

Sistemi linearnih enačb, 2.del

Predoločeni sistemi

Spomnimo se, da je sistem linearnih enačb **predoločen**, če ima več enačb kot spremenljivk. Spomnimo se tudi, da je predoločen sistem običajno nerešljiv. Radi bi posplošili definicijo rešitve sistema tako, da bo vsak predoločen sistem posplošeno rešljiv.

Definicija posplošene rešitve sistema linearnih enačb

Posplošena rešitev sistema linearnih enačb

$$\begin{array}{rcl} a_{1,1}x_1 + \dots + a_{1,n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n & = & b_m \end{array} \quad (1)$$

je taka n -terica realnih števil $(x_1, \dots, x_n)$ pri kateri je vektor levih strani

$$(a_{1,1}x_1 + \dots + a_{1,n}x_n, \dots, a_{m,1}x_1 + \dots + a_{m,n}x_n)$$

najbližje vektorju desnih strani $(b_1, \dots, b_m)$.

Opomba: Če je sistem (1) rešljiv v običajnem smislu, potem se posplošena rešitev ujema z običajno rešitvijo.

Opomba: Posplošeni rešitvi pravimo tudi (posplošena) **rešitev po metodi najmanjših kvadratov**, ker minimizira izraz

$$(a_{1,1}x_1 + \dots + a_{1,n}x_n - b_1)^2 + \dots + (a_{m,1}x_1 + \dots + a_{m,n}x_n - b_m)^2 \quad (2)$$

Poglejmo si zdaj, kako poiščemo posplošeno rešitev. Izraz (2) najprej predelamo v vektorsko obliko

$$\|x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n - \mathbf{b}\|^2 \quad (3)$$

kjer je $\mathbf{a}_1 = (a_{1,1}, \dots, a_{m,1}), \dots, \mathbf{a}_n = (a_{1,n}, \dots, a_{m,n}), \mathbf{b} = (b_1, \dots, b_m)$.

Geometrijsko gledano je minimizacija izraza (3) ekvivalentna pravokotni projekciji vektorja $\mathbf{b}$ na množico vseh linearnih kombinacij vektorjev $\mathbf{a}_1, \dots, \mathbf{a}_n$. Poiskati moramo torej take skalarje $x_1, \dots, x_n$, da je vektor $x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n - \mathbf{b}$ pravokoten na vektorje $\mathbf{a}_1, \dots, \mathbf{a}_n$.

Če pravokotnostne pogoje

$$\langle x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n - \mathbf{b}, \mathbf{a}_1 \rangle = \dots = \langle x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n - \mathbf{b}, \mathbf{a}_n \rangle = 0$$

razpišemo, dobimo sistem linearnih enačb

$$\begin{array}{rcl} \langle \mathbf{a}_1, \mathbf{a}_1 \rangle x_1 + \dots + \langle \mathbf{a}_n, \mathbf{a}_1 \rangle x_n & = & \langle \mathbf{b}, \mathbf{a}_1 \rangle \\ \vdots & & \vdots \\ \langle \mathbf{a}_1, \mathbf{a}_n \rangle x_1 + \dots + \langle \mathbf{a}_n, \mathbf{a}_n \rangle x_n & = & \langle \mathbf{b}, \mathbf{a}_n \rangle \end{array} \quad (4)$$

Posplošene rešitve sistema (1) se torej ujemajo z običajnimi rešitvami sistema (4).

Opomba: Izkaže se, da je sistem (4) vedno rešljiv. Enolično je rešljiv natanko tedaj, ko so vektorji $\mathbf{a}_1, \dots, \mathbf{a}_n$ linearno neodvisni. Kadar ni enolično rešljiv, poiščemo njegovo najkrajšo rešitev. S temi vprašanji se bomo ukvarjali na koncu drugega semestra v razdelku o psevdoinverzih.

