Zbrani zapiski za 3. letnik

Patrik Žnidaršič

Prevedeno dne 13. april 2024

Zahvala Jakobu Schraderju za raznovrstne popravke v zapiskih.

Kazalo

1 Analiza 3 **7**
1.1 Splošno . . . 8
1.2 Linearna NDE prvega reda . . . 9
1.3 Prvi integral enačbe . . . 11
1.4 Parametrično reševanje . . . 12
1.4.1 Lagrangeova in Clairontova enačba . . . 13
1.4.2 Ovojnice družin krivulj . . . 14
1.5 Enačbe drugega reda . . . 15
1.6 Eksistenčni izrek . . . 16
1.7 Sistemi linearnih NDE . . . 20
1.8 Linearne NDE višjega reda . . . 26
1.8.1 Enačbe s konstantnimi koeficienti . . . 28
1.8.2 Linearizacija . . . 30
1.9 Variacijski račun . . . 30
1.9.1 Vezani ekstremi . . . 34

2 Mehanika **37**
2.1 Osnove Newtonove mehanike . . . 38
2.2 Premočrtno gibanje . . . 44
2.3 Gibanje po krivulji . . . 46
2.4 Gibanje v polju centralne sile . . . 48
2.5 Relativno gibanje . . . 52
2.6 Sistem materialnih točk . . . 57
2.7 Togo telo . . . 58
2.7.1 Prosta vrtavka . . . 63
2.7.2 Eulerjevi koti . . . 64

3 Uvod v numerične metode **67**
3.1 Računske napake . . . 68
3.2 Nelinearne enačbe . . . 69
3.2.1 Bisekcija . . . 70
3.2.2 Navadna iteracija . . . 70
3.2.3 Tangentna metoda . . . 71
3.2.4 Sekantna metoda . . . 72
3.2.5 Ostale metode . . . 73
3.2.6 Ničle polinomov . . . 73

3.2.7 Durand-Kernerjeva metoda . . . 75
3.3 Sistemi linearnih enačb . . . 75
3.3.1 Matrične norme . . . 75
3.3.2 Občutljivost sistema linearnih enačb . . . 79
3.3.3 LU razcep . . . 80
3.3.4 Razcep Choleskega . . . 84
3.4 Sistemi nelineranih enačb . . . 85
3.5 Linearni problemi najmanjših kvadratov . . . 87
3.5.1 Normalni sistem . . . 87
3.5.2 QR razcep . . . 88
3.5.3 Gram-Schmittova ortogonalizacija . . . 88
3.5.4 Givensove rotacije . . . 89
3.5.5 Householderjeva zrcaljenja . . . 90
3.6 Lastne vrednosti . . . 91
3.6.1 Potenčna metoda . . . 92
3.6.2 Inverzna iteracija . . . 93
3.6.3 Ortogonalna iteracija . . . 94
3.6.4 QR iteracija . . . 95
3.7 Polinomska interpolacija . . . 98
3.7.1 Lagrangeova oblika . . . 98
3.7.2 Deljene diference . . . 99
3.8 Numerično integriranje . . . 100
3.8.1 Newton-Colesove formule . . . 101
3.8.2 Napake pri numeričnem integriranju . . . 102
3.8.3 Gaussove kvadraturne formule . . . 103
3.9 Diferencialne enačbe . . . 104
3.9.1 Runge-Kutta metode . . . 105

4 Verjetnost **107**
4.1 Izidi, dogodki, verjetnosti . . . 108
4.1.1 Pogojna verjetnost in neodvisnost . . . 110
4.1.2 Neodvisnost dogodkov . . . 111
4.2 Slučajne spremenljivke in porazdelitve . . . 112
4.2.1 Slučajni vektorji . . . 117
4.2.2 Neodvisnost slučajnih spremenljivk . . . 117
4.2.3 Pričakovana vrednost diskretnih spremenljivk . . . 118
4.2.4 Večrazsežne zvezne porazdelitve . . . 119
4.2.5 Pogojne pričakovane vrednosti . . . 122
4.3 Rodovne funkcije . . . 123
4.3.1 Procesi razvejanja . . . 125
4.3.2 Panjerjeva rekurzija . . . 125
4.4 Tabele . . . 126

5 Algebra 3 **127**
5.1 Reševanje polinomskih enačb . . . 128
5.1.1 Rešljive grupe . . . 132
5.1.2 Rešljivost polinomskih enačb z radikali . . . 133
5.2 Moduli . . . 134

6 Analiza 4 **137**
6.1 Osnovni tipi PDE . . . 138
6.2 Kvazilinearne enačbe prvega reda v dveh spremenljivkah . . . 139
6.2.1 Linearna PDE . . . 142
6.2.2 Ovojnica družine ravnin . . . 142
6.3 Nelinearne enačbe prvega reda . . . 143
6.3.1 Cauchyjeva naloga za PDE prvega reda . . . 145

7 Izbrane teme iz analize podatkov **147**
7.1 Linearna regresija . . . 149
7.2 Logistična regresija . . . 150
7.3 Najbližji sosedi . . . 151

8 Numerična linearna algebra **153**
8.1 Singularni razcep . . . 154

9 Statistika **157**
9.1 Centralni limitni izrek . . . 158
9.2 Konvergenca porazdelitev . . . 162

1 Analiza 3

1.1 Splošno

Definicija. Naj bo $F : I \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ zvezna funkcija in I interval v $\mathbb{R}$. Navadna diferencialna enačba prvega reda je enačba oblike $F(x, y(x), y'(x)) = 0$, kjer je $y(x)$ neka funkcija. Rešitev enačbe je vsaka funkcija $y_r(x) : I \to \mathbb{R}$, za katero velja enačba.

Opomba. NDE n-tega reda definiramo podobno kot enačbo oblike

$$F(x, y, y', y'', \dots, y^{(n)}) = 0.$$

Opomba. Smiselno je opazovati tudi enačbe, kjer je $F = (F_1, \dots, F_m)$ vektorska funkcija. Temu pravimo sistem NDE.

Opomba. Naj bo $y^{(n)} = F(x, y, y', \dots, y^{(n-1)})$ enačba reda n. Ta enačba je ekvivalentna primernemu sistemu $n \times n$ prvega reda; definirajmo $y_1 = y$, $y_2 = y'$, ..., $y_{n-1} = y^{n-2}$. Tedaj dobimo enačbo $y_n' = F(x, y_1, \dots, y_n)$.

Definicija. Integralska krivulja γ vektorskega polja $F : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ skozi točko $x_0 \in \Omega$ je krivulja $\gamma : [0, b) \to \Omega$, za katero velja

- v vsaki točki t je $\dot{\gamma}(t) = F(\gamma(t))$,
- $\gamma(0) = x_0$.

Vprašanje 1. Definiraj integralske krivulje.

Če prvi pogoj iz definicije zapišemo v koordinatah,

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix}(t) = \begin{bmatrix} F_1(x_1, \dots, x_n) \\ \vdots \\ F_n(x_1, \dots, x_n) \end{bmatrix},$$

dobimo sistem n NDE prvega reda z n neznankami. Ta sistem ni eksplicitno odvisen od t; takim sistemom pravimo avtonomni sistemi. Pokazali bomo, da za vsako izbiro x_0 obstajata interval $[0, a)$ in krivulja γ, za katero veljata pogoja v definiciji.

Vsak neavtonomen sistem lahko prepišemo v avtonomnega, z uvedbo nove odvisne spremenljivke $v(t) = t$. Dobimo nov sistem

$$\begin{aligned} \dot{v} &= 1, \\ \dot{x}_1 &= F_1(v, x_1, \dots, x_n), \\ &\vdots \\ \dot{x}_n &= F_n(v, x_1, \dots, x_n). \end{aligned}$$

Partikularna rešitev tega sistema je tedaj integralska krivulja vektorskega polja $\vec{F}(v, \vec{x})$ v razširjenem faznem prostoru $\mathbb{R} \times \Omega$ (Ω je običajen fazni prostor), ki ustreza primernemu začetnemu pogoju.

Vprašanje 2. Kako spremenimo neavtonomni sistem v avtonomnega?

V nekaterih primerih poznamo rešitev NDE. Če imamo enačbo z ločljivima spremenljivkama

$$\dot{x} = f(t)g(x),$$

lahko enačbo delimo z $g(x)$, in definiramo $h(x) = 1/g(x)$. Dobimo

$$h(x)\dot{x} = f(t).$$

Sedaj definiramo $H(x)$ kot primitivno funkcijo $h(x)$, in $F(t)$ kot primitivno funkcijo $f(t)$. Velja $\dot{H}(x) = \dot{F}(t)$, torej je $x(t) = H^{-1}(F(t) + C)$.

Vprašanje 3. Kako rešiš enačbo z ločljivima spremenljivkama?

Če imamo enačbo s homogeno desno stranjo, torej $\dot{x} = f(t, x)$, kjer velja $f(t, x) = f(\lambda t, \lambda x)$ za $\lambda \in \mathbb{R}\backslash\{0\}$, potem velja $f(t, x) = f(1, {}^{x}\!/_{t})$. Vpeljemo novo spremenljivko $v = {}^{x}\!/_{t}$, in s kratkim računom pridemo do $\dot{x} = t\dot{v} + v$. Po drugi strani velja $\dot{x} = f(1, v)$, torej

$$\dot{v} = \frac{1}{t}\left(f(1, v) - v\right).$$

To je enačba z ločljivima spremenljivkama, ki jo znamo rešiti.

Vprašanje 4. Kako rešiš enačbo s homogeno desno stranjo?

1.2 Linearna NDE prvega reda

Linearna NDE prvega reda je enačba oblike

$$y' = f(x)y + g(x),$$

kjer sta $f(x)$ in $g(x)$ znani funkciji. Ta enačba je nehomogena z nehomogenostjo $g(x)$. Njena homogenizacija je enačba

$$y' = f(x)y.$$

Predpostavimo, da je $y(x) \in \mathcal{C}^1([a, b])$. Oglejmo si operator $A : \mathcal{C}^1([a, b]) \to \mathcal{C}([a, b])$, definiran kot

$$Ay(x) = y'(x) - f(x)y.$$

Trditev. *Preslikava A je linearen operator.*

Dokaz je trivialen, in zato izpuščen. Vidimo, da je y rešitev homogene enačbe natanko tedaj, ko je $A(y) = 0$. Rešitev homogene enačbe je torej jedro preslikave A. Homogena enačba je enačba z ločljivimi spremenljivkami, torej jo znamo rešiti. Rešitve so oblike

$$y(x) = C \exp\left(\int_a^x f(\xi) d\xi\right)$$

za $C \in \mathbb{R}$. Množico teh rešitev označimo z R_h.

Trditev. *Naj bosta y_1, y_2 rešitvi nehomogene enačbe. Tedaj je $y(x) = y_1(x) - y_2(x)$ rešitev homogene enačbe.*

Dokaz. Izračun odvoda nam da

$$(y_1(x) - y_2(x))' = f(x)y_1(x) + g(x) - (f(x)y_2(x) + g(x)) = f(x)(y_1(x) - y_2(x)).$$

□

Definicija. Naj bo V nek vektorski prostor in $W \subseteq V$. Če obstaja tak vektorski podprostor $H \subseteq V$, da za poljubna $w_1, w_2 \in W$ velja $w_1 - w_2 \in H$, je W AFIN PODPROSTOR v V, modeliran z vektorskim podprostorom H.

Rešitve nehomogene enačbe so torej afin prostor, modeliran s prostorom R_h rešitev homogene enačbe. Če želimo poiskati splošno rešitev, poiščemo rešitev homogenega sistema, in neko partikularno rešitev. Partikularno rešitev dobimo z nastavkom

$$y_p(x) = C(x) \exp\left(\int_a^x f(\xi)d\xi\right),$$

temu postopku pravimo VARIACIJA KONSTANTE.

Vprašanje 5. Kako rešiš linearno NDE prvega reda? Utemelji postopek.

S tem znanjem lahko rešimo še dve posebni NDE. Prva je Bernoulijeva enačba

$$p(x)y' + q(x)y = r(x)y^\alpha(x)$$

za $\alpha \in \mathbb{R}$. V primeru $\alpha = 0$ ali $\alpha = 1$, je to nehomogena linearna enačba prvega reda. Sicer vpeljemo $z(x) = (y(x))^{1-\alpha}$ in računamo

$$p(x)z'(x)\frac{1}{1-\alpha} + q(x)z(x) = r(x)$$

oziroma

$$z'(x) + \frac{q(x)}{p(x)}(1-\alpha)z(x) = \frac{r(x)}{p(x)}(1-\alpha).$$

To je nehomogena linearna NDE prvega reda, torej jo znamo rešiti.

Vprašanje 6. Kako rešiš Bernoulijevo enačbo?

Druga taka enačba je Riccatijeva enačba

$$y'(x) = a(x)y^2(x) + b(x)y(x) + c(x),$$

ki je v splošnem ne znamo rešiti. Poznamo pa dva načina obravnave, ki nas lahko včasih pripeljeta do rešitve. Denimo, da uganemo neko partikularno rešitev $y_p(x)$. Enačbo tedaj rešujemo z nastavkom $y(x) = y_p(x) + z(x)$ za neko neznano funkcijo z. Če to vstavimo v enačbo, dobimo

$$y_p' + z' = ay_p^2 + 2ay_pz + az^2 + by_p + bz + c,$$

členi y_p', ay_p^2, by_p in c odpadejo, ker tvorijo rešitev enačbe. Ostane torej

$$z' = (2ay_p + b)z + az^2,$$

kar je Bernoulijeva enačba, ki jo znamo rešiti.

Drug način za reševanje Riccatijeve enačbe je s pretvorbo na linearni sistem prvega reda. Vpeljemo $y = {}^{u}\!/_{v}$, s čimer dobimo

$$u'v - uv' = au^2 + buv + cv^2.$$

Ker imamo dve neznanki, potrebujemo še eno enačbo. Izberemo $u'v = buv + cv^2$. Iz tega izpeljemo $v' = -au$ in $u' = bu + cv$. Zapisano matrično

$$\begin{bmatrix} v' \\ u' \end{bmatrix} = \begin{bmatrix} 0 & -a \\ c & b \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix}$$

Sistema v splošnem ne znamo rešiti, ker funkcije a, b, c niso konstantne. Lahko pa rešitev zapišemo v obliki neskončne vrste.

Vprašanje 7. Kako rešiš Riccatijevo enačbo?

1.3 Prvi integral enačbe

Splošna rešitev enačbe $y' = f(x, y)$ je enoparametrična družina funkcij $y = \phi(x, C)$. Denimo, da obstaja taka funkcija $u(x, y) : [a, b] \times M \to \mathbb{R}$ na razširjenem faznem prostoru, da zanjo velja $u(x, y(x, C)) = \text{konst.}$ za vsak C (konstanta je lahko drugačna za različne C). Taki funkciji pravimo PRVI INTEGRAL ENAČBE. Recimo, da velja $\partial_y u \neq 0$. Potem lahko iz enakosti izračunamo funkcijo $y(x, D)$, da velja $u(x, y(x, D)) = D$.

Trditev. *Vsaka krivulja $y(x)$, ki je implicitno podana z enačbo $u(x, y) = D$, kjer je u prvi integral enačbe $y' = f(x, y)$, je rešitev te enačbe.*

Dokaz. Naj bo $y_0(x)$ dana krivulja. Tedaj velja $u(x, y_0(x)) = D$. Če odvajamo po x, dobimo

$$\partial_x u + y_0' \partial_y u = 0,$$

torej

$$y_0' = -\frac{\partial_x u}{\partial_y u}.$$

Po drugi strani za vsako rešitev velja $y' = f(x, y)$, iz česar izpeljemo

$$f(x, y) = -\frac{\partial_x u}{\partial_y u}.$$

Sledi, da je y_0 res rešitev enačbe. $\square$

Vsaka diferencialna enačba ima neskončno mnogo prvih integralov, vsakega lahko še transformiramo s poljubno $\psi : \mathbb{R} \to \mathbb{R}$.

Vprašanje 8. Kaj je prvi integral enačbe? Kako iz njega dobiš rešitev enačbe?

Imejmo dano vektorsko polje $F : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, podano z

$$F(x,y) = \begin{bmatrix} P(x,y) \\ Q(x,y) \end{bmatrix}$$

Poiščimo družino krivulj, ortogonalnih na polje $F(x,y)$, in jih parametrizirajmo z $\gamma(x) = (x, y(x))$. Izpeljemo lahko pogoj

$$y'(x) = -\frac{P(x,y)}{Q(x,y)},$$

kar je diferencialna enačba prvega reda. Recimo, da je polje potencialno. Tedaj obstaja taka funkcija $u : [a,b] \times \mathbb{R} \to \mathbb{R}$, da velja $\partial_x u = P$ in $\partial_y u = Q$. Krivulje, ki so ortogonalne na $\vec{\nabla}.u$, so natanko izohipse ploskve $(x, y, u(x,y))$. Za izohipso velja $u(x, y(x)) = C$, iz česar z odvajanjem izpeljemo

$$y' = -\frac{\partial_x u}{\partial_y u}.$$

Potencial u je torej prvi integral zgornje enačbe. Ker lahko potencial poiščemo z integralom, enačbo v tem primeru znamo rešiti.

Če polje ni potencialno, imamo še vedno ortogonalne krivulje, torej še vedno velja $y' = -{}^{P}/_{Q}$. Denimo, da je $y(x,C)$ splošna rešitev. Če lahko poiščemo prvi integral u, mora obstajati funkcija $\lambda(x,y)$, za katero je polje $(\lambda P, \lambda Q)$ potencialno, oziroma $\partial_x u = \lambda P$ in $\partial_y u = \lambda Q$. Taki funkciji pravimo INTEGRIRUJOČI MNOŽITELJ. Če ga lahko najdemo, lahko rešimo enačbo; tega pa v splošnem ne znamo.

Vprašanje 9. Kako poiščeš družino krivulj, pravokotnih na dano vektorsko polje $F = (P, Q)$? Kaj je integrirujoči množitelj?

1.4 Parametrično reševanje

Naj bo NDE prvega reda podana implicitno,

$$F(x, y, y') = 0.$$

Denimo, da y' ne moramo eksplicitno izraziti z x, y, ali pa je ekspliciten izraz nepripraven. Na F poglejmo nekoliko drugače; vsaka dovolj lepa funkcija treh spremenljivk podaja družino ploskev, $F(\xi, \eta, \zeta) = C$ je implicitna enačba ploskve v $\mathbb{R}^3$ za vsak $C \in \mathbb{R}$. Tako podano ploskev lahko parametriziramo. Naj bo $(u,v) \mapsto (\varphi(u,v), \psi(u,v), \chi(u,v))$ neka parametrizacija. Imamo tri pristope za reševanje, v odvisnosti od F.

