m\times n}$ једнак n.

7.3 Питања

1. Како гласи правоугаона матрична једначина?
2. Формулисати критеријум решивости правоугаоног система линеарних једначина.
3. Представити систем од m линеарних једначина са n непознатих у матричном облику.
4. Како се дефинише Мур-Пенроузов инверз правоугаоне матрице.
5. Формулисати најједноставнији проблем апроксимације помоћу норме.
6. Како се дефинише норма на $\mathbb{R}^m$?
7. Како се дефише и шта представља резидуал проблема апроксимације помоћу норме?
8. Објаснити проблем апроксимације методом најмањих квадрата.

7.4 Додатак

7.4.1 Матричне факторизације правоугаоних матрица

Врло често нам је корисно да правоугаону матрицу A представимо у погоднијем облику, користећи матричне факторизације:

– LU факторизација правоугаоне матрице;

– QR факторизација правоугаоне матрице;

– SVD факторизација правоугаоне матрице.

У наставку ћемо детаљније описати неке од наведених факторизација.

Ненегативан реалан број σ је сингуларна вредност матрице $A \in \mathbb{R}^{m\times n}$ ако постоје јединични вектори $u \in \mathbb{R}^m$ и $v \in \mathbb{R}^n$, такви да је

$$\begin{aligned} Av =&\sigma u;\\ A^{\mathrm{T}}u =&\sigma v. \end{aligned}$$

Овде се вектори u и v називају лево сингуларни и десно сингуларни за сингуларну вредност σ, респективно.

LU факторизација правоугаоне матрице

LU факторизација матрице $A \in \mathbb{R}^{n\times r}$ такве да је $r(A) = r$, гласи

$$A = PLU, \quad \text{број неопходних операција}^1 \text{ је } \frac{2}{3}r^3 + r^2(n-r);$$

где је

$P \in \mathbb{R}^{n\times n}$ пермутациона матрица;
$L \in \mathbb{R}^{n\times r}$ је јединична доња троугаона матрица т.j. $l_{ij} = 0,$ за $i < j$ и $l_{ii} = 1$;
$U \in \mathbb{R}^{r\times r}$ је несингуларна и горња троугаона матрица.

LU факторизација ретке несингуларне матрице[2] $A \in \mathbb{R}^{n\times r}$, чији је ранг $r(A) = r$ укључује пермутацију врста и колона, односно гласи

$$A = P_1LUP_2,$$

где су $P_1, P_2 \in \mathbb{R}^{r\times r}$ пермутационе матрице. LU факторизација ретке правоугаоне матрице се знатно ефикасније одређује, у смислу да је потребан мањи број операција за њено одређивање него за густе матрице.

Наведене LU факторизације могу се искористити за решавање недовољно одређеног система линеарних једначина.

[2] Појам ретке матрице је на енглеском “sparse matrix”.

QR факторизација правоугаоне матрице

QR факторизација матрице $A \in \mathbb{R}^{n\times r}$ такве да је $r(A) = r \leqslant n$ гласи

$$A = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix} \quad \text{број неопходних операција је } 2r^2\left(n - \frac{r}{3}\right);$$

где матрице $Q_1 \in \mathbb{R}^{n\times r}$ и $Q_2 \in \mathbb{R}^{n\times(n-r)}$ задовољавају

$$Q_1{}^{\mathrm{T}}Q_1 = I, \quad Q_2{}^{\mathrm{T}}Q_2 = I, \quad Q_1{}^{\mathrm{T}}Q_2 = 0,$$

док је $R \in \mathbb{R}^{r\times r}$ горња троугаона матрица са ненула елементима на главној дијагонали. Оваква факторизација се назива QR факторизација матрице A.

Метода која користи QR факторизацију је најкоришћенија метода за решавање недовољно одређених система једначина. Недостатак ове методе је тешкоћа да се изрази реткост. Фактор Q је обично густа матрице, чак и када је A веома ретка матрица.

SVD факторизација правоугаоне матрице

SVD факторизација матрице $A \in \mathbb{R}^{m\times n}$ такве да је $r(A) = r$ гласи

$$A = U\Sigma V^{\mathrm{T}},$$

где матрица $U \in \mathbb{R}^{m\times r}$ задовољава $U^{\mathrm{T}}U = I$, а матрица $V \in \mathbb{R}^{n\times r}$ задовољава $V^{\mathrm{T}}V = I$, док је $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$, при чему важи

$$\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0.$$

Колоне матрице U се зову леви сингуларни вектори од A, док су колоне матрице V десни сингуларни вектори од A, а бројеви $\sigma_1, \sigma_2, \ldots, \sigma_r$ су сингуларне вредности матрице A.

Библиографија

[1] **Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, Cambridge, UK, 2009. Seventh Printing with Corrections.**

Индекс

Листа симбола

$\approx$ релација апроксимације