Preprost primer predoločenega sistema

Poiščimo posplošeno rešitev 3×2 sistema

$$x = 0, \quad y = 0, \quad x + y = 1.$$

Sistem je očitno nerešljiv. Zapišimo ga v vektorski obliki

$$(0, 0, 1) = (x, y, x + y) = x(1, 0, 1) + y(0, 1, 1)$$

Razliko leve in desne strani skalarno množimo z $(1, 0, 1)$ in z $(0, 1, 1)$:

$$\langle x(1, 0, 1) + y(0, 1, 1) - (0, 0, 1), (1, 0, 1) \rangle = 0$$

$$\langle x(1, 0, 1) + y(0, 1, 1) - (0, 0, 1), (0, 1, 1) \rangle = 0.$$

Ko uredimo, dobimo sistem enačb

$$2x + y = 1, \quad x + 2y = 1.$$

Običajna rešitev tega sistema je $x = y = \frac{1}{3}$. To je potem tudi posplošena rešitev prvotnega sistema.

Regresijska premica in posplošitve

Iščemo tako premico $y = ax + b$, ki se najbolj prilega danim točkam

$$(x_1, y_1), \dots, (x_m, y_m).$$

Če bi šla iskana premica skozi vse dane točke, bi koeficienta a in b zadoščala naslednjemu $m \times 2$ sistemu linearnih enačb

$$ax_1 + b = y_1 \quad \dots \quad ax_m + b = y_m \quad (5)$$

Seveda se pri $m > 2$ skoraj nikoli ne zgodi, da bi dane točke ležale na isti premici. Zato se bomo zadovoljili s posplošeno rešitvijo sistema (5), ki jo dobimo tako, da minimiziramo izraz

$$(ax_1 + b - y_1)^2 + \dots + (ax_m + b - y_m)^2 \quad (6)$$

Opomba: Izraz (6) je enak $d_1^2 + \dots + d_m^2$, kjer je $d_i := |ax_i + b - y_i|$ razdalja med dano točko (x_i, y_i) in točko $(x_i, ax_i + b)$ na iskani premici.

Izraz (6) predelamo v vektorsko obliko

$$\|a(x_1, \dots, x_m) + b(1, \dots, 1) - (y_1, \dots, y_m)\|^2 \quad (7)$$

Iščemo taka a in b , pri katerih je izraz (7) minimalen. To je poseben primer naloge o pravokotni projekciji točke na parametrično podano ravnino v $\mathbb{R}^m$. Vemo, da mora biti vektor $a(x_1, \dots, x_m) + b(1, \dots, 1) - (y_1, \dots, y_m)$ pravokoten na vektorja $(x_1, \dots, x_m)$ in $(1, \dots, 1)$. Ko razpišemo izraza

$$\langle a(x_1, \dots, x_m) + b(1, \dots, 1) - (y_1, \dots, y_m), (x_1, \dots, x_m) \rangle = 0$$

$$\langle a(x_1, \dots, x_m) + b(1, \dots, 1) - (y_1, \dots, y_m), (1, \dots, 1) \rangle = 0$$

dobimo sistem dveh enačb za a in b :

$$\begin{aligned} a\left(\sum_{i=1}^m x_i^2\right) + b\left(\sum_{i=1}^m x_i\right) &= \sum_{i=1}^m x_i y_i \\ a\left(\sum_{i=1}^m x_i\right) + bm &= \sum_{i=1}^m y_i \end{aligned} \quad (8)$$

Rešitev tega sistema vstavimo v $y = ax + b$ in dobimo iskano premico.

Primer

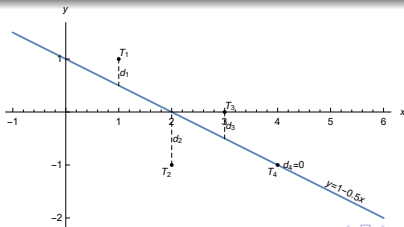
Poišči premico, ki se najboljše prilega točkam $(1, 1)$, $(2, -1)$, $(3, 0)$, $(4, -1)$.
Najprej izračunamo

$$\sum_{i=1}^4 x_i^2 = 30, \quad \sum_{i=1}^4 x_i = 10, \quad \sum_{i=1}^4 x_i y_i = -5, \quad \sum_{i=1}^4 y_i = -1$$

in rezultate vstavimo v sistem (8). Dobimo sistem enačb

$$30a + 10b = -5, \quad 10a + 4b = -1$$

katerega rešitev je $a = -\frac{1}{2}$, $b = 1$. Iskana premica je torej $y = 1 - 0.5x$.