Če y ne nastopa eksplicitno, torej $F(x, y') = 0$, nam enačba definira krivuljo. Parametriziramo jo z $\xi = \varphi(t)$ in $\eta = \psi(t)$, da velja $F(\varphi, \psi) = 0$. Za poljubno rešitev $t \mapsto (x(t), y(t))$ velja $\dot{y} = y'\dot{x}$, torej za $\varphi(t) = x(t)$ in $\psi(t) = y'(t)$ dobimo $\dot{y} = \chi\dot{\varphi}$, oziroma

$$y(t) = \int_0^t \chi(\tau)\dot{\varphi}(\tau)d\tau.$$

Dobimo parametrično izraženo rešitev $t \mapsto (\varphi(t), y(t))$.

Vprašanje 10. Kako parametrično rešiš enačbo $F(x, y') = 0$?

Če x ne nastopa eksplicitno, torej $F(y, y') = 0$, dobimo enačbo krivulje $F(\xi, \eta) = 0$, ki jo parametriziramo s $t \mapsto (\chi(t), \psi(t))$. Če označimo $\psi = y$ in $\chi = y'$, velja $\dot{\psi} = \chi\dot{x}$ oziroma

$$x(t) = \int_0^t \frac{\dot{\psi}(\tau)}{\chi(\tau)} d\tau,$$

torej je $t \mapsto (x(t), \psi(t))$ parametrično podana rešitev.

Vprašanje 11. Kako parametrično rešiš enačbo $F(y, y') = 0$?

V splošnem nam $F(x, y, y') = 0$ definira ploskev. Parametriziramo jo kot zgoraj z

$$x = \varphi(u, v) \qquad y = \psi(u, v) \qquad y' = \chi(u, v)$$

Naj bo $t \mapsto (x(u(t), v(t)), y(u(t), v(t)))$ neka rešitev naše enačbe. Potem je $t \mapsto (\varphi, \psi, \chi)$ krivulja na ploskvi. V nadaljevanju predpostavimo, da je preslikava $(u, v) \mapsto (x, y)$ obrnljiva, in izračunajmo

$$\begin{aligned} \dot{y} &= y_u\dot{u} + y_v\dot{v} = \psi_u\dot{u} + \psi_v\dot{v}, \\ \dot{x} &= \varphi_u\dot{u} + \varphi_v\dot{v}. \end{aligned}$$

Ker tudi tu velja $\dot{y} = y'\dot{x}$, izrazimo

$$u' = \frac{\dot{u}}{\dot{v}} = -\frac{\psi_v - \chi\varphi_v}{\psi_u - \chi\varphi_u}.$$

Dobili smo eksplicitno enačbo prvega reda v spremenljivki $u = u(v)$.

Vprašanje 12. Kako parametrično rešiš $F(x, y, y') = 0$?

1.4.1 Lagrangeova in Clairontova enačba

Lagrangeova enačba je enačba oblike

$$y = x\varphi(y') + \psi(y').$$

Rešujemo jo parametrično; $x = u$, $y' = v$ in $y = u\varphi(v) + \psi(v)$. Velja $dy = y'dx$, iz česar izpeljemo

$$(\varphi(v) - v)du + (u\varphi'(v) + \psi'(v))dv = 0.$$

Če je $\varphi(v) \neq v$, dobimo

$$(\phi(v) - v)\frac{du}{dv} + u\varphi'(v) + \psi'(v) = 0,$$

kar je linearna diferencialna enačba prvega reda, če pa je $\varphi(v) = v$, pa imamo Clairontovo enačbo

$$y = xy' + \psi(y').$$

To predelamo v

$$(u + \psi'(v))dv = 0,$$

in obravnavamo dva primera. Če je $dv = 0$, je y' konstanta, torej dobimo družino rešitev $y = Cx + \psi(C)$ (to vstavimo v enačbo; ni nujno vsak C dober). Če pa je $u + \psi'(v) = 0$, pa dobimo še eno rešitev.

Vprašanje 13. Kaj sta Lagrangeova in Clairontova enačba? Kako ju rešimo?

1.4.2 Ovojnice družin krivulj

Imejmo družino krivulj, podano implicitno z enačbo $F(x, y, C) = 0$. Denimo, da je družina taka, da obstaja krivulja, ki se v vsaki svoji točki dotika natanko enega člana družine. Taki krivulji pravimo OVOJNICA družine. Smiselno jo je parametrizirati z $C \mapsto (x(C), y(C))$, pri čemer se ovojnica v točki $(x(C), y(C))$ dotika člana družine s tem C. Denimo, da je parametrizacija regularna, torej za vsak C

$$(\partial_C x)^2 + (\partial_C y)^2 \neq 0.$$

Definirajmo

$$\phi(C) = F(x(C), y(C), C).$$

Ker funkcija izračuna F v točki na krivulji, je $\phi = 0$. Torej

$$\phi'(C) = \partial_x F \partial_C x + \partial_y F \partial_C y + \partial_C F = 0.$$

Če je $t \mapsto (x(t), y(t))$ parametrizacija C_0-tega člana družine, velja

$$\partial_t F(x(t), y(t), C_0) = \partial_x F \dot{x} + \partial_y F \dot{y} = 0$$

v točki dotika z ovojnico. Vektor $[\dot{x}, \dot{y}]^T$ je vzporeden z $[\partial_C x, \partial_C y]^T$ v tej točki, torej je

$$\begin{bmatrix} \partial_x F \\ \partial_y F \end{bmatrix} \perp \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \parallel \begin{bmatrix} \partial_C x \\ \partial_C y \end{bmatrix}$$

in zato

$$\partial_x F \partial_C x + \partial_y F \partial_C y = 0.$$

Torej v točki dotika velja

$$\partial_C F = 0.$$

Iz para enačb $F(x, y, C) = 0$ in $\partial_C F = 0$ dobimo vse točke na ovojnici.

Vprašanje 14. Kaj je ovojnica družine krivulj? Kako jo izračunaš? Izpelji.

1.5 Enačbe drugega reda

Najpomembnejša enačba drugega reda je drugi Newtonov zakon. Malce posplošeno ima obliko

$$\ddot{x}_i = F_i(x_1, \dots, x_n)$$

za $i = 1, \dots, n$. Če vpeljemo $p = \dot{x}$ in $q = x$, dobimo sistem prvega reda

$$\begin{aligned} \dot{q}_i &= p_i \\ \dot{p}_i &= F_i(q_1, \dots, q_n) \end{aligned}$$

Sistem lahko še posplošimo. Naj bosta $F, G : M \subseteq \mathbb{R}^{2n} \to \mathbb{R}^n$ preslikavi iz faznega prostora M. Zanima nas časovni razvoj sistema, ki je podan z enačbami

$$\begin{aligned} \dot{q} &= G(q, p), \\ \dot{p} &= F(q, p). \end{aligned}$$

Posebej pomembni so sistemi, za katere obstaja funkcija $H : M \to \mathbb{R}$, za katero velja

$$\begin{aligned} \frac{\partial H}{\partial q_i} &= -F_i, \\ \frac{\partial H}{\partial p_i} &= G_i. \end{aligned}$$

Taki funkciji pravimo Hamiltonian.

Izrek. *Če Hamiltonian obstaja, potem je prvi integral sistema.*

Dokaz. Naj bo $t \mapsto (q(t), p(t))$ neka rešitev sistema. Potem imamo

$$\partial_t H(q, p) = \partial_q H \dot{q} + \partial_p H \dot{p} = \begin{bmatrix} -F & F \end{bmatrix} \cdot \begin{bmatrix} G \\ G \end{bmatrix} = 0$$

$\square$

Vprašanje 15. Kaj je Hamiltonian? Dokaži, da je prvi integral.

Definicija. Naj bo podana funkcija $H : M \subseteq \mathbb{R}^{2n} \to \mathbb{R}$. Sistem enačb

$$\begin{aligned} \dot{q} &= \partial_p H \\ \dot{p} &= -\partial_q H \end{aligned}$$

je hamiltonski sistem s hamiltonsko funkcijo H.

Da bo sistem $\dot{q} = G(q, p), \dot{p} = F(q, p)$ Hamiltonski, mora obstajati funkcija H, za katero velja $\partial_p H = G$ in $\partial_q H = -F$. Zapisano v drugačni obliki

$$\begin{bmatrix} \dot{q} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \cdot \begin{bmatrix} \partial_q H \\ \partial_p H \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \vec{\nabla}.H,$$

torej

$$\vec{\nabla}.H = \begin{bmatrix} -F \\ G \end{bmatrix}.$$

Da bo sistem hamiltonski, mora biti torej polje $[-F, G]^T$ potencialno.

Vprašanje 16. Pod katerim pogojem je sistem hamiltonski? Dokaži.

Definicija. Eliptični integral prve vrste z modulom m je funkcija, podana s predpisom

$$F(x; m) = \int_0^x \left(1 - m \sin^2 \xi\right)^{-1/2} d\xi.$$

Inverzna funkcija te funkcije se imenuje Jacobijeva amplituda, velja

$$y = F(x; m) \Leftrightarrow x = \mathrm{am}(y; m).$$

Vprašanje 17. Obravnavaj gravitacijsko nihalo.

Odgovor: Gravitacijsko nihalo je oblike

$$\ddot{q} = -\sin q.$$

Temu sistemu pripada hamiltonska funkcija

$$H(q, p) = \frac{1}{2}p^2 - (\cos q - 1).$$

Vzdolž neke rešitve $t \mapsto (q(t), p(t))$ je to konstanta, in velja

$$\frac{1}{2}\dot{q}^2 - \cos q + 1 = E.$$

Enačbo lahko prevedemo v

$$\frac{dq}{\sqrt{1 - \frac{2}{E}\sin^2 q/2}} = \sqrt{2E}dt.$$

Rešitev je

$$q = 2\,\mathrm{am}\left(\sqrt{\frac{E}{2}}t + C; \frac{2}{E}\right).$$

⊠

1.6 Eksistenčni izrek

Izrek (Eksistenčni)**.** *Naj bo vektorsko polje $F(t, x)$ podano na valju*

$$\mathcal{C}_{a,b} = \{(t, x) \mid |t - t_0| \le a, \|x - x_0\| \le b\}$$

za neki par t_0, x_0 in $a, b \in \mathbb{R}^+$. Naj bo $F(t,x)$ na $\mathcal{C}_{a,b}$ zvezno in naj bo $F(t,x) : \mathcal{C}_{a,b} \to \mathbb{R}^n$ Lipschitzova glede na x pri vsakem t. Alternativno je lahko preslikava $F(t,x)$ odvedljiva po x pri vsakem t in $\|D_x F\|$ na $\mathcal{C}_{a,b}$ omejeno število. Potem obstaja natanko ena rešitev začetnega problema

$$\dot{x} = F(t,x) \qquad\qquad x(t_0) = x_0$$

za vsak x_0. Rešitev $\varphi(t)$ obstaja na $[t_0 - a', t_0 + a']$ za nek $a' \leq a$. Še več: za družino začetnih problemov $\dot{x} = F(t,x), x(t_0) = \hat{x}$, kjer je $\|x_0 - \hat{x}\|$ dovolj majhno, obstajata $0 < a' \leq a$ in funkcija $g(t,x) : \mathcal{C}_{a',b'} \to \mathbb{R}^n$, za katero velja

- *je zvezna na obe spremenljivki,*
- *$\partial_t g(t,x) = F(t,x)$,*
- *$g(t_0, \hat{x}) = \hat{x}$.*

Vprašanje 18. Formuliraj eksistenčni izrek.

Začetni problem $\dot{x} = F(t,x), x(t_0) = x_0$ je ekvivalenten integralski enačbi

$$x(t) = x_0 + \int_{t_0}^{t} F(\tau, x(\tau)) d\tau.$$

Dokaz bo uporabil Picardov operator

$$Af(x) = x_0 + \int_{t_0}^{t} F(\tau, f(\tau)) d\tau$$

na posebnem funkcijskem prostoru.

Naj bo

$$\mathcal{C}_{a,b} = \{(t,x) \mid |t - t_0| \leq a, \|x - x_0\| \leq b\}.$$

Predpostavimo, da F ustreza predpostavkam izreka. Naj bo L Lipschitzova konstanta za F na $\mathcal{C}_{a,b}$ in

$$C = \max_{(t,x) \in \mathcal{C}_{a,b}} \|F(t,x)\|.$$

Velikokrat lahko za L vzamemo kar maksimum norme Jacobijeve matrike na $\mathcal{C}_{a,b}$. Označimo s K_0 stožec

$$K_0 = \{(t,x) \mid |t - t_0| \leq a', \|x - x_0\| \leq C\,|t - t_0|\},$$

kjer je a' dovolj majhen, da velja $K_0 \subseteq \mathcal{C}_{a,b}$. Sedaj sprostimo začetno vrednost x_0. Naj bo $\|\hat{x} - x_0\| < b'$, $K_{\hat{x}} = K_0 + (\hat{x} - x_0)$ premik prostora in

$$K = \bigcup_{\|\hat{x} - x_0\| < b'} K_{\hat{x}},$$

kjer je b' dovolj majhen, da je $K \subseteq \mathcal{C}_{a,b}$. Za nas bo pomemben nov prostor $\mathcal{C}_{a',b'}$.

Začetni problem zapišimo nekoliko drugače. Spomnimo se: Iščemo $g(t, \hat{x})$, da bo $\partial_t g = F(t, g)$ in $g(t_0, \hat{x}) = \hat{x}$. Vpeljimo novo funkcijo $h(t, x) : \mathcal{C}_{a,b} \to \mathbb{R}^n$,

$$g(t, x) = x + h(t, x).$$

Velja

$$\begin{aligned} \partial_t h(t, x) &= F(t, g(t, x)) \\ h(t_0, x) &= g(t_0, x) - x = 0. \end{aligned}$$

Torej je h pri vsakem x rešitev začetnega problema $\dot{h} = F(t, x + h(t, x)), h(t_0, x) = 0$.

Definiramo

$$M = \{h(t, x) : \mathcal{C}_{a',b'} \to \mathbb{R}^n \mid h \,\text{zvezna}, \|h(t, x)\| \leq C\, |t - t_0|\}.$$

Te preslikave zavzemajo vrednosti v stožcu K_0.

Vprašanje 19. Povej postopek konstrukcije funkcijskega prostora v dokazu eksistenčnega izreka.

Opremimo M z maksimum normo (in s tem z metriko in topologijo)

$$\|h\| = \max_{(t,x) \in \mathcal{C}_{a',b'}} \|h(t, x)\|.$$

Trditev. *Prostor M je poln metrični prostor.*

Dokaz. Naj bo $\{h_n(t, x)\}_n$ Cauchyjevo zaporedje v M. Ker je $\mathbb{R}^n$ poln, obstaja limita $\lim_{n\to\infty} h_n(t, x)$ za poljubna t, x. Ker je norma definirana z maksimumom, je konvergenca glede na to normo enakomerna, torej je

$$h(t, x) = \lim_{n\to\infty} h_n(t, x)$$

zvezna. Če je $h_n \in M$, velja $\|h_n(t, x)\| \leq C\, |t - t_0|$. To očitno velja tudi v limiti. □

Vprašanje 20. Dokaži, da je ta funkcijski prostor poln.

Našo rešitev poiščemo kot limito iteracij Picardove preslikave. Označimo

$$h_n(t, x) = A^n(h_0(t, x))$$

za $h_0 = 0$. Dokazati moramo, da za vsak $n \in \mathbb{N}$ velja $\|h_n(t, x)\| \leq C\, |t - t_0|$. To naredimo z indukcijo na n. Pri $n = 0$ to očitno velja, indukcijski korak pa pokažemo z računom

$$\|h_{n+1}(t, x)\| = \left\| \int_{t_0}^{t} F(\tau, x + h_n(\tau, x)) d\tau \right\| \leq \int_{t_0}^{t} \|F(\tau, x + h_n(\tau, x))\| \, d\tau.$$

Po indukcijski predpostavki vemo, da točka $h_n(t,x)$ leži v K_0 za vsak (t,x), zato $x + h_n(t,x)$ leži v $K_x \subseteq K \subseteq \mathcal{C}_{a,b}$. Sledi

$$\|F(\tau, x + h_n(\tau, x))\| \leq C,$$

zato

$$\|h_{n+1}(t,x)\| \leq \left| \int_{t_0}^{t} C d\tau \right| = C \left| t - t_0 \right|.$$

Pokazali smo, da je $h_n \in M$ za vsak n. Ker je M poln, je tudi limita v M, če obstaja.

Pokazati moramo še, da je A na M skrčitev. Naj bosta $h_1, h_2 \in M$ poljubni. Oglejmo si

$$\begin{aligned}\|Ah_1(t,x) - Ah_2(t,x)\| &= \left\| \int_{t_0}^{t} F(\tau, x + h_1(\tau,x)) - F(\tau, x + h_2(\tau,x)) d\tau \right\| \\ &\leq \int_{t_0}^{t} \|F(\tau, x + h_1(\tau,x)) - F(\tau, x + h_2(\tau,x))\| \, d\tau.\end{aligned}$$

Ker je F Lipschitzova glede na x, velja

$$\begin{aligned}\|Ah_1(t,x) - Ah_2(t,x)\| &\leq \int_{t_0}^{t} L \, \|h_1(\tau,x) - h_2(\tau,x)\| \, d\tau \\ &\leq \int_{t_0}^{t} L \, \|h_1 - h_2\| \, d\tau \\ &= L \, \|h_1 - h_2\| \, |t - t_0| \\ &\leq L \, \|h_1 - h_2\| \, a'\end{aligned}$$

Po potrebi še zmanjšamo a', da bo $La' < 1$.

Ker je A skrčitev, limita zaporedja h_n obstaja in je fiksna točka preslikave A. Torej je rešitev začetnega problema

$$\partial_t h = F(t, x + h(t,x)) \qquad\qquad h(t_0, x) = 0,$$

iz katere dobimo preslikavo $g(t,x) = x + h(t,x)$. Naša limita je po konstrukciji zvezna (ker leži v M), torej je tudi g zvezna glede na oba argumenta.

To je konec dokaza eksistenčnega izreka.

Vprašanje 21. Dokaži eksistenčni izrek.

