

# Funktsioonide lähendamine. Vähimruutude meetod.

Loeng 10

MTMM.00.005 Numbrilised meetodid

22.04.2019

- 1 Lineaarsete süsteemide lahendamine vähimruutude meetodil
- 2 Funktsioonide lähendamine vähimruutude meetodil
- 3 Polünoomidega lähendamine vähimruutude meetodil

Ülesanded 37-38

Vaatleme süsteemi

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = f_1, \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = f_m, \end{cases} \quad (1)$$

ehk  $Ax = f$ , kus

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}.$$

Üldiselt  $m \neq n$  ning süsteemil (1) ei pruugi olla lahendit.

Defineerime  $x \in \mathbb{R}^n$  korral vektori

$$R(x) = f - Ax = \begin{pmatrix} f_1 - \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ f_m - \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

ning vaatleme saadud vektori 2-normi ruutu (ehk eukleidilise normi ruutu)

$$\|R(x)\|^2 = \sum_{i=1}^m \left( f_i - \sum_{j=1}^n a_{ij}x_j \right)^2.$$

## Definitsioon

Vektorit  $x \in \mathbb{R}^n$  nimetatakse süsteemi (1) lahendiks vähimruutude mõttes (vähimruutude lahendiks), kui tema korral

$$\|R(x)\|^2 = \min_{y \in \mathbb{R}^n} \|R(y)\|^2.$$

Vähimruutude mõttes lahend  $x$  on tavaline lahend parajasti siis, kui  $\|R(x)\| = 0$ .

# Lineaarsete süsteemide lahendamine

Vaatleme funktsiooni  $x \rightarrow \|R(x)\|^2$  ehk  $n$  muutuja funktsiooni

$$g(x_1, \dots, x_n) = \sum_{i=1}^m \left( f_i - \sum_{j=1}^n a_{ij} x_j \right)^2.$$

Tegemist on ruutpolünoomiga. Kui  $x$  on  $g$  miinimumpunkt, siis

$$\frac{\partial g}{\partial x_k}(x) = 0, \quad k = 1, \dots, n,$$

ehk

$$\sum_{i=1}^m 2 \left( f_i - \sum_{j=1}^n a_{ij} x_j \right) (-a_{ik}) = 0, \quad k = 1, \dots, n,$$

ehk

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij} a_{ik} x_j = \sum_{i=1}^m a_{ik} f_i, \quad k = 1, \dots, n,$$

ehk

$$\sum_{j=1}^n \left( \sum_{i=1}^m a_{ik} a_{ij} \right) x_j = \sum_{i=1}^m a_{ik} f_i, \quad k = 1, \dots, n,$$

ehk

$$A^T A x = A^T f, \quad (2)$$

kus

$$A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

on transposeeritud maatriks.

Süsteemi (2) nimetatakse **normaalvõrrandite süsteemiks**. Maatriks  $A^T A$  on  $n \times n$  maatriks, sest  $A$  on  $m \times n$  maatriks ja  $A^T$  on  $n \times m$  maatriks, lisaks  $x \in \mathbb{R}^n$  ja  $A^T f \in \mathbb{R}^n$ .

# Lineaarsete süsteemide lahendamine

Oleme näidanud, et kui  $x$  on ülesande (1) lahend vähimruutude mõttes, siis ta on süsteemi (2) lahend. Näitame ka vastupidist.

Olgu  $x \in \mathbb{R}^n$  süsteemi (2) lahend, siis suvalise  $y \in \mathbb{R}^n$  korral

$$\begin{aligned}\|f - Ay\|^2 &= \|f - Ax + Ax - Ay\|^2 \\&= \langle f - Ax + A(x - y), f - Ax + A(x - y) \rangle = \\&= \|f - Ax\|^2 + 2\langle A(x - y), f - Ax \rangle + \|A(x - y)\|^2 = \\&= \|f - Ax\|^2 + 2\langle x - y, A^T(f - Ax) \rangle + \|A(x - y)\|^2 \\&\geq \|f - Ax\|^2,\end{aligned}$$

sest  $A^T(f - Ax) = 0$  ja  $\|A(x - y)\|^2 \geq 0$ . Lisaks on siin kasutatud skalaarkorrutise aditiivsust, kommutatiivsust ning omadust  $\langle Ax, b \rangle = \langle x, A^T b \rangle$ .