Opomba: Premici, ki se najboljše prilega danim točkam pravimo tudi **regresijska premica**.

Opomba: Na podoben način poiščemo parabolo $y = ax^2 + bx + c$, ki se najboljše prilega danim točkam $(x_1, y_1), \dots, (x_m, y_m)$. Rešiti moramo sistem

$$a \sum x_i^4 + b \sum x_i^3 + c \sum x_i^2 = \sum x_i^2 y_i$$

$$a \sum x_i^3 + b \sum x_i^2 + c \sum x_i = \sum x_i y_i$$

$$a \sum x_i^2 + b \sum x_i + c \sum 1 = \sum y_i$$

Opomba: Podobno poiščemo ravnino $z = ax + by + c$, ki se najboljše prilega danim točkam $(x_1, y_1, z_1), \dots, (x_n, y_n, z_n)$. Rešiti moramo sistem

$$a \sum x_i^2 + b \sum x_i y_i + c \sum x_i = \sum x_i z_i$$

$$a \sum x_i y_i + b \sum y_i^2 + c \sum y_i = \sum y_i z_i$$

$$a \sum x_i + b \sum y_i + c \sum 1 = \sum z_i$$

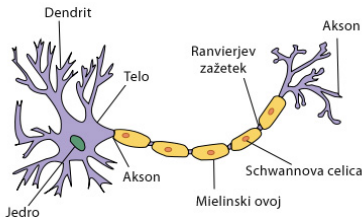
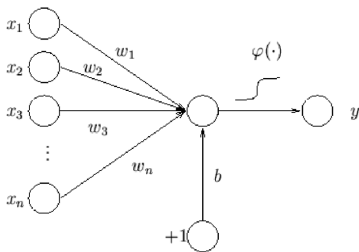
Tako nalogo dobimo, kadar treniramo perceptron.

Perceptron je poenostavljen model biološkega nevrona.

- Dendriti prejmejo vhodne signale $x_1, \dots, x_n$.
- Do telesa pride signal $z = w_1x_1 + \dots + w_nx_n + b$, kjer so w_i, b uteži.
- Akson pošlje izhodni signal $y = \varphi(z)$, kjer je φ aktivacijska funkcija. (Običajno je $\phi(z) = 1$, če $z > 0$ in 0 sicer. Pri nas je $\phi(z) = z$.)

Radi bi nevron natrenirali tako, da se bodo dobljeni izhodi čimbolj ujemali s pričakovanimi izhodi $(x_{1,1}, \dots, x_{1,n}, y_1), \dots, (x_{m,1}, \dots, x_{m,n}, y_m)$, kjer je $m > n + 1$. Iščemo torej take uteži $w_1, \dots, w_n, b$, ki "rešijo" sistem

$$w_1x_{i,1} + \dots + w_nx_{i,n} + b = y_i, \quad i = 1, \dots, m.$$



Poddoločeni sistemi

Spomnimo se, da je sistem linearnih enačb **poddoločen**, če ima manj enačb kot neznank.

Trditev

Če je poddoločen sistem linearnih enačb rešljiv, potem ima neskončno rešitev.

Dokaz: To sledi iz naslednjih dveh rezultatov o homogenih sistemih.

- Vsak poddoločen homogen sistem ima neskončno rešitev.
- Če je **p** rešitev nehomogenega sistema in če je **h** rešitev pripadajočega homogenega sistema, potem je **p + h** rešitev nehomogenega sistema.

Opomba: Seveda se lahko zgodi, da je nehomogen poddoločen sistem nerešljiv. Recimo 2×3 sistem $x + y + z = 1, x + y + z = 2$ je nerešljiv. Tudi 1×2 sistem $0x + 0y = 1$ je nerešljiv.