Trditev. *Naj bo $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ zvezno odvedljiva na konveksnem kompaktu $M \subseteq U$. Potem je na M Lipschitzova.*

Dokaz. Naj bosta $x, y \in M$ poljubni točki. Definiramo $z(t) = x + t(y - x)$ kot daljico med x in y. Velja

$$f(y) - f(x) = \int_0^1 \partial_\tau f(z(\tau)) d\tau = \int_0^1 Df(z(\tau))(y - x) d\tau.$$

Ker je f zvezno odvedljiva na kompaktu, norma Jacobijeve matrike doseže maksimum, torej velja

$$\|f(x) - f(y)\| = \left\| \int_0^1 Df(z(\tau))(y-x)d\tau \right\| \leq \int_0^1 \|Df\| \, \|y-x\| \, d\tau = \|Df\| \, \|y-x\| \, .$$

□

Vprašanje 22. Dokaži, da je zvezno odvedljiva funkcija na konveksnem kompaktu Lipschitzova.

Imejmo NDE $\dot{x} = F(t,x)$. Tok te enačbe je preslikava

$$\phi : (a,b) \times (\alpha, \beta) \times U \to \mathbb{R}^n,$$

definirana s predpisom

$$\phi(t, t_0, x) = \gamma(t),$$

kjer je γ rešitev začetnega problema $\dot{\gamma}(t) = F(t, \gamma(t)), \gamma(t_0) = x$.

Trditev. *Za tok enačbe velja*

- *Za vsak t_0, za katerega rešitve začetnih problemov $\gamma(t_0) = x$ obstajajo, in za vsak t dovolj blizu t_0, je $x \mapsto \phi(t, t_0, x)$ difeomorfizem U na svojo sliko.*
- *Za t_1, t_2 dovolj blizu t_0 velja*

$$\phi(t_2, t_1, \phi(t_1, t_0, x)) = \phi(t_2, t_0, x).$$

Dokaz. Prva točka: Uporabimo izrek o inverzni preslikavi na $\phi(t, t_0, \cdot)$. Ker je $\phi(t_0, t_0, x) = x$, obstaja okolica t_0, v kateri je $\det D_x(t, t_0, x) \neq 0$, in dobimo difeomorfizem.

Druga točka sledi iz edinosti, ki nam jo da eksistenčni izrek. □

Vprašanje 23. Kaj je tok enačbe? Kakšne lastnosti ima?

1.7 Sistemi linearnih NDE

Naj bodo podane funkcije $a_{ij} : [a,b] \to \mathbb{R}$ in $b_k : [a,b] \to \mathbb{R}$, ki so na $[a,b]$ omejene. Sistem NDE prvega reda s koeficienti $a_{ij}(t)$ in desno stranjo $b_k(t)$ je sistem

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + \ldots + a_{1n}x_n \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + \ldots + a_{nn}x_n \end{aligned}$$

Naj bo matrika $A(t)$ podana s koeficienti a_{ij} in b vektor podan s komponentami b_k. Za $x = [x_1 \ldots x_n]^T$ sistem zapišemo kot $\dot{x} = Ax + b$. Če je $b = 0$ pravimo, da je sistem homogen.

Izrek. *Če je $A : [a,b] \to \mathbb{R}^{n\times n}$ zvezna in omejena, je množica rešitev homogenega sistema $\dot{x} = Ax$ n-dimenzionalen vektorski prostor v prostoru $\mathcal{C}^1([a,b])$.*

Dokaz. Naj bo R prostor rešitev. Najprej moramo pokazati, da je R vektorski podprostor. Za vsak par rešitev x_1, x_2 in $\alpha, \beta \in \mathbb{R}$ velja

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2 = \partial_t(\alpha x_1 + \beta x_2).$$

Naj bo $t_0 \in [a,b]$. Po eksistenčnem izreku obstaja rešitev vsakega začetnega problema $\dot{x} = Ax, x(t_0) = c \in \mathbb{R}^n$. Naj bo $e_1, \ldots, e_n$ kanonična baza v $\mathbb{R}^n$. Obstajajo torej rešitve x_i začetnih problemov $\dot{x} = Ax, x(t_0) = e_i$. Dokazati moramo še, da so x_i tudi globalne rešitve; ta del pustimo za kasneje, preostanek dokaza je lokalen.

Trdimo, da je $\{x_i\}_i$ baza R. Linearna neodvisnost je trivialna. Naj bo x rešitev sistema in $c = x(t_0)$. Vektor c razvijemo po bazi e_i v $c = \alpha_1 e_1 + \ldots + \alpha_n e_n$ in definiramo

$$\tilde{x} = \alpha_1 x_1 + \ldots + \alpha_n x_n.$$

Ker sta tako x kot $\tilde{x}$ rešitvi začetnega problema $\dot{x} = Ax, x(t_0) = c$, sta po eksistenčnem izreku enaki. □

Vprašanje 24. Kaj je množica rešitev homogenega sistema linearnih NDE? Dokaži.

Definicija. FUNDAMENTALNA MATRIKA sistema $\dot{x} = Ax$ je matrika

$$\phi(t,t_0) = \begin{bmatrix} x_{11}(t) & \ldots & x_{1n}(t) \\ \vdots & \ddots & \vdots \\ x_{n1}(t) & \ldots & x_{nn}(t) \end{bmatrix},$$

v kateri je i-ti stolpec enak i-ti rešitvi iz dokaza.

Za matriko $\phi(t,t_0)$ velja $\phi(t_0,t_0) = I$. Naj bo x rešitev začetnega problema $\dot{x} = Ax, x(t_0) = c$. Potem velja $x = \phi(t,t_0)c$.

Trditev. *Za vsak t veja* $\det \phi(t,t_0) \neq 0$.

Dokaz. Recimo, da obstaja t_1, da je $\det \phi(t_1,t_0) = 0$. Potem je matrika $\phi(t_1,t_0)$ singularna, zato ima netrivialno jedro, torej obstaja vsaj en neničeln vektor $d \in \mathbb{R}^n$, da $\phi(t_1,t_0)d = 0$. Torej

$$\phi(t_1,t_0)d = \sum_{i=1}^n x_i(t_1) d_i = 0.$$

Definirajmo

$$z(t) = \sum_{i=1}^n d_i . x_i(t).$$

Ta funkcija je rešitev sistema, zanjo velja $z(t_1) = 0$. Tudi funkcija $w(t) = 0$ je rešitev začetnega problema $\dot{x} = Ax, x(t_1) = 0$, torej po eksistenčnem izreku $z = 0$. Ker so v točki t_0 vektorji x_i linearno neodvisni, velja $d_i = 0$. □

Vprašanje 25. Kaj je fundamentalna matrika homogenega sistema linearnih NDE? Dokaži, da je nesingularna.

Oglejmo si preslikavo $\mathcal{F} : (t_0 - \varepsilon, t_0 + \varepsilon) \times \mathbb{R}^n \to \mathbb{R}^n$, definirano z

$$\mathcal{F}(t, y) = \phi(t, t_0)y.$$

To je tok enačbe $\dot{x} = Ax$. Za vsak dovolj majhen t je preslikava $y \mapsto \mathcal{F}(t, y)$ difeomorfizem, saj je obrnljiva linearna preslikava $\mathbb{R}^n \to \mathbb{R}^n$. Vzemimo $t_0 = 0$ in označimo $\phi(t, 0) = \phi(t)$.

Trditev. *Velja* $\phi(t_1 + t_2) = \phi(t_1)\phi(t_2)$.

Dokaz. Po eni strani imamo za vsak $x \in \mathbb{R}^n$

$$\begin{aligned} \phi(t_1)x &= \mathcal{F}(t_1, x), \\ \phi(t_2)\phi(t_1)x &= \mathcal{F}(t_2, \phi(t_1)x) = \mathcal{F}(t_2 + t_1, x), \end{aligned}$$

ker je $\mathcal{F}$ tok, po drugi strani pa

$$\phi(t_1 + t_2)x = \mathcal{F}(t_1 + t_2, x).$$

□

Vprašanje 26. Pokaži, da velja $\phi(t_1 + t_2) = \phi(t_1)\phi(t_2)$.

Trditev. *Splošna rešitev sistema* $\dot{x} = Ax + b$ *je afin podprostor v* $\mathcal{C}^1([a, b])$*, modeliran nad prostorom* R *rešitev homogenega sistema.*

To pomeni, da obstajajo vektorji $x_p \in \mathcal{C}^1([a, b])$, da je množica rešitev $\dot{x} = Ax + b$ enaka

$$W = \{x_h + x_p \mid x_h \in R\}.$$

Če imamo R in želimo poiskati W, potrebujemo eno partikularno rešitev nehomogenega sistema. To dobimo z variacijo konstante. Za vsak konstanten vektor $c \in \mathbb{R}^n$ je $\phi(t)c$ rešitev homogenega sistema. Poskusimo poiskati kakšno rešitev $\dot{x} = Ax + b$ z nastavkom $x_p = \phi(t)c(t)$, kjer je $c(t) : \mathbb{R} \to \mathbb{R}^n$ neznana funkcija. Začnemo z

$$\dot{x}_p = \dot{\phi}c + \phi\dot{c} = Ax_p + b,$$

iz česar dobimo $\phi\dot{c} = b$, oziroma $\dot{c} = \phi(-t)b(t)$. Sledi

$$c(t) = \int_0^t \phi(-\tau)b(\tau)d\tau$$

in

$$x_p(t) = \phi(t)c(t) = \int_0^t \phi(t - \tau)b(\tau)d\tau.$$

Dokazali smo

Trditev. *Splošna rešitev nehomogenega problema $\dot{x} = Ax + b$ je*

$$x(t,c) = \phi(t)c + \int_0^t \phi(t-\tau)b(\tau)d\tau,$$

kjer je c začetni pogoj pri $t_0 = 0$.

Vprašanje 27. Kako poiščeš množico rešitev $\dot{x} = Ax + b$?

Kako pa izračunamo ϕ? V zaključeni obliki za splošen A izračunati ne moremo, lahko pa dobimo izrazitev z Dysonovo vrsto. Oglejmo si začetni problem

$$\dot{\phi} = A\phi, \phi(0) = I.$$

Ta je ekvivalenten integralski enačbi

$$\phi(t) = I + \int_0^t A(\tau)\phi(\tau)d\tau.$$

To lahko razvijemo naprej v

$$\phi(t) = I + \int_0^t A(\tau_1)\left(I + \int_0^{\tau_1} A(\tau_2)\phi(\tau_2)\right)d\tau_1,$$

in nadaljujemo. Na koncu dobimo

$$\phi(t) = I + \sum_{n=1}^{\infty} \int_0^t A(\tau_1) \int_0^{\tau_1} A(\tau_2) \dots \int_0^{\tau_{n-1}} A(\tau_n) d\tau_n \dots d\tau_1$$

Trditev. *Naj bo matrična funkcija $A(t) : [0,T] \to \mathbb{R}^{n\times n}$ omejena po normi $\|A(t)\| \leq M$, Potem Dysonova vrsta konvergira.*

Dokaz. Ocenimo lahko

$$\|\phi\| \leq 1 + \sum_{n=1}^{\infty} \int_{\Delta_n(t)} \|A(\tau_1)\dots A(\tau_n)\|\, d\tau \leq 1 + \sum_{n=1}^{\infty} \int_{\Delta_n(t)} M^n d\tau,$$

kjer je $\Delta_n(t)$ urejeni n-simpleks. Nadalje velja

$$\|\phi\| \leq 1 + \sum_{n=1}^{\infty} V(\Delta_n(t))M^n = 1 + \sum_{n=1}^{\infty} \frac{t^n}{n!}M^n = e^{Mt} \leq e^{MT}.$$

□

Vprašanje 28. Pod katerim pogojem Dysonova vrsta konvergira? Dokaži.

Recimo, da je A konstanta matrika. V tem primeru se Dysonova matrika glasi

$$\phi(t) = e^{At}.$$

Vprašanje 29. Kakšna je Dysonova vrsta, če je matrika koeficientov konstanta?

Trditev. *Naj bo $A : [0,T] \to \mathbb{R}^{n\times n}$ matrična funkcija, za katero velja $A(t_1)A(t_2) = A(t_2)A(t_1)$ za vsaka $t_1, t_2 \in [0,T]$. Potem velja*

$$\phi(t) = \exp\left(\int_0^t A(\tau)d\tau\right).$$

Dokaz. Označimo $\Box_n(t) = [0,t]^n$. Oglejmo si

$$\int_{\Box_n(t)} A(\tau_1)\dots A(\tau_n)d\tau.$$

Za skoraj vsak $\tau \in \Box_n(t)$ obstaja natanko ena permutacija $\sigma \in S_n$, da velja $\sigma\cdot\tau \in \Delta_n(t)$. Označimo

$$\Delta^\sigma(t) = \{\tau \in \Box_n(t) \,|\, \sigma\cdot\tau \in \Delta_n(t)\}.$$

Razen na množici z mero 0 velja

$$\Box_n(t) = \bigcup_{\sigma\in S_n} \Delta^\sigma(t),$$

zato za vsako funkcijo $\mathcal{A} : \Box_n(t) \to \mathbb{R}^{n\times n}$ velja

$$\int_{\Box_n(t)} \mathcal{A}(\tau)d\tau = \sum_{\sigma\in S_n}\int_{\Delta^\sigma(t)} \mathcal{A}(\tau)d\tau = \sum_{\sigma\in S_n}\int_{\Delta_n(t)} \mathcal{A}(\tau_{\sigma(1)},\dots,\tau_{\sigma(n)})\underbrace{|\det\sigma|}_{=1}d\tau.$$

Če matrike komutirajo, torej velja

$$\int_{\Box_n(t)} A(\tau_1)\dots A(\tau_n)d\tau = n!\int_{\Delta_n(t)} A(\tau_1)\dots A(\tau_n)d\tau.$$

Torej je

$$\int_{\Delta_n(t)} A(\tau_1)\dots A(\tau_n)d\tau = \frac{1}{n!}\int_0^t A(\tau_1)d\tau_1\dots\int_0^t A(\tau_n)d\tau_n = \frac{1}{n!}\left(\int_0^t A(\tau)d\tau\right)^n.$$

□

Vprašanje 30. Kakšna je fundamentalna matrika, če $A(t_1)$ komutira z $A(t_2)$ za vsaka t_1, t_2? Dokaži.

Trditev (Liouvilloeva formula)**.** *Za fundamentalno matriko $\phi(t)$ sistema $\dot{x} = Ax + b$ velja*

$$\det\phi(t) = \exp\left(\int_0^t \mathrm{sl}(A(\tau))d\tau\right)$$

Dokaz. Naj bo

$$\phi(t) = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix}.$$

Če definicijo determinante odvajamo, dobimo

$$\partial_t \det \phi(t) = \sum_{\pi \in S_n} (-1)^{s(\pi)} \sum_{i=1}^{n} x_{1,\pi(1)} \ldots \dot{x}_{i,\pi(i)} \ldots x_{n,\pi(n)}.$$

Velja $\dot{\phi} = A\phi$, torej

$$\begin{bmatrix} \dot{x}_{11} & \cdots & \dot{x}_{1n} \\ \vdots & \ddots & \vdots \\ \dot{x}_{n1} & \cdots & \dot{x}_{nn} \end{bmatrix} = \begin{bmatrix} \sum_i a_{1i}x_{i1} & \cdots & \sum_i a_{1i}x_{in} \\ \vdots & \ddots & \vdots \\ \sum_i a_{ni}x_{i1} & \cdots & \sum_i a_{ni}x_{in} \end{bmatrix},$$

iz česar dobimo $\dot{x}_{ij} = \sum_k a_{ik}x_{kj}$. To vstavimo v prejšnji zapis

$$\partial_t \det \phi(t) = \sum_{\pi \in S_n} (-1)^{s(\pi)} \sum_{i=1}^{n} x_{1,\pi(1)} \ldots \left(\sum_{j=1}^{n} a_{ij} x_{j,\pi(j)} \right) \ldots x_{n,\pi(n)},$$

ki ga prvo seštejemo po π. Pri vsakem i dobimo determinanto

$$\begin{vmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ \sum_j a_{ij}x_{j1} & \cdots & \sum_j a_{ij}x_{jn} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{vmatrix} = \begin{vmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ a_{ii}x_{i1} & \cdots & a_{ii}x_{in} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{vmatrix} = a_{ii} \det \phi$$

Torej

$$\partial_t \det \phi(t) = a_{11} \det \phi + a_{22} \det \phi + \ldots + a_{nn} \det \phi = \operatorname{sl} A \det \phi.$$

To je diferencialna enačba, katere rešitev je

$$\det \phi = \exp \left(\int_0^t \operatorname{sl} A d\tau \right).$$

□

Vprašanje 31. Povej in dokaži Liouvilloevo formulo.

1.8 Linearne NDE višjega reda

Obravnavamo enačbe oblike

$$a_n(t)x^{(n)} + \ldots + a_1(t)\dot{x} + a_0(t)x = b(t).$$

Definicija. Linearni diferencialni operator s koeficienti $a_i(t)$ je preslikava

$$L : \mathcal{C}^1([a,b]) \to \mathcal{C}([a,b]),$$

podana s predpisom

$$Lx(t) = a_n(t)x^{(n)} + \ldots + a_0(t)x.$$

Splošna rešitev homogene enačbe $Lx = 0$ je $\ker L$. Vemo, da je splošna rešitev n-dimenzionalni vektorski prostor. Enačbo s substitucijo

$$\begin{aligned} x_1 &= x \\ x_2 &= \dot{x} \\ &\vdots \\ x_n &= x^{(n-1)} \end{aligned}$$

prepišemo v sistem

$$\begin{aligned} \dot{x}_0 &= x_1 \\ \dot{x}_1 &= x_2 \\ &\vdots \\ \dot{x}_n &= \frac{1}{a_n}(b - a_0x_1 - a_2x_1 - \ldots - a_{n-1}x_{n-2}) \end{aligned}$$

Označimo

$$p_i(t) = \frac{a_i(t)}{a_n(t)},$$

s čimer izrazimo matriko koeficientov zgornjega sistema

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-1} \end{bmatrix}.$$

Izrek. *Naj bo $L : \mathcal{C}^n([a,b]) \to \mathcal{C}^0([a,b])$ regularen diferencialni operator. Za vsak začetni pogoj $x(t_0) = c_0, \dot{x}(t_0) = c_1, \ldots, x^{(n-1)}(t_0) = c_{n-1}$ ima enačba $Lx = b$ natanko eno rešitev na vsem intervalu $[a,b]$.*

Opomba. Diferencialni operator: $Lx = a_n x^{(n)} + \ldots + a_1\dot{x} + a_0 x = b$ je regularen, če je $a_n(t) \neq 0$ za vsak t in če so $a_i(t)$ omejene.

Množica rešitev homogene linearne enačbe $Lx = 0$ je n-dimenzionalen vektorski prostor v $\mathcal{C}^n([a,b])$. Vsaka rešitev $x(t)$ namreč na enoličen način določa rešitev sistema $\dot{\vec{x}} = A\vec{x}$ za

$$\dot{\vec{x}} = \begin{bmatrix} x(t) \\ \dot{x}(t) \\ \vdots \\ x^{(n-1)}(t) \end{bmatrix}.$$

Tudi obratno je res: vsak vektor $\vec{x}$ na enoličen način določa svojo prvo komponento.