Niisiis on süsteemi (1) lahendamine vähimruutude mõttes samaväärne normaalkorrandite süsteemi (2) lahendamisega.



- 37. \*(1 p) Tõestada, et normaalkvõrrandite süsteemil on alati olemas lahend. Soovitus: tõestada, et  $\text{ran } A^T A = \text{ran } A^T$ , kus  $m \times n$  maatriksi  $A$  korral defineeritakse  $\text{ran } A = \{Ax \mid x \in \mathbb{R}^n\}$ .
- 38. Tõestada, et võrdusest  $A^T Ay = 0$  järeldeb, et  $Ay = 0$ .

## Teoreem

Normaalvõrrandite süsteem (2) on üheselt lahenduv parajasti siis, kui maatriksi  $A$  veerud on lineaarselt sõltumatud.

Tõestus. Olgu

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, a_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix},$$

siis  $y \in \mathbb{R}^n$  korral

$$\begin{aligned} Ay &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11}y_1 + \dots + a_{1n}y_n \\ \dots & \dots & \dots \\ a_{m1}y_1 + \dots + a_{mn}y_n \end{pmatrix} \\ &= y_1 a_1 + \dots + y_n a_n. \end{aligned}$$

## Nüüd

$$\begin{aligned} & a_1, \dots, a_n \text{ on lineaarselt sõltumatud} \quad \Leftrightarrow \\ \Leftrightarrow & \{y_1 a_1 + \dots + y_n a_n = 0 \Rightarrow y_1 = \dots = y_n = 0\} \quad \Leftrightarrow \\ \Leftrightarrow & \{Ay = 0 \Rightarrow y = 0\} \stackrel{(*)}{\Leftrightarrow} \{A^T Ay = 0 \Rightarrow y = 0\} \quad \Leftrightarrow \\ \Leftrightarrow & \text{süsteem (2) on üheselt lahenduv.} \quad \square \end{aligned}$$

(\*) Kui  $Ay = 0$ , siis  $A^T Ay = 0$ . Vastupidine on tõestatud ülesandes 38.

Seepärast leiab aset võrdus  $\ker A^T A = \ker A$ , kus  $\ker A = \{x \in \mathbb{R}^n \mid Ax = 0\}$ . Seda võrdust kasutab ka ülesande 37 üks võimalik lahendus.

## Näide 1

Vaatleme süsteemi

$$\begin{cases} x_1 = 6, \\ x_1 + x_2 = 0, \\ x_1 + 2x_2 = 0 \end{cases} \Leftrightarrow Ax = f, \text{ kus } A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, f = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

Sellel süsteemil lahend puudub. Kirjutame välja normaalvõrrandite süsteemi

$$A^T Ax = A^T f \Leftrightarrow \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}.$$

Sel süsteemil on ühene lahend  $x_1 = 5$ ,  $x_2 = -3$  ning see on algse süsteemi lahend vähimruutude mõttes, st iga  $y \in \mathbb{R}^2$ ,  $y \neq x$ , korral

$$\begin{aligned} \|f - Ay\|^2 &> \|f - Ax\|^2 = (6 - x_1)^2 + (0 - (x_1 + x_2))^2 + (0 - (x_1 + 2x_2))^2 \\ &= 1^2 + (-2)^2 + 1^2 = 6. \end{aligned}$$