Kadar je poddoločen rešljiv, nas pogosto zanima njegova najkrajša rešitev.

Geometrijsko gledano iščemo pravokotno projekcijo izhodišča na presek hiperravnin. Ponovimo najprej, kako to naredimo za eno hiperravnino.

Iščemo presek hiperravnine $\langle \mathbf{n}, \mathbf{r} \rangle = b$ s premico $\mathbf{r} = t\mathbf{n}$, ki gre skozi izhodišče in je pravokotna na hiperravnino. Vstavimo drugo enačbo v prvo in izrazimo $t = \frac{b}{\langle \mathbf{n}, \mathbf{n} \rangle}$. Torej je iskana najkrajša rešitev $\mathbf{r} = \frac{b}{\langle \mathbf{n}, \mathbf{n} \rangle} \mathbf{n}$.

Podobno poiščemo pravokotno projekcijo izhodišča na presek dveh hiperravnin $\langle \mathbf{n}_1, \mathbf{r} \rangle = b_1$ in $\langle \mathbf{n}_2, \mathbf{r} \rangle = b_2$. Naloge se lotimo v dveh korakih:

- Poiščemo ravnino, ki gre skozi izhodišče in je pravokotna na presek obeh hiperravnin. To je ravnina $\mathbf{r} = t_1\mathbf{n}_1 + t_2\mathbf{n}_2$.
- Izračunamo presek te ravnine s presekom obeh hiperravnin. Enačbo ravnine vstavimo v enačbi hiperravnin in uredimo. Dobimo sistem

$$t_1 \langle \mathbf{n}_1, \mathbf{n}_1 \rangle + t_2 \langle \mathbf{n}_1, \mathbf{n}_2 \rangle = b_1, \quad t_1 \langle \mathbf{n}_2, \mathbf{n}_1 \rangle + t_2 \langle \mathbf{n}_2, \mathbf{n}_2 \rangle = b_2.$$

Rešimo po t_1 in t_2 in rešitev vstavimo v enačbo ravnine.

Enak postopek deluje tudi pri preseku več kot dveh hiperravnin.

Primer iskanja najkrajše rešitve poddoločenega sistema

Poišči najkrajšo rešitev sistema

$$\begin{aligned}2x - y + 3u - 2v &= 1, \\ x + y + 2u - 2v &= 0, \\ x + 2u - 3v &= -2.\end{aligned}$$

Najprej sistem zapišemo v obliki

$$\langle \mathbf{n}_1, \mathbf{r} \rangle = 1, \quad \langle \mathbf{n}_2, \mathbf{r} \rangle = 0, \quad \langle \mathbf{n}_3, \mathbf{r} \rangle = -2,$$

kjer je $\mathbf{r} = (x, y, u, v)$ in

$$\mathbf{n}_1 = (2, -1, 3, -2), \quad \mathbf{n}_2 = (1, 1, 2, -2), \quad \mathbf{n}_3 = (1, 0, 2, -3).$$

Rešitev iščemo z nastavkom $\mathbf{r} = t_1 \mathbf{n}_1 + t_2 \mathbf{n}_2 + t_3 \mathbf{n}_3$.

Dobimo sistem

$$\begin{aligned}18t_1 + 11t_2 + 14t_3 &= 1, \\11t_1 + 10t_2 + 11t_3 &= 0, \\14t_1 + 11t_2 + 14t_3 &= -2,\end{aligned}$$

katerega rešitev je

$$t_1 = \frac{3}{4}, \quad t_2 = \frac{22}{19}, \quad t_3 = -\frac{137}{76}.$$

Ko te vrednosti vstavimo v nastavek, dobimo

$$\mathbf{r} = t_1\mathbf{n}_1 + t_2\mathbf{n}_2 + t_3\mathbf{n}_3 = \left(\frac{65}{76}, \frac{31}{76}, \frac{73}{76}, \frac{121}{76}\right).$$

To je iskana najkrajša rešitev prvotnega sistema.

Opomba: Drug način reševanja je, da sistem najprej rešimo in nato projiciramo izhodišče na množico rešitev.