Oglejmo si fundamentalno matriko

$$\phi(t) = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ \dot{x}_1 & \dot{x}_2 & \cdots & \dot{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(n-1)} & x_2^{(n-1)} & \cdots & x_n^{(n-1)} \end{bmatrix}.$$

Denimo, da so funkcije $x_1(t), x_2(t), \ldots, x_n(t)$ baza rešitev enačbe $Lx = 0$. Za bazo velikokrat vzamemo take vektorje $\vec{x}_1, \ldots, \vec{x}_n$, da je $\phi(t = 0) = I$. Če je ta baza dobljena iz baze rešitev enačbe $Lx = 0$, potem za te rešitve velja $x_i(0) = 0, \ldots, x_i^{(i-1)}(t) = 0, x_i^{(i)}(0) = 1, x_i^{(i+1)}(0) = 0, \ldots, x_i^{(n-1)}(0) = 0$. Determinanta

$$W(t) = \det \phi(t)$$

se imenuje DETERMINANTA WRONSKEGA. V tem primeru se Liouvilleova formula glasi

$$W(t) = W(t_0) \exp\left(-\int_{t_0}^{t} p_{n-1}(\tau) d\tau \right).$$

Rešitev nehomogene enačbe dobimo s pomočjo variacije konstante;

$$\vec{x} = \int_{t_0}^{t} \phi(t)\phi^{-1}(\tau)\vec{b}(\tau) d\tau,$$

oziroma, ker ima $\vec{b}$ v tem primeru le eno neničelno komponento $b(t)$, bo prva komponenta $\vec{x}$ enaka

$$x(t) = \int_{t_0}^{t} \sum_{i=1}^{n} \phi_{1i}(t)\phi_{in}^{-1}(\tau)b(\tau)d\tau = \sum_{i=1}^{n} x_i(t) \int_{t_0}^{t} \phi_{in}^{-1}(\tau)b(\tau)d\tau,$$

kjer je $(x_i(t))_i$ baza rešitev homogene enačbe $Lx = 0$.

Vprašanje 32. Kako izračunaš rešitev linearne NDE višjega reda?

1.8.1 Enačbe s konstantnimi koeficienti

Naj bo sedaj linearen diferencialni operator L podan z

$$Lx = x^{(n)} + a_1 x^{(n-1)} + \ldots + a_n x.$$

Oglejmo si enačbo $Lx = 0$. Če vstavimo nastavek $x(t) = e^{\lambda t}$:

$$\lambda^n e^{\lambda t} + a_1 \lambda^{n-1} e^{\lambda t} + \ldots + a_{n-1} \lambda e^{\lambda t} + a_n e^{\lambda t} = 0$$

oziroma (ker $e^{\lambda t} \neq 0$ tudi za $\lambda \in \mathbb{C}$)

$$\lambda^n + a_1 \lambda^{n-1} + \ldots + a_{n-1} \lambda + a_n = 0.$$

Ta polinom imenujemo KARAKTERISTIČNI POLINOM ENAČBE $Lx = 0$ in označimo s $P(\lambda)$.

Trditev. *Če so $\lambda_1, \ldots, \lambda_n$ različne ničle karakterističnega polinoma $P(\lambda)$, potem so funkcije $x_i(t) = e^{\lambda_i t}$ baza rešitev homogene enačbe $Lx = 0$.*

Dokaz. Vemo, da je rešitev sistema vektorski prostor, dokazati moramo samo, da so te rešitve linearno neodvisne. Priredimo našim rešitvam pripadajoče rešitve sistema, ki je prirejen $Lx = 0$,

$$x_i(t) \mapsto \vec{x}_i(t) = \begin{bmatrix} x_i(t) \\ \dot{x}_i(t) \\ \vdots \\ x_i^{n-1}(t) \end{bmatrix}.$$

Te stolpce zložimo v matriko in dobimo kandidatko za fundamentalno matriko $\phi(t)$

$$\phi(t) = \begin{bmatrix} e^{\lambda_1 t} & e^{\lambda_2 t} & \cdots & e^{\lambda_n t} \\ \lambda_1 e^{\lambda_1 t} & \lambda_2 e^{\lambda_2 t} & \cdots & \lambda_n e^{\lambda_n t} \\ \lambda_1^2 e^{\lambda_1 t} & \lambda_2^2 e^{\lambda_2 t} & \cdots & \lambda_n^2 e^{\lambda_n t} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} e^{\lambda_1 t} & \lambda_2^{n-1} e^{\lambda_2 t} & \cdots & \lambda_n^{n-1} e^{\lambda_n t} \end{bmatrix}$$

Matrika $\phi(0)$ je vandermondova, torej

$$W(t) = \det \phi(t) = W(0) \exp\left(\int_0^t \mathrm{sl}(A) d\tau \right) = \prod_{i>j} (\lambda_i - \lambda_j) \cdot e^{-a_1 t}.$$

Ker so vsi λ_i različni, velja $W(0) \neq 0$, torej je $W(t) \neq 0$ za vsak t, in so funkcije x_i res linearno neodvisne in so baza prostora rešitev enačbe $Lx = 0$. □

Vprašanje 33. Kaj je karakteristični polinom homogene linearne NDE višjega reda? Kako z njim poiščemo rešitve enačbe, če so vse ničle različne? Dokaži.

Z razvojem po prvem stolpcu lahko izračunamo

$$\det(A-\lambda I) = \begin{vmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\lambda & 1 & \cdots & 0 \\ 0 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1-\lambda \end{vmatrix} = \lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n.$$

Lastne vrednosti matrike A so torej res ničle karakterističnega polinoma $P(\lambda)$ enačbe $Lx = 0$.

Kaj pa če ima A večkratne lastne vrednosti?

Trditev. *Naj bo λ k-kratna ničla polinoma $P(\lambda)$. Potem so funkcije $x_0(t) = e^{\lambda t}, x_1(t) = te^{\lambda t}, \ldots, x_{k-1}(t) = \lambda^{k-1}e^{\lambda t}$ linearno neodvisne rešitve $Lx = 0$.*

Dokaz. Opazimo, da velja $x_i(t) = \frac{\partial^i}{\partial\lambda^i}e^{\lambda t}$ za $i = 0, \ldots, k-1$. Kot že vemo, velja $Lx_0(t) = P(\lambda)e^{\lambda t}$. Odvedemo to enačbo i-krat po λ. Na levi dobimo

$$\frac{\partial^i}{\partial\lambda^i}Lx_0(t) = L(\frac{\partial^i}{\partial\lambda^i}e^{\lambda t}) = L(x_i(t)).$$

To je res, ker so vsi koeficienti a_i konstantni na t in λ. Imamo torej

$$\begin{aligned} Lx_i(t) = \frac{\partial^i}{\partial\lambda^i}Lx_0(t) &= \frac{\partial^i}{\partial\lambda^i}(P(\lambda)e^{\lambda t}) \\ &= \sum_{l=0}^{i}\binom{i}{l}P^{(l)}(\lambda)\frac{\partial^{i-l}}{\partial\lambda^{i-l}}e^{\lambda t} \\ &= \sum_{l=0}\binom{i}{l}P^{(l)}(\lambda)t^{i-l}e^{\lambda t} \\ &= \sum_{l=0}^{i}\binom{i}{l}P^{(l)}(\lambda)x_{i-l}(t). \end{aligned}$$

Naj bo sedaj λ ničla k-te stopnje in $i \leq k$. Potem velja $P^{(l)}(\lambda) = 0$, torej $Lx_i(t) = 0$, funkcija $x_i(t) = t^ie^{\lambda t}$ je torej res rešitev enačbe $Lx = 0$ za $i = 0, \ldots, k-1$. □

Vprašanje 34. Kako s karakterističnim polinomom poiščemo rešitve linearne NDE višjega reda s konstantnimi koeficienti, če niso vse ničle različne? Dokaži.

Naj bo sedaj λ kompleksna ničla $\lambda = a + ib$. Če so koeficienti operatorja L realni, potem je tudi $\overline{\lambda}$ ničla P. Če je $\lambda = a + ib$ ničla k-tega reda, je tudi $\overline{\lambda}$ ničla k-tega reda. Ti dve lastni vrednosti dasta $2k$ baznih rešitev. Če jih želimo na najpreprostejši način izraziti z realnimi funkcijami, dobimo bazo

$$
\begin{gathered}
x_0(t) = e^{ta}\cos(bt), x_1(t) = e^{ta}\sin(bt), \\
x_2(t) = te^{ta}\cos(bt), x_3(t) = te^{ta}\sin(bt), \\
\vdots \\
x_{2k-2} = t^{k-1}e^{ta}\cos(bt), x_{2k-1} = t^{k-1}e^{ta}\sin(bt).
\end{gathered}
$$

1.8.2 Linearizacija

Imejmo nelinearen sistem NDE $\dot{\vec{x}} = F(t,x)$. Naj bo $\vec{x}_0(t)$ neka rešitev tega sistema. Definiramo nelinearen operator $\mathcal{F}(\vec{x}(t)) = \dot{\vec{x}}(t) - F(t,\vec{x})$. Funkcija $\vec{x}_0$ je rešitev sistema natanko tedaj, ko je $\mathcal{F}(\vec{x}_0) = 0$. Splošna rešitev $\mathcal{S}$, ki je n-parametrični nelinearen prostor, je nivojska ploskev $\mathcal{S} = \mathcal{F}^{-1}(0)$. Naj bo sedaj $s \mapsto \vec{x}(t,s) \in \mathcal{S}$ pot v prostoru rešitev, za katero velja $\vec{x}(t,0) = \vec{x}_0(t)$. Oglejmo si odvod

$$
\left.\frac{d}{ds}\right|_{s=0} \mathcal{F}(\vec{x}(t,s)) = D_{\vec{x}_0}\mathcal{F}\left(\left.\frac{d}{ds}\right|_{s=0} \vec{x}(t,s)\right) =: D_{\vec{x}_0}\mathcal{F}(\vec{u}(t)).
$$

Če je $\vec{x}(t,s)$ rešitev za vsak s, potem velja $F(\vec{x}(t,s)) = 0$, zato

$$
\left.\frac{d}{ds}\right|_{s=0} \mathcal{F}(\vec{x}(t,s)) = D_{\vec{x}_0}\mathcal{F}(\vec{u}(t)) = 0.
$$

Operator $D_{\vec{x}_0}\mathcal{F}$ je linearen. Če v predpis odvoda vstavimo definicijo $\mathcal{F}$, dobimo

$$
\left.\frac{d}{ds}\right|_{s=0} \mathcal{F}(\vec{x}(t,s)) = \left.\frac{d}{ds}\right|_{s=0} \Big(\dot{\vec{x}}(t,s) - F(t,\vec{x}(t,s))\Big) = \dot{\vec{u}} - \left.\frac{d}{ds}\right|_{s=0} F(t,\vec{x}(t,s)).
$$

Če desni člen posredno odvajamo, dobimo sistem $\dot{\vec{u}} - A(t)\vec{u} = 0$ za

$$
A(t) = \begin{bmatrix} \partial_{x_1}F_1(t,\vec{x}_0(t)) & \cdots & \partial_{x_n}F_1(t,\vec{x}_0(t)) \\ \vdots & \ddots & \vdots \\ \partial_{x_1}F_n(t,\vec{x}_0(t)) & \cdots & \partial_{x_n}F_n(t,\vec{x}_0(t)) \end{bmatrix}.
$$

Dobili smo linearizacijo začetnega sistema okoli rešitve $\vec{x}_0$. Splošna rešitev te linearizacije je jedro operatorja $D_{\vec{x}_0}\mathcal{F}$, oziroma tangentni prostor na $\mathcal{S}$ v točki $\vec{x}_0$.

Vprašanje 35. Izpelji linearizacijo nelinearnega sistema NDE. Kakšne so rešitve linearizacije?

1.9 Variacijski račun

Definicija. Naj bosta U in V Banachova prostora in $P : U \to V$ operator. Naj bo $u \in U$ točka v prostoru. Gateauxov odvod P v u in v smeri $v \in U$ je podan s predpisom

$$
D_u^G P(v) = \left.\frac{d}{dt}\right|_{t=0} P(u+tv) = \lim_{t\to 0} \frac{P(u+tv) - P(u)}{t}.
$$

Ta odvod se imenuje tudi SMERNI ali ŠIBKI odvod.

Definicija. Naj bosta U in V Banachova prostora in $P : U \to V$ operator. FRECHETOV ali KREPKI odvod P v točki u je omejen linearen operator $\mathcal{A} : U \to V$, za katerega velja $P(u + tv) = P(u) + \mathcal{A}(v) + o(u, v)$ in

$$\lim_{\|v\|\to 0} \frac{\|o(u, v)\|}{\|v\|} = 0.$$

Pišemo $\mathcal{A} = D_u^F P$.

Vprašanje 36. Definiraj šibki in krepki odvod. Kako se še imenujeta?

Trditev. *Če je P v $u \in U$ Frechetovo odvedljiv, je tam tudi Gateauxovo odvedljiv. Tedaj sta odvoda enaka.*

Dokaz. Za vsak $v \in U$ velja

$$P(u + tv) = Pu + D_u^F P(tv) + o(u, tv) = Pu + tD_u^F P(v) + o(u, tv),$$

iz česar pride

$$\frac{P(u + tv) - Pu}{t} = D_u^F P(v) + \frac{o(u, tv)}{t}.$$

Brez škode za splošnost predpostavimo $\|v\| = 1$. Ker Frechetov odvod obstaja, velja

$$\lim_{t\to 0} \frac{P(u + tv) - Pu}{t} = D_u^F P(v) + \lim_{\|tv\|\to 0} \frac{o(u, tv)}{\|tv\|} = D_u^F P(v).$$

$\square$

Trditev. *Če obstaja $D_u^G P$ in je limita*

$$\lim_{t\to 0} \frac{P(u + tv) - Pu}{t} = D_u^G P(v)$$

enakomerna glede na v na enotski sferi $S \subseteq U$, potem obstaja tudi $D_u^F P$.

Dokaz. Enakomernost pomeni, da za vsak $\varepsilon > 0$ obstaja $\delta > 0$, da $|t| < \delta$ implicira

$$\left\| \frac{P(u + tv) - Pu}{t} - D_u^G P(v) \right\| < \varepsilon$$

ne glede na $v \in S \subseteq U$, oziroma

$$\left\| P(u + tv) - Pu - D_u^G P(tv) \right\| < \varepsilon t.$$

Označimo $h = tv$ in dobimo

$$\left\| P(u + h) - Pu - D_u^G P(h) \right\| < \varepsilon \|h\|.$$

Po definiciji limite potem velja

$$\lim_{\|h\|\to 0} \frac{\left\|P(u+h) - Pu - D_u^G P(h)\right\|}{\|h\|} = 0.$$

□

Vprašanje 37. Povej zadostni pogoj za obstoj Frechetovega odvoda.

Trditev. *Naj bo $u \in U$ lokalni minimum funkcionala $\mathcal{L} : U \to \mathbb{R}$, in naj bo $\mathcal{L}$ krepko odvedljiv. Potem velja $D_u\mathcal{L}(v) = 0$ za vsak v.*

Dokaz. Ker je v u dosežen minimum, velja za vsak v in vsak dovolj majhen t $\mathcal{L}(u+tv) > \mathcal{L}(u)$. Ker je $\mathcal{L}(u+tv) = \mathcal{L}(u) + D_u\mathcal{L}(tv) + o(tv)$, velja $D_u\mathcal{L}(tv) + o(tv) > 0$. Če je $t > 0$, je $tD_u\mathcal{L}(v) + o(tv) > 0$ in

$$D_u\mathcal{L}(v) + \frac{o(tv)}{t} > 0.$$

Ker drug člen limitira k 0, je za dovolj majhen t predznak izraza odvisen le od predznaka $D_u\mathcal{L}(v)$, ki mora torej biti pozitiven. Če enako naredimo za $t < 0$, dobimo, da mora biti predznak $D_u\mathcal{L}(v)$ negativen; torej je enak nič. □

Vprašanje 38. Dokaži, da je odvod funkcionala v lokalnem minimumu enak 0.

Če je

$$\mathcal{L} = \int_a^b L(x, u, u')dx,$$

je odvod funkcionala enak

$$D_u\mathcal{L}(v) = \int_a^b \left(\frac{\partial L}{\partial u}(x)v(x) + \frac{\partial L}{\partial u'}(x)v'(x)\right) dx.$$

Definicija. Naj bo U Banachov prostor funkcij $y : [a, b] \to \mathbb{R}$ z metriko, porojeno iz maksimum norme. Naj bosta $A, B \in \mathbb{R}$ konstanti. Prostor

$$V = \{y \in U \mid y(a) = A, y(b) = B\}.$$

imenujemo PROSTOR DOPUSTNIH FUNKCIJ.

Nekoliko splošneje: če so $l_i : U \to \mathbb{R}$ omejeni linearni funkcionali, je prostor dopustnih funkcij podan z

$$V = \{y \in U \mid \forall i \,.\, l_i(y) = A_i\}.$$

Definicija. DOPUSTNA VARIACIJA je vsaka funkcija $v \in U$, za katero velja $u + tv \in V$ za vsak t.

Za dopustne variacije velja $l_i(u+tv) = l_i(u) + tl_i(v)$. Če je $u+tv \in V$, je $l_i(u+tv) = A_i$, in torej $l_i(v) = 0$ za vsak i. Prostor dopustnih variacij je torej podan s predpisom

$$\mathrm{Var} = \{v \in U \mid l_i(v) = 0 \,\forall i\} = \bigcap_i \ker l_i.$$

Prostor Var je torej linearen podprostor v U, prostor V pa je afin podprostor v U, modeliran s prostorom Var. Če se vrnemo k osnovnemu variacijskemu problemu, velja $l_1(u) = u(a)$ in $l_2(u) = u(b)$.

Definicija. TESTNE FUNKCIJE na intervalu $[a, b]$ so funkcije $\varphi : [a, b] \to \mathbb{R}$, za katere velja

- $\varphi \in \mathcal{C}^\infty$,
- $\overline{\operatorname{supp} \varphi} \subsetneq [a, b]$.

Vprašanje 39. Kaj so dopustne funkcije, dopustne variacije in testne funkcije? Kaj mora veljati za dopustne variacije?

Izrek (Osnovni izrek variacijskega računa)**.** *Naj bo $f : [a, b] \to \mathbb{R}$ zvezna funkcija. Če za vsako testno funkcijo φ velja*

$$\int_a^b f(x)\varphi(x) = 0,$$

potem je $f(x) = 0$ na $[a, b]$.