## Näide 2

Vaatleme süsteemi

$$\begin{cases} x_1 + x_2 + x_3 = 6, \\ x_2 + 2x_3 = 0 \end{cases} \Leftrightarrow Ax = f,$$

kus  $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ ,  $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ ,  $f = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$ . Sellel süsteemil on lõpmata

palju lahendeid  $x = \begin{pmatrix} 6+t \\ -2t \\ t \end{pmatrix}$ ,  $t \in \mathbb{R}$ . Kirjutame välja normaalvõrrandite süsteemi

$$A^T Ax = A^T f \Leftrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix}.$$

## Näide 2 jätkub

Normaalvõrrandite süsteemil on lõpmata palju lahendeid, samad, mis algsel süsteemil. Iga lahendi  $x$  korral  $\|f - Ax\|^2 = 0$ .

Meenutame, et normaalvõrrandite süsteem on alati lahenduv, kuid ta on üheselt lahenduv parajasti siis, kui  $A$  veerud on lineaarselt sõltumatud.

## Näide 3

Vaatleme süsteemi

$$\begin{cases} x_1 + x_2 + x_3 = 6, \\ x_1 + x_2 + x_3 = 0 \end{cases} \Leftrightarrow Ax = f,$$

kus  $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ ,  $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ ,  $f = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$ . Sellel süsteemil lahendid

puuduvad. Kirjutame välja normaalkõrrandite süsteemi

$$A^T Ax = A^T f \Leftrightarrow \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix}.$$

Sel süsteemil on lõpmata palju lahendeid, mis rahuldavad tingimust  $x_1 + x_2 + x_3 = 3$ .

# Funktsioonide lähendamine

Olgu antud sõlmed  $x_0, \dots, x_m$  ja vastavad arvud  $f_0, \dots, f_m$  (näiteks mingi funktsiooni väärtused või lähisväärtused sõlmedes). Vaatleme lähendit

$$\varphi(x) = \sum_{j=0}^n c_j \varphi_j(x),$$

kus koordinaatfunktsioonid  $\varphi_0, \dots, \varphi_n$  on antud. Kui  $m = n$  ja kordajad  $c_j$  leitakse süsteemist  $\varphi(x_i) = f_i$ ,  $i = 0, \dots, m$ , ehk

$$\sum_{j=0}^n c_j \varphi_j(x_i) = f_i, \quad i = 0, \dots, m, \quad (3)$$

siis saadakse interpolant. Kui süsteemil (3) lahend puudub (mis on loomulik nähtus, kui  $m > n$ ), siis lahendades süsteemi (3) tundmatute  $c_j$  suhtes vähimruutude meetodil, räägitakse lähendamisest vähimruutude mõttes.



Antud juhul on süsteemi (3) maatriks

$$A = \begin{pmatrix} \varphi_0(x_0) & \dots & \varphi_n(x_0) \\ \dots & \dots & \dots \\ \varphi_0(x_m) & \dots & \varphi_n(x_m) \end{pmatrix},$$

s.t.  $a_{ij} = \varphi_j(x_i)$ . Seega on normaalkvõrrandite süsteemiks

$$\sum_{j=0}^n \left( \sum_{i=0}^m \varphi_k(x_i) \varphi_j(x_i) \right) c_j = \sum_{i=0}^m \varphi_k(x_i) f_i, \quad k = 0, \dots, n.$$

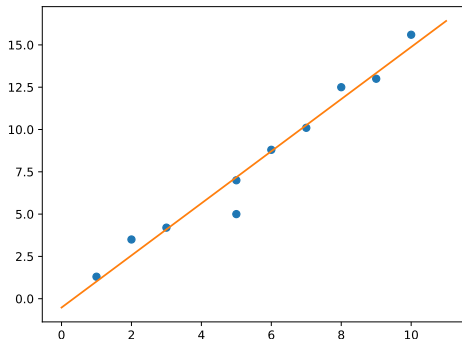
Normaalkvõrrandite süsteem on siin üheselt lahenduv parajasti siis, kui

maatriksi  $A$  veerud  $\begin{pmatrix} \varphi_j(x_0) \\ \vdots \\ \varphi_j(x_m) \end{pmatrix}$ ,  $j = 0, \dots, n$ , on lineaarselt sõltumatud.