Dokaz. Denimo, da ni tako, da obstaja $x_0 \in (a, b)$, za katerega je $f(x_0) = c > 0$. Zaradi zveznosti f obstaja $\delta > 0$, da velja $f(x) > c/2$ za $x \in (x_0 - \delta, x_0 + \delta)$. Naj bo φ testna funkcija, za katero je $\operatorname{supp} \varphi \subset (x_0 - \delta, x_0 + \delta)$ in $\varphi(x) > 0$ na $\operatorname{supp} \varphi$. Potem imamo

$$\int_a^b f(x)\varphi(x)dx = \int_{x_0-\delta}^{x_0+\delta} f(x)\varphi(x)dx > \frac{c}{2}\int_{x_0-\delta}^{x_0+\delta} \varphi(x)dx > 0,$$

kar je protislovno. □

Vprašanje 40. Povej in dokaži osnovni izrek variacijskega računa.

Predpostavimo, da je funkcija $\partial_{u'} L(x, u, u')$ odvedljiva po x, in z integracijo po delih izračunamo

$$\int_a^b \partial_{u'} L v' dx = -\int_a^b \frac{d}{dx}\left(\partial_{u'} L\right) v dx + \partial_{u'} L v\big|_a^b.$$

Odvod je torej enak

$$D_u\mathcal{L}(v) = \int_a^b \left(\partial_u L - \frac{d}{dx}\partial_{u'} L\right) v dx + \partial_{u'} L v\big|_a^b = 0.$$

Ker so testne funkcije tudi dopustne variacije, dobimo, da za vse testne funkcije v velja

$$D_u\mathcal{L}(v) = \int_a^b \left(\partial_u L - \frac{d}{dx}\partial_{u'} L\right) v dx = 0,$$

oziroma, po osnovnem izreku variacijskega računa,

$$\partial_u L(x, \hat{u}(x), \hat{u}'(x)) - \frac{d}{dx}\partial_{u'} L(x, \hat{u}(x), \hat{u}'(x)) = 0.$$

Temu pravimo Euler-Lagrangeova enačba.

Vprašanje 41. Izpelji Euler-Lagrangeovo enačbo.

Če imamo samo en robni pogoj $u(a) = A$, bo za testne funkcije še vedno veljala Euler-Lagrangeova enačba, za ostale dopustne variacije pa dobimo novo enačbo

$$\partial_{u'} L(b) = 0.$$

Temu pravimo DINAMIČNI POGOJ.

1.9.1 Vezani ekstremi

Naj bodo $\phi_i : V \to \mathbb{R}$ nelinearni funkcionali. Množica $W \subseteq V$ naj bo podana s predpisom

$$W = \{u \in V \mid \forall i . \phi(u) = l_i\}.$$

Označimo $\vec{\phi} : V \to \mathbb{R}^n$ s komponentami $\vec{\phi} = (\phi_1, \ldots, \phi_n)$ in $\vec{l} = (l_1, \ldots, l_n)$. Velja $W = \vec{\phi}^{-1}(\vec{l})$. Predpostavljamo, da so v vseh točkah $u \in W$, ki nas bodo zanimale, funkcionali $\{D_u\phi_i\}_i$ linearno neodvisni v dualnem prostoru Var*. Ker je za pot $v \in V$ z $v(0) = u$ odvod $\dot{v}(0)$ v tangentnem prostoru $T_uV = \text{Var}$, in ker velja

$$\left.\frac{d}{dt}\right|_{t=0} \phi_i(v(t)) = D_u\phi_i(\dot{v}(0)),$$

so $D_u\phi_i$ res v Var*.

Naj bo $\mathcal{L} : V \to \mathbb{R}$ funkcional. Iščemo njegove ekstreme na podmnožici $W \subseteq V$. Videli smo, da je $\hat{u} \in W$ stacionarna točka $\mathcal{L} : W \to \mathbb{R}$, če za vsako krivuljo $v : (-\varepsilon, \varepsilon) \to W$, za katero je $v(0) = \hat{u}$, velja

$$D_u\mathcal{L}(\dot{v}(0)) = 0.$$

En način iskanja bi bil, da za vsak $u \in W$ poiščemo T_uW in najdemo tisti $\hat{u}$, za katerega je $D_{\hat{u}}\mathcal{L} = 0$ na prostoru $T_{\hat{u}}W$. Tangentni prostor je enak

$$T_uW = \{\dot{v}(0) \mid \forall i \, . \, D_u\phi_i(\dot{v}(0)) = 0\} = \bigcap_{i=1}^n \ker D_u\phi_i.$$

Ker sta tako odvod kot tangentni prostor v vsaki točki različna, je ta način iskanja nepraktičen.

Vprašanje 42. Povej in razloži primitiven način iskanja vezanih ekstremov.

Boljši način je, da $\mathcal{L}$ modificiramo tako, da bo za novi $\tilde{\mathcal{L}}$ veljal sklep, da če je $D_{\hat{u}}\tilde{\mathcal{L}}(v) = 0$ za vse dopustne variacije v, bo $\hat{u}$ stacionarna točka $\mathcal{L}$. Vzemimo poljubno dopustno variacijo v in jo dekomponirajmo v obliko $v = v_u + v_u^\perp$, kjer je $v_u \in T_uW$, $v_u^\perp$ pa pravokoten nanj. Naj bo $\{\varphi_1(u), \ldots, \varphi_n(u)\}$ dualna baza baze $\{D_u\phi_i\}_i$, torej $D_u\phi_i \cdot \varphi_j(u) = \delta_{ij}$. Definiramo

$$v_u^\perp = \sum_{i=1}^n D_u\phi_i(v) \cdot \varphi_i(u)$$

in trdimo, da velja $v - v_u^\perp \in T_uW$. To je res, ker je

$$\begin{aligned} D_u\phi_j(v - v_u^\perp) &= D_u\phi_j(v) - D_u\phi_j\left(\sum_{i=1}^n D_u\phi_i(v) \cdot \varphi_i(u)\right) \\ &= D_u\phi_j(v) - \left(\sum_{i=1}^n D_u\phi_i(v) \cdot D_u\phi_j(\varphi_i(u))\right) \\ &= D_u\phi_j(v) - D_u\phi_j(v) \\ &= 0. \end{aligned}$$

Če definiramo

$$\tilde{\mathcal{L}}(u) = \mathcal{L}(u) - \sum_{i=1}^n D_u\mathcal{L}(\varphi_i(u))\phi_i(u),$$

bo veljalo $D_u\tilde{\mathcal{L}}(v_u^\perp) = 0$, kar lahko preverimo s podobnim računom. Označimo $\lambda_i(u) = D_u\mathcal{L}(\varphi_i(u))$. Odvajanje nam da

$$D_u\tilde{\mathcal{L}} = D_u\mathcal{L} - \sum_{i=1}^n \lambda_i(u) D_u\phi_i - \sum_{i=1}^n D_u\lambda_i \cdot \phi_i(u).$$

Če ustrezno modificiramo λ_i, bo res veljalo $D_{\hat{u}}\lambda_i = 0$ v ekstremnih točkah, vendar to vodi v zelo kompliciran predpis. Alternativno lahko rečemo, da so λ_i konstantne.

Vprašanje 43. Izpelji drug način reševanja problemov vezanih ekstremov.

Strategija reševanja variacijskih problemov z vezmi je tedaj takšna: Prvo rešimo Euler-Lagrangeovo enačbo za

$$\tilde{\mathcal{L}} = \mathcal{L} - \sum_{i=1}^n \lambda_i\phi_i,$$

kjer dobimo rešitev $\hat{u}$, odvisno od λ_i. Parametre poiščemo s pomočjo pogojev $\phi_i(\hat{u}) = l_i$.

2 Mehanika

2.1 Osnove Newtonove mehanike

Definicija. AFIN PROSTOR $\mathcal{A}$ nad vektorskim prostorom V je množica z binarno operacijo $+ : \mathcal{A} \times V \to \mathcal{A}$, za katero velja:

- Za poljuben $\mathbf{A} \in \mathcal{A}$ ter $a, b \in V$ velja $(\mathbf{A} + a) + b = \mathbf{A} + (a + b)$
- Za poljubna $\mathbf{A}, \mathbf{B} \in \mathcal{A}$ obstaja natanko določen $a \in V$, da je $\mathbf{B} = \mathbf{A} + a$.

DIMENZIJA afinega prostora je enaka dimenziji vektorskega prostora V.

Definicija. Naj bo $\mathcal{A}$ afin prostor nad vektorskih prostorom V. Definiramo operacijo odštevanja $\mathcal{A} \times \mathcal{A} \to V$ s predpisom

$$\mathbf{B} - \mathbf{A} = a \Leftrightarrow \mathbf{B} = \mathbf{A} + a.$$

Trditev. *V afinem prostoru veljajo naslednje zveze:*

- $\mathbf{A} - \mathbf{A} = 0.$
- $(\mathbf{A} - \mathbf{B}) + (\mathbf{B} - \mathbf{A}) = 0.$
- $(\mathbf{A} - \mathbf{B}) + (\mathbf{B} - \mathbf{C}) + (\mathbf{C} - \mathbf{A}) = 0.$
- $(\mathbf{A} - \mathbf{B}) + a = (\mathbf{A} + a) - \mathbf{B}.$
- $(\mathbf{A} - \mathbf{B}) + \mathbf{C} = (\mathbf{C} - \mathbf{B}) + \mathbf{A}.$

Definicija. Preslikava $g : \mathcal{A} \to \mathcal{A}'$ med afinima prostoroma je AFINA, če obstaja $dg \in L(V, V')$, da za vsaka $\mathbf{A}, \mathbf{B} \in \mathcal{A}$ velja $g(\mathbf{A}) - g(\mathbf{B}) = dg(\mathbf{A} - \mathbf{B})$.

Za afino preslikavo g si lahko izberemo POL $\mathbf{O}$, ter izpeljemo

$$g(\mathbf{A}) = g(\mathbf{O}) + dg(\mathbf{A} - \mathbf{O}).$$

Vrednosti funkcije seveda niso odvisne od izbire pola.

Vprašanje 1. Definiraj afin prostor in afino preslikavo.

Definicija. GALILEJEVA STRUKTURA je trojica $\mathcal{G} = (\mathcal{A}, \mathfrak{t}, \rho)$, kjer je $\mathcal{A}$ štirirazsežni afin prostor nad V, $\mathfrak{t} \in L(V, \mathbb{R})$ in ρ ekvlidska metrika na $\ker \mathfrak{t}$, porojena z normo $\|\cdot\|$. Funkciji $\mathfrak{t}$ pravimo ČASOVNOST, elementom $\mathcal{A}$ pa pravimo DOGODKI. Pretečeni čas med dogodkoma $\mathbf{A}$ in $\mathbf{B}$ označimo s $\mathfrak{t}(\mathbf{A}, \mathbf{B})$. Dogodka sta ISTOČASNA, če je $\mathfrak{t}(\mathbf{A}, \mathbf{B}) = 0$. Za istočasne dogodke lahko definiramo razdaljo $\rho(\mathbf{A}, \mathbf{B}) = \|\mathbf{B} - \mathbf{A}\|$ (uporabimo isto oznako kot za metriko v $\ker \mathfrak{t}$).

Definicija. Galilejevi strukturi $\mathcal{G} = (\mathcal{A}, \mathfrak{t}, \rho)$ in $\mathcal{G}' = (\mathcal{A}', \mathfrak{t}', \rho')$ sta EKVIVALENTNI, če obstaja afina bijekcija $g : \mathcal{A} \to \mathcal{A}'$, ki ohranja časovnost in razdaljo med istočasnimi dogodki;

$$\mathfrak{t}'(g(\mathbf{A}) - g(\mathbf{B})) = \mathfrak{t}(\mathbf{A}, \mathbf{B}), \qquad \rho'(g(\mathbf{A}), g(\mathbf{B})) = \rho(\mathbf{A}, \mathbf{B}).$$

Taki transformaciji pravimo GALILEJEVA TRANSFORMACIJA.

Vprašanje 2. Definiraj Galilejevo strukturo in Galilejeve transformacije.

Modelni primer je naravna Galilejeva struktura na $\mathcal{A} = \mathbb{R} \times \mathbb{E}$, kjer je $\mathbb{E}$ trirazsežni Evklidski prostor. Za elemente $A_i = (t_i, \mathbf{P}_i) \in \mathcal{A}$ naravne strukture velja

- $\mathfrak{t}(A_1 - A_2) = t_1 - t_2$,
- $\rho(A_1, A_2) = \|\mathbf{P}_1 - \mathbf{P}_2\|$.

Definicija. KOORDINATNI SISTEM na $\mathcal{A}$ je bijekcija $\phi : \mathcal{A} \to \mathbb{R} \times \mathbb{E}$ s komponentami $\phi(A) = (\tau\phi(A), \pi\phi(A))$, in pri kateri je $\tau \circ \phi$ linearna preslikava.

Opomba. Če sta ϕ in ϕ' koordinatna sistema, je preslikava $\phi' \circ \phi^{-1} : \mathbb{R} \times \mathbb{E} \to \mathbb{R} \times \mathbb{E}$ bijekcija.

Vprašanje 3. Kaj je koordinantni sistem?

Izrek. *Galilejeva transformacija $g : \mathbb{R} \times \mathbb{E} \to \mathbb{R} \times \mathbb{E}$ je oblike*

$$g(t, \mathbf{P}) = (t_0' + t, \mathbf{P}_0' + \vec{c}t + Q(\mathbf{P} - \mathbf{P}_0)),$$

kjer je $Q \in O(3)$ ortogonalna transformacija.

Dokaz. Ker je g afina preslikava, jo lahko zapišemo kot

$$g(t, \mathbf{P}) = g(t_0, \mathbf{P}_0) + dg(t - t_0, \mathbf{P} - \mathbf{P}_0),$$

kjer je $dg \in L(\mathbb{R}^4, \mathbb{R}^4)$. Če označimo $g(t_0, \mathbf{P}_0) = (t_0', \mathbf{P}_0')$, in zapišemo dg kot bločno matriko, dobimo

$$g(t, \mathbf{P}) = (t_0', \mathbf{P}_0') + \begin{bmatrix} \alpha & \vec{a}^T \\ \vec{c} & Q \end{bmatrix} \begin{bmatrix} t - t_0 \\ \mathbf{P} - \mathbf{P}_0 \end{bmatrix} = (t_0', \mathbf{P}_0') + \begin{bmatrix} \alpha(t - t_0) + \vec{a} \cdot (\mathbf{P} - \mathbf{P}_0) \\ (t - t_0).\vec{c} + Q(\mathbf{P} - \mathbf{P}_0) \end{bmatrix}.$$

Za dogodka $(t_1, \mathbf{P}_1)$ in $(t_2, \mathbf{P}_2)$ zahtevamo

$$t_2 - t_1 = \tau(g(t_2, \mathbf{P}_2) - g(t_1, \mathbf{P}_1)).$$

Če razvijemo desno stran zahteve po izpeljani formuli, dobimo pogoj

$$t_2 - t_1 = \alpha(t_2 - t_1) + \vec{a} \cdot (\mathbf{P}_2 - \mathbf{P}_1).$$

Iz tega sledi $\alpha = 1$ in $\vec{a} = \vec{0}$. Drug pogoj je, da se mora razdalja med istočasnimi dogodki ohranjati. Iz spodnjega dela bločne matrike dobimo pogoj

$$\|\mathbf{P}_2 - \mathbf{P}_1\| = \|Q(\mathbf{P}_2 - \mathbf{P}_1)\|,$$

torej mora biti Q ortogonalna. $\square$

Vprašanje 4. Kakšno obliko imajo Galilejeve transformacije $\mathbb{R}\times\mathbb{E}\to\mathbb{R}\times\mathbb{E}$? Dokaži.

Če definiramo $\vec{v}=\dot{\mathbf{P}}$ in $\vec{a}=\dot{\vec{v}}$, lahko opazujemo, kako se ti količini obnašata pri Galilejevi transformaciji. V koordinatnem sistemu $\phi'(t',\mathbf{P}')$ velja $\vec{v}'=\partial_{t'}\mathbf{P}'=\dot{\mathbf{P}}'$ in $\vec{a}'=\dot{\vec{v}}'$. Izpeljemo $\vec{v}'=\vec{c}+Q\dot{\mathbf{P}}(t'-t_0')=\vec{c}+Q\dot{\mathbf{P}}(t)$ in $\vec{a}'=Q\ddot{\mathbf{P}}(t)$.

Za sistem materialnih točk $\mathcal{P}=\{\mathbf{P}_1,\ldots,\mathbf{P}_n\}$ lahko definiramo

$$
\begin{aligned}
\underline{\mathbf{P}} &= (\mathbf{P}_1,\ldots,\mathbf{P}_n)\\
\underline{\mathbf{P}}_0' &= (\mathbf{P}_0',\ldots,\mathbf{P}_0')\\
\underline{\vec{c}} &= (\vec{c},\ldots,\vec{c})\\
\underline{\mathbf{P}}' = (\mathbf{P}_1',\ldots,\mathbf{P}_n') &= \underline{\mathbf{P}}_0'+\underline{\vec{c}}t+\underline{Q}(\underline{\mathbf{P}}-\underline{\mathbf{P}}_0)
\end{aligned}
$$

Gibanje lahko tedaj zapišemo s tremi principi.

- *Princip determiniranosti*: Trajektorija sistema materialnih točk $\mathcal{P}$ je v danem koordinantnem sistemu natanko določena z začetnim položajem in hitrostjo. To pomeni, da obstaja funkcija interakcije $\underline{\vec{f}}$, da velja

$$\ddot{\underline{\mathbf{P}}}=\underline{\vec{f}}(t,\underline{\mathbf{P}},\dot{\underline{\mathbf{P}}}).$$

- *Princip relativnosti*: Obstaja tak razred koordinatnih sistemov, v katerem je funkcija interakcije invariantna na Galilejeve transformacije. Temu razredu pravimo RAZRED INERCIALNIH KOORDINATNIH SISTEMOV. To pomeni, da je funkcija interakcije invariantna v tem razredu,

$$\ddot{\underline{\mathbf{P}}}'=\underline{\vec{f}}(t',\underline{\mathbf{P}}',\dot{\underline{\mathbf{P}}}').$$

- *Princip o sorazmernosti*: Obstajajo pozitivne konstante α_{ij}, da za vsako interakcijo med materialnimi točkami sistema $\mathcal{P}=(\mathbf{P}_1,\ldots,\mathbf{P}_n)$ velja

$$\vec{f}_i=-\sum_{j\neq i}\alpha_{ji}\vec{f}_j.$$

  Te konstante so enake za vse možne interakcije v sistemu.

Vprašanje 5. Kateri so principi gibanja?

Z ozirom na princip relativnosti izpeljemo $\underline{Q}\ddot{\underline{\mathbf{P}}}=\underline{Q}\underline{\vec{f}}(t,\underline{\mathbf{P}},\dot{\underline{\mathbf{P}}})$. Če v to enakost vstavimo vrednosti $t'=t_0'+t$, $\vec{c}=\vec{0}$, $Q=I$ ter $\mathbf{P}_0'=\mathbf{P}_0$, dobimo $\underline{\vec{f}}(t_0'+t,\underline{\mathbf{P}},\dot{\underline{\mathbf{P}}})=\underline{\vec{f}}(t,\underline{\mathbf{P}},\underline{\mathbf{P}}_0)$, kar mora veljati za vsak t_0'. Sledi, da funkcija $\vec{f}$ ne more biti eksplicitno odvisna od časa. Tej ugotovitvi pravimo HOMOGENOST ČASA.

Če sedaj vstavimo $\vec{c}=\vec{0}$, $Q=I$ in $\mathbf{P}_0'=\mathbf{P}_0+\vec{a}$, kjer je $\vec{a}$ poljuben vektor (in ne pospešek), izpeljemo $\mathbf{P}'=\mathbf{P}+\vec{a}$, in sledi $\underline{\vec{f}}(\underline{\mathbf{P}}+\underline{\vec{a}},\dot{\underline{\mathbf{P}}})=\underline{\vec{f}}(\underline{\mathbf{P}},\dot{\underline{\mathbf{P}}})$, torej $\vec{f}$ ne more biti odvisna od absolutnih položajev. Seveda je še vedno lahko odvisna od relativnih položajev (v tem primeru se $\vec{a}$ odšeteje). Tej lastnosti pravimo HOMOGENOST PROSTORA.

S poljubno izbiro vektorja $\vec{c}$ in $Q = I$ lahko podobno izpeljemo, da je $\underline{\vec{f}}$ lahko odvisna le od relativnih hitrosti, čemur pravimo HOMOGENOST PROSTORA HITROSTI.

Če nenazadnje relaksiramo še pogoj na Q, dobimo

$$\underline{\vec{f}}(Q(\mathbf{P}_i - \mathbf{P}_j), Q(\dot{\mathbf{P}}_i, \dot{\mathbf{P}}_j)) = \underline{Q\vec{f}}(\mathbf{P}_i - \mathbf{P}_j, \dot{\mathbf{P}}_i - \dot{\mathbf{P}}_j).$$

Funkcijam, ki zadoščajo temu pogoju, pravimo IZOTROPIČNE FUNKCIJE.

V posebnem primeru za $n = 1$ je $\vec{f}$ konstantna funkcija (ker ne more biti odvisna od ničesar). Ker za vsak $Q \in O(3)$ velja $\vec{f} = Q\vec{f}$, mora biti $\vec{f} = \vec{0}$. Torej se prosta materialna točka v inercialnem koordinantem sistemu premika premočrtno s konstantno hitrostjo. To je ena od implikacij v prvem Newtonovem zakonu.

Vprašanje 6. Izpelji homogenost časa in faznega prostora iz principov gibanja.

Definicija. Interakcija $\vec{f}$ je PARSKA, če lahko zapišemo

$$\vec{f}_i = \sum_{j \neq i} \vec{f}_{ji}(\mathbf{P}_i - \mathbf{P}_i, \dot{\mathbf{P}}_j - \dot{\mathbf{P}}_i)$$

za vse indekse i.

Definicija. Interakcija $\vec{f}$ je LOKALNA, če je parska in če velja

$$\lim_{\mathbf{P}_i - \mathbf{P}_j \to \infty} \vec{f}_{ji} = \vec{0}.$$

Vprašanje 7. Definiraj parske in lokalne interakcije.

Lema. *Za števila α_{ij} iz principa sorazmernosti velja*

- $\alpha_{ij}\alpha_{ji} = 1$,
- $\alpha_{ij}\alpha_{jk}\alpha_{kj} = 1$.

Dokaz. Prva točka: Izberemo si take interakcije $\vec{f}_k$, ki so parske in lokalne in ki so neodvisne od relativnih hitrosti. Vse točke razen i in j pošljemo v neskončnost, da je njihov vpliv ničeln. Tedaj velja $\vec{f}_i = -\alpha_{ji}\vec{f}_j$ in $\vec{f}_j = -\alpha_{ij}\vec{f}_i$, torej $\vec{f}_i = \alpha_{ij}\alpha_{ji}\vec{f}_i$.

Druga točka: Izberemo si indekse i, j, k in podobno kot prej pošljemo druge točke v neskončnost. Ob predpostavki parske in lokalne interakcije tako dobimo

$$\begin{aligned} \vec{f}_i &= -\alpha_{ji}\vec{f}_j - \alpha_{ki}\vec{f}_k, \\ \vec{f}_j &= -\alpha_{ij}\vec{f}_i - \alpha_{kj}\vec{f}_k. \end{aligned}$$

Če vstavimo drugo enačbo v prvo,

$$\vec{f}_i = \alpha_{ji}\alpha_{ij}\vec{f}_i + \alpha_{ji}\alpha_{kj}\vec{f}_k - \alpha_{ki}\vec{f}_k,$$

nam člen na levi in prvi člen na desni po prvi točki odpadeta. Dobljeno enačbo še pomnožimo z α_{ik} in nam ostane

$$\vec{f}_k = \alpha_{ji}\alpha_{kj}\alpha_{ik}\vec{f}_k.$$

□

Lema. *Naj za pozitivna števila α_{ij} velja ugotovitev prejšnje leme. Potem obstajajo števila m_i, da je $\alpha_{ji} = {}^{m_j}/_{m_i}$.*

Dokaz. Števila α_{ij} so definirana le za $i \neq j$. Definicijo lahko razširimo, da je $\alpha_{ii} = 1$. Definiramo $l_{ij} = \log \alpha_{ij}$. Velja $l_{ii} = 0$ in $l_{ij} = -l_{ji}$, poleg tega pa tudi $l_{ij} + l_{jk} + l_{ki} = 0$.

Izberemo si indeks i_0, ki nam bo definiral enoto mase. Velja $l_{i_0j} + l_{jk} + l_{ki_0} = 0$, kar odštejemo od prejšnje vsote treh členov in dobimo

$$l_{ij} - l_{i_0j} + l_{ki} - l_{ki_0} = 0.$$

Od tu izpeljemo, da za poljubna j in k velja

$$l_{ij} - l_{i_0j} = l_{ik} - l_{i_0k},$$

torej je $n_{ii_0} = l_{ij} - l_{i_0j}$ dobro definirana količina. Opazimo, da za $i = j$ velja $n_{ii_0} = l_{ii_0}$.

Definiramo $m_i = \exp n_{ii_0}$. Sledi

$$\log \alpha_{ij} = l_{ij} = l_{i_0j} + n_{ii_0} = -l_{ji} + n_{ii_0} = -n_{ji_0} + n_{ii_0} = \log m_i - \log m_j = \log \frac{m_i}{m_j}.$$

□

Opomba. Številom m_i pravimo INERCIJSKE MASE.

Vprašanje 8. Kaj so inercijske mase? Dokaži, da res obstajajo.

Produktu $m\vec{f} = \vec{F}$ pravimo SILA. Iz parskosti sledi

$$\vec{F}_i = \sum_{j \neq i} \vec{F}_{ji}(\mathbf{P}_i - \mathbf{P}_j, \dot{\mathbf{P}}_i - \dot{\mathbf{P}}_j).$$

Naj velja $\sum_{i \neq j} \vec{F}_{ji} = \vec{0}$. Predpostavimo, da so sile lokalne, in fiksiramo indeksa $k \neq l$. Če vsa ostala telesa pošljemo v neskončnost, ostane

$$\vec{F}_{kl} + \vec{F}_{lk} = \vec{0}.$$

S tem smo dokazali tretji Newtonov zakon.

Trditev (tretji Newtonov zakon)**.** *Če so vse sile parske in lokalne, velja $\vec{F}_{kl} = -\vec{F}_{lk}$.*

Za nadaljevanje potrebujemo še dodaten princip gibanja, ki ga imenujemo *princip o masi.* Pravi, da je inercijska masa enaka v vseh koordinatnih sistemih.

Vprašanje 9. Kaj je princip o masi?

Najpreprostejši primer sile je gravitacija. Med točkama $(m_1, \mathbf{P}_1)$ in $(m_2, \mathbf{P}_2)$ deluje sila

$$\vec{F}_{21} = \frac{\kappa M_1 M_2}{|\mathbf{P}_1 - \mathbf{P}_2|^2} \frac{\mathbf{P}_2 - \mathbf{P}_1}{|\mathbf{P}_2 - \mathbf{P}_1|}.$$

Številoma M_1 in M_2 pravimo GRAVITACIJSKI MASI. Z eksperimentiranjem je Newton ugotovil, da so pravzaprav enake inercijskim masam.

Definicija. Zunanja sila $\vec{F} = \vec{F}(t, \mathbf{P}, \dot{\mathbf{P}})$ je POTENCIALNA, če obstaja potencial U, da je $\vec{F} = -\vec{\nabla}_{\mathbf{P}} U$.

Definicija. DELO sile $\vec{F}$ pri gibanju materialne točke od $\mathbf{P}_1$ do $\mathbf{P}_2$ je krivuljni integral

$$A = \int_{\mathbf{P}_1}^{\mathbf{P}_2} \vec{F} \cdot d\mathbf{P} = \int_{t_1}^{t_2} \vec{F} \cdot \dot{\mathbf{P}} dt,$$

kjer smo pot parametrizirali s $\mathbf{P}(t)$. Produktu $\vec{F} \cdot \dot{\mathbf{P}}$ pravimo MOČ.

Definicija. KINETIČNA ENERGIJA T je enaka $\frac{1}{2} m \left|\dot{\mathbf{P}}\right|^2$.

Za rezultanto vseh sil $\vec{F}$ na telo m lahko izpeljemo

$$A = \int_{t_1}^{t_2} \vec{F} \cdot \dot{\mathbf{P}} dt = \int_{t_1}^{t_2} m \ddot{\mathbf{P}} \cdot \dot{\mathbf{P}} dt = m \int_{t_1}^{t_2} \partial_t \left(\frac{1}{2} \dot{\mathbf{P}} \cdot \dot{\mathbf{P}} \right) dt = T_2 - T_1,$$

kar lahko zapišemo v izrek.

Izrek (izrek o delu)**.** *Delo rezultante vseh sil je enako razliki kinetične energije telesa.*

Vprašanje 10. Povej in dokaži izrek o delu.

Definicija. Sila je KONZERVATIVNA v danem razredu inercialnih koordinatnih sistemov, če obstaja inercialni koordinatni sistem, v katerem je $\vec{F}$ potencialna in odvisna samo od položaja.

Tedaj je $\vec{F}$ potencialna, torej velja $\vec{F} = -\vec{\nabla}.U$ za nek potencial U, ki mu pravimo POTENCIALNA ENERGIJA. Velja

$$A = \int_{t_1}^{t_2} \vec{F} \cdot \dot{\mathbf{P}} dt = -\int_{t_1}^{t_2} \vec{\nabla}.U \cdot \dot{\mathbf{P}} dt = U(\mathbf{P}_1) - U(\mathbf{P}_2).$$

Vidimo, da je delo odvisno le od začetnega in končnega položaja. Sledi $T_2 - T_1 = U_1 - U_2$, torej je $T_1 + U_1 = T_2 + U_2 = E_0$ konstantna vrednost.

Izrek (izrek o energiji)**.** *Če je rezultanta vseh sil konzervativna, je vsota kinetične in potencialne energije konstanta gibanja.*

Vprašanje 11. Povej in dokaži izrek o energiji.

2.2 Premočrtno gibanje

Definicija. Gibanje je PREMOČRTNO, če ima pospešek konstantno smer.

Primer takega gibanja je poševni met. Opazimo, da lahko vedno izberemo koordinatni sistem, v katerem tir poti leži na premici: Če je $\vec{a} = a\vec{e}$, kjer je $\vec{e}$ konstanten vektor, velja

$$\vec{v} = \vec{e}\int_{t_0}^{t} a dt + \vec{v}_0.$$

Izberemo lahko sistem, kjer je $\vec{v}_0$ enak $\vec{0}$, in bo torej $\vec{v}$ vzporeden $\vec{e}$.

Če gibanje poteka pod vplivom konzervativne sile, lahko zapišemo potencial U, in velja izrek o energiji

$$\frac{1}{2}m\dot{x}^2 + U(x) = E_0.$$

Od tod izpeljemo

$$\dot{x} = \pm\sqrt{\frac{2}{m}\left(E_0 - U(x)\right)}.$$

Enačbo z ločljivimi spremenljivkami tedaj integriramo in dobimo

$$\pm\int_{x_0}^{x} \frac{dx}{\sqrt{\frac{2}{m}\left(E_0 - U(x)\right)}} = \int_{t_0}^{t} dt = t - t_0.$$

Dobimo funkcijo $t = t(x)$. Če se na poti ne ustavimo, po izreku o inverzni preslikavi obstaja funkcija $x = x(t)$. Pravimo, da je premočrtno gibanje INTEGRABILNO.

Če v kvalitativni analizi ugotovimo, da je neko gibanje periodično med točkama a in b, lahko periodo izračunamo kot

$$\begin{aligned} T &= \int_{x_0}^{b} \frac{dx}{\sqrt{\frac{2}{m}\left(E_0 - U(x)\right)}} - \int_{a}^{b} \frac{dx}{\sqrt{\frac{2}{m}\left(E_0 - U(x)\right)}} + \int_{a}^{x_0} \frac{dx}{\sqrt{\frac{2}{m}\left(E_0 - U(x)\right)}} \\ &= \sqrt{2m}\int_{a}^{b} \frac{dx}{\sqrt{E_0 - U(x)}}. \end{aligned}$$

Ker je $E = U(x)$ v krajiščih, je to posplošen integral. Situacija v obeh krajiščih je simetrična, torej preverimo le za levo krajišče, da integral res konvergira. V prvem koraku razvijemo preslikavo U v Taylorjev polinom prve stopnje v točki a, kjer se pojavi vrednost odvoda U v neki točki ξ blizu a. Ker je odvod zvezen, obstaja tak $\delta > 0$, da za $\xi \in [a, a+\delta)$ velja $2\partial_x U(a) < \partial_x U(\xi) < \frac{1}{2}\partial_x U(a)$, torej

$$\int_{a}^{a+\delta} \frac{dx}{\sqrt{E_0 - U(x)}} = \int_{a}^{a+\delta} \frac{dx}{\sqrt{-\partial_x U(\xi)(x-a)}} \leq \int_{a}^{a+\delta} \frac{1}{\sqrt{-\frac{1}{2}\partial_x U(a)}}\frac{1}{\sqrt{x-a}}dx < \infty.$$

Vprašanje 12. Izpelji izraz za periodo premočrtnega potencialnega gibanja.

Lema. *Za $a < b$ velja*

$$\int_a^b \frac{dx}{\sqrt{(b-x)(x-a)}} = \pi.$$

Dokaz. Uvedemo novo spremenljivko $x = \frac{1}{2}(a+b)+\frac{1}{2}(b-a)z$, s čimer se integral spremeni v

$$\int_{-1}^1 \frac{dz}{\sqrt{(1-z)(1+z)}} = \pi.$$

□

Primer. Oglejmo si harmonični oscilator, ki deluje pod potencialom $U = \frac{1}{2}kx^2$. Tedaj velja $F = -\partial_x U = -kx$, torej $m\ddot{x} = -kx$. Rešitev tega sistema je $x = A\cos\omega t + B\sin\omega t$ za $\omega = \sqrt{k/m}$. Iz tega lahko kar direktno preberemo $T = 2\pi/\omega$. Posebnost harmoničnega oscilatorja je, da je T neodvisen od E_0. Takemu gibanju pravimo IZOHRONIČNO, harmonični potencial je edini primer izohroničnega potenciala, ki je simetričen glede na svoj minimum.

Če ima potencial lokalni minimum v x_0, lahko za določanje potenciala uporabimo harmonično aproksimacijo. Zapišemo

$$\hat{U}(x) = U(x_0) + \partial_x U(x_0)(x-x_0) + \frac{1}{2}\partial_{x^2}U(x_0)(x-x_0)^2,$$

kar je harmonični potencial s periodo

$$T = 2\pi\sqrt{\frac{m}{\partial_{x^2}U(x_0)}}.$$

Ta aproksimacija je dobra, če velja $E_0 - U(x) \ll 1$.

Vprašanje 13. Izpelji harmonično aproksimacijo.

Druga vrsta aproksimacije, ki jo lahko uporabimo, je LIBRACIJSKA. Računamo

$$t = \operatorname{sgn}\dot{x}\int_{x_0}^x \frac{dx}{\sqrt{\frac{2}{m}(E_0 - U(x))}} = \sqrt{\frac{m}{2}}\operatorname{sgn}\dot{x}\int_{\theta_0}^{\theta} \frac{\frac{1}{2}(b-a)(-\sin\theta)d\theta}{\frac{1}{2}(b-a)\sqrt{\chi(\theta)}\sqrt{1-\cos^2\theta}}$$

za substitucijo $x = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)\cos\theta$ in $E_0 - U(x) = (x-a)(b-x)\chi(x)$. Pri tem smo si x predstavljali kot kosinus kota v krožnici, ki poteka skozi točki a in b in ima središče na njuni zveznici. Če računamo dalje, dobimo

$$t = \sqrt{\frac{m}{2}}\int_{\theta_0}^{\theta} \frac{1}{\sqrt{\chi(\theta)}}d\theta.$$

Če želimo dobiti periodo gibanja, bo θ tekel od 0 do π.

$$T = 2\sqrt{\frac{m}{2}}\int_0^{\pi} \frac{d\theta}{\sqrt{\chi(\theta)}}.$$

Ta integral aproksimiramo s trapezno formulo, ki je natančna v primeru, da je funkcija v integralu afina. Rešitev je tedaj

$$T \doteq \pi\sqrt{\frac{m}{2}}\left(\frac{1}{\sqrt{\chi(a)}}+\frac{1}{\sqrt{\chi(b)}}\right).$$

Vprašanje 14. Izpelji libracijsko aproksimacijo.