# Funktsioonide lähendamine

Vähimruutude mõttes lähendamisel võib olla sõlmede  $x_0, \dots, x_m$  hulgas omavahel võrdseid, näiteks tehakse igal väärtusel  $x_i$  teatud arv (10 või 100) mõõtmisi. Seejuures võib olla  $x_i = x_j$ , aga  $f_i \neq f_j$ .

## Näide 4



Valime  $\varphi_0(x) = 1$ ,  $\varphi_1(x) = x$ ,  $\dots$ ,  $\varphi_n(x) = x^n$ , siis lähendiks on polünoom  $P_n(x) = c_0 + c_1x + \dots + c_nx^n$ . Süsteemi (3) kordajad on siin  $\varphi_j(x_i) = x_i^j$ , normaalvõrrandite süsteem on

$$\sum_{j=0}^n \left( \sum_{i=0}^m x_i^{k+j} \right) c_j = \sum_{i=0}^m x_i^k f_i, \quad k = 0, \dots, n. \quad (4)$$

## Teoreem

Normaalvõrrandite süsteem (4) on üheselt lahenduv parajasti siis, kui sõlmede  $x_0, \dots, x_m$  hulgas on vähemalt  $n + 1$  omavahel erinevat.

**Tõestus.** Tugineme eespool tõestatud teoreemile normaalvõrrandite süsteemi ühesest lahenduvusest. Selleks on tarvilik ja piisav, et lähtesüsteemi maatriksi

$$A = \begin{pmatrix} 1 & x_0 & \dots & x_0^n \\ \dots & \dots & \dots & \dots \\ 1 & x_m & \dots & x_m^n \end{pmatrix}$$

veerud oleksid lineaarselt sõltumatud. See leiab aset parajasti siis, kui maatriksi astak  $r(A)$  on võrdne veergude arvuga  $n + 1$ . Kui on olemas  $n + 1$  omavahel erinevat sõlme, siis on olemas  $n + 1$  järku nullist erinev miinor (Vandermonde'i determinant) ning seega  $r(A) = n + 1$ . Kui ei ole võimalik leida  $n + 1$  omavahel erinevat sõlme, siis igas  $n + 1$  järku miinoris on vähemalt kaks rida võrdsed ja miinor ise võrdne nulliga, mis tähendab, et  $r(A) < n + 1$ . □

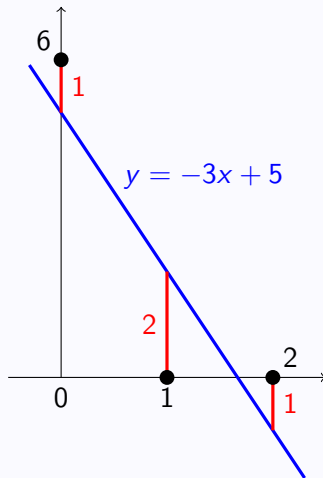
## Näide 1b

Olgu antud punktid  $(0, 6)$ ,  $(1, 0)$  ja  $(2, 0)$ . Leida sirge  $y = ax + b$ , mis lähendab neid andmeid vähimruutude mõttes. Peaksime lahendama süsteemi

$$\begin{cases} b = 6, \\ a + b = 0, \\ 2a + b = 0 \end{cases}$$

vähimruutude mõttes. Näite 1 põhjal  $a = -3$ ,  $b = 5$  ning  $y = -3x + 5$ .

## Näide 1b jätkub



## Näide 1b jätkub

Sirge  $y = ax + b$  on lahend vähimruutude mõttes  $\Leftrightarrow$

$$\sum_{i=0}^2 (y_i - (ax_i + b))^2 \rightarrow \min$$

Põhimõtteliselt võiks vaadelda ka ülesannet, kus

a)  $\max_{0 \leq i \leq 2} |y_i - (ax_i + b)| \rightarrow \min$  (lõpmatusnormi minimiseerimine)

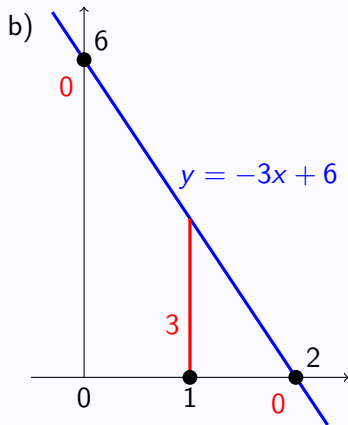
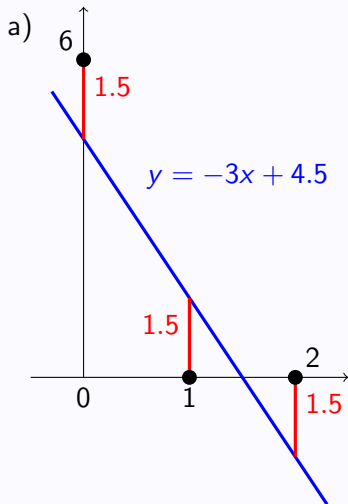
või

b)  $\sum_{i=0}^2 |y_i - (ax_i + b)| \rightarrow \min$  (1-normi minimiseerimine).

Variant a) on minimaks ülesanne, mida üldjuhul elementaarsete vaheditega lahendada ei ole võimalik, b) korral on üldjuhul probleemiks absoluutväärtus, mis ei ole diferentseeruv.

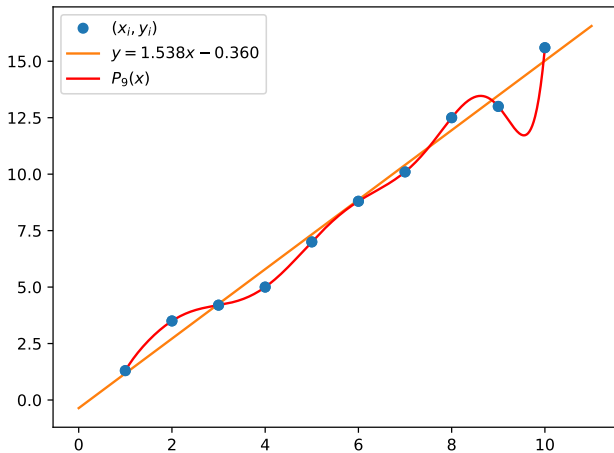
# Polünoomidega lähendamine

## Näide 1b jätkub



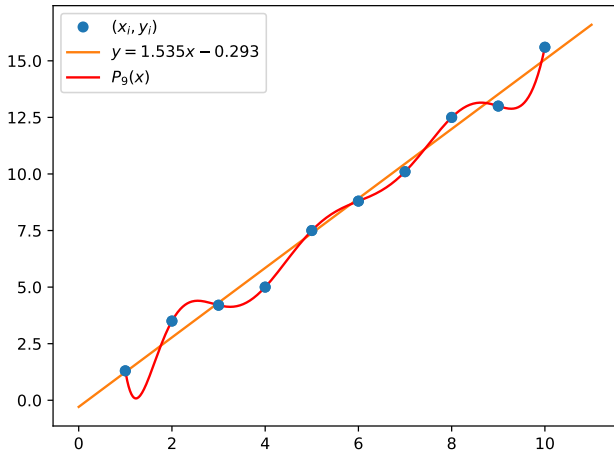


## Näide 5



# Polünoomidega lähendamine

## Näide 5 jätkub



# Polünoomidega lähendamine

## Näide 5 jätkub