Trditev. *Za premočrtno potencialno periodično gibanje velja* $\frac{dS}{dE_0} = \frac{T}{m}$ *za ploščino* S *faznega diagrama.*

Dokaz. Velja $S = 2\sqrt{\frac{m}{2}}\int_a^b \sqrt{E_0 - U}dx$, torej

$$\frac{dS}{dE_0} = 2\sqrt{\frac{m}{2}}\left(b'\sqrt{E_0 - U(b)} - a'\sqrt{E_0 - U(a)} + \int_a^b \frac{dx}{2\sqrt{E_0 - U}}\right).$$

Ker je $U(a) = U(b) = E_0$, sta prva dva člena v oklepaju enaka 0, torej

$$\frac{dS}{dE_0} = \sqrt{\frac{m}{2}}\int_a^b \frac{dx}{\sqrt{E_0 - U}} = \frac{T}{m}.$$

□

Vprašanje 15. Kako se pri nihanju ploščina faznega portreta spreminja z energijskim nivojem E_0?

2.3 Gibanje po krivulji

Dana je krivulja $\vec{r} = \vec{r}(s(t))$, kjer je s naravni parameter. Če s $\mathbf{P}$ označimo trenutno lokacijo, velja

$$\vec{v} = \frac{d\mathbf{P}}{dt} = \frac{d\mathbf{P}}{ds}\frac{ds}{dt} = \vec{e}_t\dot{s},$$

kjer je $\vec{e}_t$ enotski vektor, tangenten na krivuljo, in

$$\vec{a} = \ddot{s}\vec{e}_t + \dot{s}\frac{d\vec{e}_t}{dt} = \ddot{s}\vec{e}_t + \dot{s}^2\kappa\vec{e}_n.$$

V enačbi κ predstavlja ukrivljenost, $\vec{e}_n$ pa normalo na krivuljo. Drugi Newtonov zakon poleg rezultante vseh sil vsebuje tudi silo vezi $\vec{S}$. Razpisan v smereh krivuljnega koordinatnega sistema ima obliko

$$\begin{aligned} m\ddot{s} &= \vec{F}\cdot\vec{e}_t + \vec{S}\cdot\vec{e}_t \\ m\kappa\dot{s}^2 &= \vec{F}\cdot\vec{e}_n + \vec{S}\cdot\vec{e}_n \\ 0 &= \vec{F}\cdot\vec{e}_b + \vec{S}\cdot\vec{e}_b \end{aligned}$$

Tu imamo štiri neznanke $(s, \vec{S})$ ter tri enačbe, torej potrebujemo še dodatno konstitutivno relacijo za silo vezi. Če se omejimo na gladke krivulje (take, kjer ni trenja), dobimo dodatno enačbo

$$\vec{S} \cdot \vec{e}_t = 0.$$

Delo take sile vezi je enako 0. Če je $\vec{F}$ konzervativna sila, $\vec{F} = -\vec{\nabla}.U$, dobimo

$$m\ddot{s} = -\frac{dU}{ds},$$

iz česar lahko izpeljemo energijsko enačbo

$$\frac{1}{2}m\dot{s}^2 + U(s) = E_0.$$

Torej je gibanje po gladki krivulji pod vplivom konzervativne sile reducibilno na premočrtno gibanje v ločni dolžini.

Vprašanje 16. Na kaj se reducira gibanje po gladki krivulji pod vplivom konzervativne sile? Izpelji.

Vprašanje 17. Obravnavaj matematično nihalo kot gibanje po gladki krožnici.

Odgovor: Če je l polmer krožnice, velja $s = l\theta$. Na točko poleg sile vezi deluje tudi teža, ki ima potencial

$$U = -m\vec{g}\vec{r} = -mgl\cos\frac{s}{l}.$$

Za dovolj majhen E_0 je gibanje periodično, in velja

$$T = \sqrt{2m}\int_{-s_0}^{s_0}\frac{ds}{\sqrt{E_0 + mgl\cos\frac{s}{l}}}.$$

Če uporabimo substitucijo $s = l\theta$ in začetno energijo zapišemo z začetnim odklonom, dobimo

$$T = \sqrt{2ml}\int_{-\theta_0}^{\theta_0}\frac{d\theta}{mgl(\cos\theta - \cos\theta_0)}.$$

Na tej točki upoštevamo, da je funkcija soda, in uporabimo $\cos\theta = 1 - 2\sin^2\frac{\theta}{2}$;

$$T = 2\sqrt{\frac{l}{g}}\int_0^{\theta_0}\frac{d\theta}{\sqrt{\sin^2\frac{\theta_0}{2} - \sin^2\frac{\theta}{2}}},$$

kar se s substitucijo $\sin\frac{\theta}{2} = u\sin\theta_0 2$ končno predela na

$$T = 4\sqrt{\frac{l}{g}}\int_0^1\frac{du}{\sqrt{(1-u^2)(1-u^2\sin^2\frac{\theta_0}{2})}}.$$

To je eliptični integral, odgovor je

$$T = 4\sqrt{\frac{l}{g}}K\left(\sin^2\frac{\theta_0}{2}\right)$$

⊠

2.4 Gibanje v polju centralne sile

Definicija. Sila $\vec{F} = \vec{F}(\mathbf{P})$ je CENTRALNA, če obstaja točka $\mathbf{O}$ (pol sile), da $\vec{F}$ deluje v smeri zveznice med $\mathbf{P}$ in $\mathbf{O}$, in da je njena velikost odvisna le od razdalje.

Trditev. *Konzervativna sila $\vec{F}$ je centralna natanko tedaj, ko obstaja pol, okoli katerega je vrtilna količina konstantna.*

Dokaz. Recimo, da je $\vec{F}$ centralna. Tedaj za vrtilno količino okoli $\mathbf{O}$, velja

$$\vec{l}(\mathbf{O}, \mathbf{P}) = (\mathbf{P} - \mathbf{O}) \times m\dot{\mathbf{P}}$$
$$\partial_t \vec{l} = \dot{\mathbf{P}} \times m\dot{\mathbf{P}} + (\mathbf{P} - \mathbf{O}) \times m\ddot{\mathbf{P}} = (\mathbf{P} - \mathbf{O}) \times \vec{F}$$

Če je $\vec{F}$ centralna, je vzporedna $\mathbf{P} - \mathbf{O}$, in se izniči tudi drugi člen.

Recimo, da je vrtilna količina konstantna. Po zgornjem izračunu $\dot{\vec{l}} = (\mathbf{P} - \mathbf{O}) \times \vec{F} = 0$, torej je $\vec{F}$ vzporedna zveznici. Pokazati moramo še, da je velikost sile odvisna le od razdalje. Po predpostavki je sila konzervativna, torej $\vec{F} = -\vec{\nabla}.U$ in

$$\vec{F} = -\left(\partial_r U \vec{e}_r + \frac{1}{r}\partial_\theta U \vec{e}_\theta + \frac{1}{r\sin\theta}\partial_\varphi U \vec{e}_\varphi\right)$$

v sferičnih koordinatah. Ker je sila vzporedna zveznici $\vec{e}_r$, mora veljati $U = U(r)$. □

Vprašanje 18. Definiraj centralno silo in jo karakteriziraj ob predpostavki, da je konzervativna. Dokaži karakterizacijo.

Trditev. *Zvezna centralna sila je konzervativna.*

Dokaz. Definiramo

$$U = \int_{|\mathbf{P}_0 - \mathbf{O}|}^{|\mathbf{P} - \mathbf{O}|} F(r)dr + U(\mathbf{P}_0).$$

□

Vprašanje 19. Dokaži: zvezna centralna sila je konzervativna.

Trditev. *Gibanje v polju centralne sile je ravninsko. Dogaja se na ravnini, ki vsebuje center sile in ima normalo v smeri $\vec{l}$.*

Dokaz. Računamo

$$(\mathbf{P} - \mathbf{P}_0) \cdot \vec{l} = (\mathbf{P} - \mathbf{O}) \cdot \vec{l} + (\mathbf{O} - \mathbf{P}_0) \cdot \vec{l} = 0 + 0 = 0,$$

upoštevaje definicijo $\vec{l}$. □

Vprašanje 20. Dokaži, da je gibanje v polju centralne sile ravninsko.

Izpeljimo kinematiko v polarnem koordinatnem sistemu. Definiramo

$$\begin{aligned} \vec{e}_r &= \cos\theta\vec{i} + \sin\theta\vec{j}, \\ \vec{e}_\theta &= \partial_\theta\vec{e}_r = -\sin\theta\vec{i} + \cos\theta\vec{j}, \end{aligned}$$

in izračunamo

$$\begin{aligned} \vec{r} &= r\vec{e}_r, \\ \vec{v} &= \dot{r}\vec{e}_r + r\dot{\theta}\vec{e}_\theta, \\ \vec{a} &= (\ddot{r} - r\dot{\theta}^2)\vec{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\vec{e}_\theta. \end{aligned}$$

Prvi komponenti hitrosti pravimo RADIALNA HITROST, drugi OBODNA HITROST. Podobno prvi komponenti pospeška pravimo RADIALNI POSPEŠEK, drugi pa OBODNI POSPEŠEK.

Za vrtilno količino velja

$$\vec{l} = \vec{r} \times m\vec{v} = mr^2\dot{\theta}\vec{k}.$$

Pogledamo lahko tudi ploščinsko hitrost

$$\dot{\vec{A}} = \frac{1}{2}\vec{r} \times \vec{v},$$

iz česar dobimo $\vec{l} = 2m\dot{\vec{A}}$ oziroma $\dot{\vec{A}} = \frac{1}{2}r^2\dot{\theta}\vec{k}$. Za gibanje v polju centralne sile je $\vec{l}$ konstantna, torej sta konstantni tudi ploščinska hitrost in DVOJNA PLOŠČINSKA HITROST

$$c_0 = r^2\dot{\theta}.$$

Vprašanje 21. Izpelji kinematiko v polarnem koordinatnem sistemu. Kaj je dvojna ploščinska hitrost?

Računamo lahko

$$\dot{r} = \frac{dr}{d\theta}\dot{\theta} = \frac{dr}{d\theta}\frac{c_0}{r^2} = -c_0\partial_\theta\left(\frac{1}{r}\right),$$

iz česar s spremenljivko $u = 1/r$, $u' = \partial_\theta u$ izpeljemo

$$\ddot{r} = -c_0 u''\dot{\theta} = -c_0^2 u^2 u''.$$

To uporabimo v BINETOVI FORMULI

$$a_r = \ddot{r} - r\dot{\theta}^2 = -c_0^2 u^2(u + u'').$$

Vprašanje 22. Izpelji Binetovo formulo.

Prvi Keplerjev zakon pravi, da se planeti gibljejo okoli Sonca v elipsah. Elipso lahko parametriziramo kot

$$r = \frac{p}{1 + \varepsilon\cos\theta}.$$

Z uporabo Binetove formule dobimo

$$a_r = -c_0^2 \frac{1}{pr^2}.$$

Za silo gravitacije $\vec{F} = -\kappa m M u^2 \vec{e}_r$ bo potem veljalo

$$\frac{c_0^2}{p} = \kappa M = \text{konst.}$$

Za tako parametrizacijo elipse velja

$$a = \frac{p}{1-\varepsilon^2}$$
$$b = \frac{p}{\sqrt{1-\varepsilon^2}}$$

Če je T perioda gibanja, je

$$A = \frac{1}{2} c_0 T,$$

torej

$$T = \frac{2\pi ab}{c_0}.$$

Drugi Keplerjev zakon pravi, da je kvadrat periode gibanja sorazmeren kubu večje polosi elipse, $T^2 = ka^3$, torej

$$\frac{p}{c_0^2} = \frac{k}{4\pi^2} = \frac{1}{\kappa M}.$$

Dobimo, da je k konstanten za vse planete.

Vprašanje 23. Povej drugi Keplerjev zakon. Pokaži, da je koeficient enak za vse planete.

Centralna sila je potencialna, $\vec{F} = -\vec{\nabla}.V$. Velja energijska enačba

$$\frac{1}{2} m v^2 + V(r) = E_0,$$

ki jo lahko predelamo v

$$\frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\theta}^2 + V(r) = E_0.$$

Upoštevaje $r^2 \dot{\theta} = c_0$ dobimo

$$\frac{1}{2} m \dot{r}^2 + \frac{l^2}{2mr^2} + V(r) = E_0.$$

Če zadnja dva člena na levi strani pospravimo v EFEKTIVNI POTENCIAL $U(r)$, smo reducirali gibanje na premočrtno s potencialom. Če to razrešimo na $r = r(t)$, lahko zapišemo

$$\theta(t) = \int_{t_0}^{t} \frac{c_0}{r^2(t)} dt.$$

Pravimo, da je gibanje v polju centralne sile INTEGRABILNO.

Vprašanje 24. Pokaži, da je gibanje v polju centralne sile integrabilno.

V primeru gravitacije integral žal ni zaprte oblike. Dobimo pa lahko enačbo trajektorije:

$$\frac{dr}{d\theta} = \frac{dr}{dt}\frac{dt}{d\theta} = \dot{r}\frac{1}{\dot{\theta}} = \pm\frac{1}{c_0}r^2\sqrt{\frac{2}{m}(E_0 - U(r))}.$$

Za $c_0 = l/m$ po integraciji dobimo

$$\theta - \theta_0 = \pm\frac{l}{\sqrt{2m}}\int_{r_0}^{r}\frac{dr}{r^2\sqrt{E_0 - U(r)}}.$$

Za gravitacijsko silo in $\gamma = \kappa m M$ velja

$$U(r) = \frac{l^2}{2mr^2} - \frac{\gamma}{r}.$$

Po integriranju dobimo

$$p = \frac{l^2}{m\gamma}, \qquad \varepsilon = \sqrt{\frac{2l^2}{m\gamma^2}E_0 + 1}.$$

Za $E_0 < 0$ je $\varepsilon < 1$, in dobimo elipso, pri $E_0 = 0$ dobimo parabolo $\varepsilon = 1$ in pri $E_0 > 0$ imamo hiperbolo za $\varepsilon > 1$.

Vprašanje 25. Kakšna je oblika tira planeta glede na energijo? Izpelji.

Če je gibanje periodično glede na efektivni potencial, se točka giblje med krožnicama med dvema APSIDNIMA RADIJEMA. Tir točke se dotika teh krožnic. Manjšemu od radijev pravimo PERICENTER, večjemu pa APOCENTER.

Trditev. *Tir je simetričen glede na apsidni radij.*

Dokaz. Velja

$$\theta^+ - \theta_0 = \frac{l}{\sqrt{2m}}\int_{r_a}^{r}\frac{dr}{r^2\sqrt{E_0 - U(r)}}$$
$$\theta^- - \theta_0 = -\frac{l}{\sqrt{2m}}\int_{r_a}^{r}\frac{dr}{r^2\sqrt{E_0 - U(r)}}$$

Sledi $\theta^+ - \theta_0 = \theta_0 - \theta^-$. □

Izračunamo lahko OVOJNO ŠTEVILO

$$\Delta\theta = \frac{l}{\sqrt{2m}}\int_{r_a}^{r_b}\frac{dr}{r^2\sqrt{E_0 - U(r)}}.$$

Tir gibanja bo zaprt takrat, ko je $\frac{\Delta\theta}{\pi} \in \mathbb{Q}$.

Trditev. *Tir je zaprt ali pa je gosta množica v kolobarju $K(0, r_a, r_b)$.*

Izrek (Bertrand)**.** *Vsi tiri v okolici krožnega tira so zaprti natanko tedaj, ko je V gravitacijski ali Hookov potencial.*

Vprašanje 26. Kakšen je tir gibanja točke v polju centralne sile?

2.5 Relativno gibanje

Definicija. Koordinatni sistem $\varphi(t, \mathbf{P})$ se GIBLJE glede na koordinatni sistem $\varphi'(t', \mathbf{P}')$, če obstaja trojica $(\mathbf{P}_0, \mathbf{P}_0', Q)$, kjer je $\mathbf{P}_0 \in \mathbb{E}$, $\mathbf{P}_0' : \mathbb{R} \to \mathbb{E}$, in $Q : \mathbb{R} \to SO(3)$, da velja $t = t'$ in

$$\mathbf{P}'(t) = \mathbf{P}_0'(t) + Q(t)(\mathbf{P} - \mathbf{P}_0).$$

Predpostavimo, da je φ' inercialen, in ga imenujmo ABSOLUTNI KOORDINATNI SISTEM, φ pa je RELATIVNI KOORDINATNI SISTEM.

Trditev. *Rotacijski del gibanja je neodvisen od izbire trojice.*

Dokaz. Recimo

$$\mathbf{P}' = \mathbf{P}_0' + Q(\mathbf{P} - \mathbf{P}_0) = \tilde{\mathbf{P}}_0' + \tilde{Q}(\mathbf{P} - \tilde{\mathbf{P}}_0).$$

Za $\mathbf{P} = \mathbf{P}_0$ dobimo

$$\mathbf{P}_0' = \tilde{\mathbf{P}}_0' + \tilde{Q}(\mathbf{P}_0 - \tilde{\mathbf{P}}_0).$$

Če to vstavimo nazaj gor, pridemo do

$$Q(\mathbf{P} - \mathbf{P}_0) = \tilde{Q}(\mathbf{P} - \mathbf{P}_0),$$

torej $Q = \tilde{Q}$. □

Vprašanje 27. Kdaj se koordinatni sistem giblje glede na nek drugi koordinatni sistem? Dokaži, da je rotacijski del neodvisen od izbire trojice.

Za odvod velja

$$\vec{v}' = \dot{\mathbf{P}}_0' + \dot{Q}(\mathbf{P} - \mathbf{P}_0) + Q\dot{\mathbf{P}} = Q(Q^T\vec{v}_0' + Q^T\dot{Q}(\mathbf{P} - \mathbf{P}_0) + \vec{v}_{\text{rel}}).$$

Prvemu členu pravimo TRANSLATORNA HITROST, drugemu ROTACIJSKA HITROST, tretjemu pa RELATIVNA HITROST.

Trditev. *$Q^T\dot{Q}$ je poševno simetrični tenzor.*

Dokaz. Velja $Q^TQ = I$. Če to odvajamo, dobimo

$$\partial_t(Q^T)Q + Q^T\dot{Q} = 0.$$

□

Izrek. *Naj bo W poševno simetričen na trirazsežnem evklidskem prostoru. Potem obstaja vektor $\vec{\omega}$ tako, da je $W\vec{a} = \vec{\omega} \times \vec{a}$ za vsak $\vec{a}$.*

Dokaz. Vemo, da obstaja lastna vrednost $W\vec{p} = \lambda\vec{p}$. Ker je

$$\lambda \left|\vec{p}\right|^2 = \vec{p} \cdot W\vec{p} = W^T\vec{p} \cdot \vec{p} = -\lambda \left|\vec{p}\right|^2,$$

torej $\lambda = 0$.

BŠS je $|\vec{p}| = 1$. Dopolnimo ga lahko do ortonormirane baze prostora $\{\vec{p}, \vec{q}, \vec{r}\}$, kjer je $\vec{r} = \vec{p} \times \vec{q}$. Ker je W poševno simetričen, je $\vec{a} \cdot W\vec{a} = 0$ za poljuben $\vec{a}$, in dobimo

$$W\vec{q} = W\vec{q} \cdot \vec{r}.\vec{r}, \qquad W\vec{r} = W\vec{r} \cdot \vec{q}.\vec{q}.$$

Za poljuben $\vec{a}$ je

$$W\vec{a} = \vec{a} \cdot \vec{q}.W\vec{q} \cdot \vec{r}.\vec{r} + \vec{a} \cdot \vec{r}.W\vec{r} \cdot \vec{q}.\vec{q}$$

Velja $\vec{q} \times \vec{r} = \vec{p}$ in $\vec{r} \times \vec{p} = \vec{q}$, torej

$$W\vec{a} = W\vec{q} \cdot \vec{r}.(\vec{a} \cdot \vec{q}.\vec{p} \times \vec{q} + \vec{a} \cdot \vec{r}.\vec{p} \times \vec{r}) = W\vec{q} \cdot \vec{r}.\vec{p} \times \vec{a}.$$

□

Vprašanje 28. Dokaži, da poševno simetričen tenzor deluje kot vektorski produkt.

Definicija. $[W_1, W_2] = W_1W_2 - W_2W_1$

Prostor poševno simetričnih tenzorjev z operacijama $+$ in $[,]$ je algebra.

Trditev. *Preslikava $W \mapsto \vec{\omega}(W)$ je homomorfizem.*

Dokaz. Pokazati želimo $\vec{\omega}([W_1, W_2]) = \vec{\omega}(W_1) \times \vec{\omega}(W_2)$. Velja

$$W_1W_2\vec{a} = \vec{\omega}_1 \times (\vec{\omega}_2 \times \vec{a}) = (\vec{\omega}_1 \cdot \vec{a}).\vec{\omega}_2 - (\vec{\omega}_1 \cdot \vec{\omega}_2).\vec{a},$$

torej dobimo

$$[W_1, W_2]\vec{a} = W_1W_2\vec{a} - W_2W_1\vec{a} = (\vec{\omega}_1 \times \vec{\omega}_2) \times \vec{a}.$$

□

Trditev. $(A\vec{a}) \times (A\vec{b}) = A^*(\vec{a} \times \vec{b})$

Dokaz. Če so $\vec{a}, \vec{b}, \vec{c}$ linearno neodvisni, velja

$$\det A = \frac{[A\vec{a}, A\vec{b}, A\vec{c}]}{[\vec{a}, \vec{b}, \vec{c}]}.$$

Tudi če niso neodvisni, velja

$$(A\vec{a} \times A\vec{b}) \cdot \vec{c} = [A\vec{a}, A\vec{b}, \vec{c}].$$

Za začetek predpostavimo, da je A obrnljiva. Tedaj

$$\begin{aligned}(A\vec{a} \times A\vec{b}) \cdot \vec{c} &= [A\vec{a}, A\vec{b}, AA^{-1}\vec{c}] \\ &= \det A \cdot [\vec{a}, \vec{b}, A^{-1}\vec{c}] \\ &= \det A \cdot (\vec{a} \times \vec{b}) \cdot A^{-1}\vec{c} \\ &= (\vec{a} \times \vec{b}) \cdot (A^*)^T \vec{c} \\ &= A^*(\vec{a} \times \vec{b}) \cdot \vec{c},\end{aligned}$$

kjer smo upoštevali $A^{-1} = (A^*)^T / \det A$. Če A ni obrnljiva, pa obstaja A_ε, da je $|A - A_\varepsilon| < \varepsilon$, torej v limiti velja za vse matrike. □

Posledica. *Če je $Q \in SO(3)$, je $Q(\vec{a} \times \vec{b}) = Q\vec{a} \times Q\vec{b}$.*

Vprašanje 29. Dokaži: če je $Q \in SO(3)$, je $Q(\vec{a} \times \vec{b}) = Q\vec{a} \times Q\vec{b}$.

Definicija. Vektor kotne hitrosti rotacije Q je vektor $\vec{\omega}' = Q\vec{\omega}$, kjer je $\vec{\omega}$ osni vektor poševno simetričnega tenzorja $W = Q^T\dot{Q}$.

Trditev. *Osni vektor poševno simetričnega tenzorja $\dot{Q}Q^T$ je vektor kotne hitrosti rotacije Q.*

Dokaz. Računamo

$$\begin{aligned}\dot{Q}Q^T\vec{a} &= QQ^T\dot{Q}Q^T\vec{a} \\ &= Q\vec{\omega} \times (Q^T\vec{a}) \\ &= Q\vec{\omega} \times (QQ^T\vec{a}) \\ &= (Q\vec{\omega}) \times \vec{a}.\end{aligned}$$

□

Trditev. *Vektor kotne hitrosti rotacije $R(\vec{e}, \varphi)$ okoli stalne osi $\vec{e}$ za kot φ je $\vec{\omega}' = \dot{\varphi}\vec{e}$.*

Dokaz. Imenujmo to rotacijo Q. Velja $Q\vec{e} = \vec{e}$; če to odvajamo po času, dobimo $\dot{Q}\vec{e} = 0$. Če to množimo z leve s Q^T, pridemo do $Q^T\dot{Q}\vec{e} = \vec{\omega} \times \vec{e} = 0$, torej je $\vec{\omega}$ vzporeden $\vec{e}$.

Dokazati moramo še, da je $\omega = \dot{\varphi}$. Naj bo $\vec{f}$ poljuben vektor dolžine 1. Velja $\vec{f} \cdot Q\vec{f} = \cos\varphi$. Če to odvajamo po času, dobimo

$$\vec{f} \cdot \dot{Q}\vec{f} = -\sin\varphi\dot{\varphi}$$

oziroma

$$Q^T\vec{f} \cdot Q^T\dot{Q}\vec{f} = -\dot{\varphi}\sin\varphi.$$

Levo stran enakosti lahko razpišemo v

$$Q^T\vec{f} \cdot Q^T\dot{Q}\vec{f} = Q^T\vec{f} \cdot (\vec{\omega} \times \vec{f}) = \vec{\omega} \cdot (\vec{f} \times Q^T\vec{f}) = \vec{\omega} \cdot (-\sin\varphi\vec{e}).$$

Sledi $\omega = \dot{\varphi}$. □

Vprašanje 30. Kaj je vektor kotne hitrosti rotacije? Kako ga izrazimo v primeru stalne osi?

Trditev. *Naj bo W poševno simetričen tenzor z enotskim osnim vektorjem $\vec{e}$. Potem je $e^{\theta W}$ rotacija okoli $\vec{e}$ za kot θ.*

Dokaz. Naj bo $A = e^{\theta W}$. Prvo dokažimo, da je $A \in SO(3)$. Če označimo $B = A^T A$ in to enakost odvajamo po θ, dobimo $B' = 0$, torej je B konstanta. Za $\theta = 0$ dobimo $B = I$. Pokazati moramo še, da je determinanta A enaka 1. Za to prvo pokažimo $\det e^{\theta W} = e^{\operatorname{sl} W}$. Če je $f(\theta) = \det e^{\theta W}$, je

$$f' = \frac{\partial \det A}{\partial \theta} * WA = A^* * WA = I * WA(A^*)^T = I * WAA^{-1} \det A = \det A \operatorname{sl} W,$$

kjer $*$ predstavlja skalarni produkt, če matrike vektoriziramo. Sledi $f' = \operatorname{sl} W \cdot f$, rešitev diferencialne enačbe je $f(\theta) = e^{\theta \operatorname{sl} W}$. Ker je $W^T = -W$, je $\operatorname{sl} W = 0$ in $\det A = 1$.

Razpis A v Taylorjevo vrsto pokaže $A\vec{e} = \vec{e}$. Če definiramo $Q(t) = e^{t\theta W}$, lahko izračunamo $Q^T \dot{Q} = \theta W$, iz česar dobimo osni vektor $\vec{\omega} = \theta \vec{e}$. To je enak osni vektor kot za rotacijo $R(\vec{e}, \theta)$. □

Izrek. $R(\vec{e}, \varphi) = \cos\varphi I + (1 - \cos\varphi)\vec{e} \otimes \vec{e} + \sin\varphi W(\vec{e})$

Dokaz. V izrazu nastopa tenzorski produkt $(\vec{a} \otimes \vec{b})\vec{c} = (\vec{b} \cdot \vec{c}).\vec{a}$.

Velja $R(\vec{e}, \varphi) = e^{\varphi W(\vec{e})}$. Kratek račun pokaže

$$W^k = \begin{cases} (-1)^{n+1}(\vec{e} \otimes \vec{e} - I) & k = 2n, n \geq 1 \\ (-1)^n W & k = 2n+1 \end{cases}$$

Če rotacijsko matriko razvijemo v Taylorjevo vrsto in zberemo lihe in sode člene, dobimo natanko želeno obliko. □

Vprašanje 31. Kako izrazimo rotacijo?

Iz vse te izražave lahko končno izračunamo

$$\begin{aligned} \dot{\mathbf{P}}' &= \dot{\mathbf{P}}_0' + \dot{Q}(\mathbf{P} - \mathbf{P}_0) + Q\dot{\mathbf{P}} \\ &= \dot{\mathbf{P}}_0' + QQ^T\dot{Q}(\mathbf{P} - \mathbf{P}_0) + Q\dot{\mathbf{P}} \\ &= \dot{\mathbf{P}}_0' + Q(\vec{\omega} \times (\mathbf{P} - \mathbf{P}_0)) + Q\dot{\mathbf{P}} \\ &= \vec{v}_0' + \vec{\omega}' \times Q(\mathbf{P} - \mathbf{P}_0) + Q\vec{v}_{\mathrm{rel}} \\ &= \vec{v}_0' + \vec{\omega}' \times \vec{\zeta}' + \vec{v}_{\mathrm{rel}}' \end{aligned}$$

Za $\zeta = \mathbf{P} - \mathbf{P}_0$. Če je $\vec{u}$ poljuben vektor, velja

$$\partial_t \vec{u}' = \partial_t(Q\vec{u}) = \dot{Q}\vec{u} + Q\dot{\vec{u}} = Q(\vec{\omega} \times \vec{u}) + Q\dot{\vec{u}},$$

torej transformacija in odvod po času komutirata le v primeru, da je $\vec{\omega}' \times \vec{u}' = 0$. Primer takega vektorja je $\vec{\omega}$, torej velja $\partial_t \vec{\omega}' = (\dot{\vec{\omega}})'$. Sedaj lahko izračunamo tudi pospešek

$$\begin{aligned}
\vec{a}' &= \partial_t \vec{v}' \\
&= \partial_t \vec{v}_0' + \dot{\vec{\omega}}' \times \vec{\zeta}' + \vec{\omega}' \times \partial_t \vec{\zeta}' + \partial_t \vec{v}_{\mathrm{rel}}' \\
&= \vec{a}_0' + \dot{\vec{\omega}}' \times \vec{\zeta}' + \vec{\omega}' \times (\vec{\omega}' \times \vec{\zeta}') + 2\vec{\omega}' \times \vec{v}_{\mathrm{rel}}' + Q\dot{\vec{v}}_{\mathrm{rel}} \\
&= \vec{a}_0' + \dot{\vec{\omega}}' \times \vec{\zeta}' + \vec{\omega}' \times (\vec{\omega}' \times \vec{\zeta}') + 2\vec{\omega}' \times \vec{v}_{\mathrm{rel}}' + \vec{a}_{\mathrm{rel}}'.
\end{aligned}$$

Prvi člen tega predpisa imenujemo POSPEŠEK IZHODIŠČA RKS, drugi člen EULERJEV POSPEŠEK, tretji CENTRIFUGALNI POSPEŠEK, četrti CORIOLISOV POSPEŠEK in peti RELATIVNI POSPEŠEK.

Vprašanje 32. Izpelji predpis za pospešek relativnega koordinatnega sistema in poimenuj člene.

Izrek. *Koordinatna sistema $\varphi(t, \mathbf{P})$ in $\varphi'(t, \mathbf{P}')$ sta v istem razredu Galilejevih koordinatnih sistemov natanko tedaj, ko koordinatna transformacija $\mathbf{P} \mapsto \mathbf{P}'$ preslika premočrtno gibanje s konstantno brzino v premočrtno gibanje s konstantno brzino.*

Dokaz. V desno je očitno iz zapisa pospeška. V levo računamo

$$\begin{gathered}
\vec{0} = \vec{a}_0 + \dot{\vec{\omega}}' \times \vec{\zeta}' + \vec{\omega}' \times (\vec{\omega}' \times \vec{\zeta}') + 2\vec{\omega}' \times \vec{v}_{\mathrm{rel}}', \\
\mathbf{P}' = \mathbf{P}_0' + t v_0 \vec{e}, \\
\vec{0} = \vec{a}_0' + t v_0 (\dot{\vec{\omega}}' \times \vec{e} + \vec{\omega}' \times (\vec{\omega}' \times \vec{e})) + 2 v_0 \vec{\omega}' \times \vec{e}.
\end{gathered}$$

To velja za vsak t, torej tudi za $t = 0$,

$$\vec{0} = \vec{a}_0' + 2 v_0 \vec{\omega}' \times \vec{e}.$$

Ta del je neodvisen od t, torej se pri poljubnem t pokrajša in dobimo

$$\vec{0} = \dot{\vec{\omega}}' \times \vec{e} + \vec{\omega}' \times (\vec{\omega}' \times \vec{e}).$$

Enačbo skalarno množimo z $\vec{e}$,

$$0 = \vec{\omega}' \times (\vec{\omega}' \times \vec{e}) \cdot \vec{e}.$$

To je mešani produkt, ki ga lahko ciklično zamenjamo in dobimo

$$0 = (\vec{\omega}' \times \vec{e}) \cdot (\vec{e} \times \vec{\omega}'),$$

iz česar sledi $\vec{\omega}' \parallel \vec{e}$. Sedaj upoštevamo prejšnjo enačbo in vidimo $\vec{a}_0' = \vec{0}$. Če vzamemo drug primer premočrtnega gibanja $\mathbf{P}' = \mathbf{P}_0' + t v_0 \vec{f}$, kjer $\vec{f}$ ni vzporeden $\vec{e}$, spet dobimo $\vec{\omega}' \parallel \vec{f}$, torej mora veljati $\vec{\omega}' = 0$. □

Vprašanje 33. Kaj velja za premočrtno gibanje pri koordinatni transformaciji? Dokaži.

2.6 Sistem materialnih točk

Imamo točke $\mathbf{P}_i, m_i$ za $i = 1, \dots, N$. Definiramo masno središče

$$\mathbf{P}_* = \mathbf{O} + \frac{1}{m} \sum_{i=1}^{N} m_i (\mathbf{P}_i - \mathbf{O}).$$

Izračunamo lahko, da je neodvisno od izbire točke $\mathbf{O}$, poleg tega pa velja tudi

$$m \ddot{\mathbf{P}}_* = \sum_{i=1}^{N} \vec{F}_i$$

za zunanje sile $\vec{F}_i$. Za vsako točko velja

$$m_i \ddot{\mathbf{P}}_i = \vec{F}_i + \sum_{j \neq i} \vec{F}_{ji}$$

Če to z leve vektorsko pomnožimo s $\mathbf{P}_i - \mathbf{O}$ in seštejemo po i, dobimo

$$\sum_{i=1}^{N} (\mathbf{P}_i - \mathbf{O}) \times m_i \ddot{\mathbf{P}}_i = \sum_{i=1}^{N} (\mathbf{P}_i - \mathbf{O}) \times \vec{F}_i + \sum_{i=1}^{N} \sum_{j \neq i} (\mathbf{P}_i - \mathbf{O}) \times \vec{F}_{ji}.$$

Prvi člen predstavlja rezultanto navora zunanjih sil, drugi pa rezultanto navora notranjih sil. Če definiramo vrtilno količino $\vec{l}(\mathbf{O}, \mathbf{P}_i) = (\mathbf{P}_i - \mathbf{O}) \times m_i \dot{\mathbf{P}}_i$ in $\vec{l} = \sum_i \vec{l}(\mathbf{O}, \mathbf{P}_i)$, lahko torej zapišemo

$$\dot{\vec{l}} = \vec{N}(\mathbf{O}) + \vec{N}_*(\mathbf{O}).$$

Definicija. Sila $\vec{F}_{ji}$ je CENTRALNA, če je oblike

$$\vec{F}_{ji} = F_{ji}(|\mathbf{P}_i - \mathbf{P}_j|) \frac{\mathbf{P}_i - \mathbf{P}_j}{\|\mathbf{P}_i - \mathbf{P}_j\|}.$$

Trditev. *Če so vse notranje sile centralne, je navor notranjih sil enak* 0.

Dokaz. Zapišemo

$$\sum_i \sum_j (\mathbf{P}_i - \mathbf{O}) \times \vec{F}_{ji} = \frac{1}{2} (\sum_i \sum_j (\mathbf{P}_i - \mathbf{O}) \times \vec{F}_{ji} + \sum_j \sum_i (\mathbf{P}_j - \mathbf{O}) \times \vec{F}_{ij}).$$

To razpišemo in dobimo 0. □

Izrek (Izrek o vrtilni količini)**.** *Če so vse notranje sile centralne, je odvod vrtilne količine enak rezultanti navora zunanjih sil,* $\dot{\vec{l}}(\mathbf{O}) = \vec{N}(\mathbf{O})$. *Pri tem predpostavimo, da je pol* $\mathbf{O}$ *fiksna točka.*

Vprašanje 34. Povej in dokaži izrek o vrtilni količini.

Če je $\mathbf{P}_0$ poljubna točka, lahko izračunamo

$$\vec{l}(\mathbf{P}_0) = \vec{l}(\mathbf{O}) + m(\mathbf{O} - \mathbf{P}_0) \times \dot{\mathbf{P}}_*.$$

V nadaljevanju predpostavimo, da so vse notranje sile centralne. Če zgornjo enačbo odvajamo, dobimo

$$\dot{\vec{l}}(\mathbf{P}_0) = \vec{l}(\mathbf{P}_0) - \dot{\mathbf{P}}_0 \times m\dot{\mathbf{P}}_*.$$

Izrek o vrtilni količini torej ohrani obliko za premikajoč pol, če velja:

- $\dot{\mathbf{P}}_0 = \vec{0}$,
- $\dot{\mathbf{P}}_* = \vec{0}$,
- $\mathbf{P}_0 = \mathbf{P}_*$.

Vprašanje 35. Za katere pole izrek o vrtilni količini ohranja obliko?

Če je sistem zaprt, velja $\vec{F} = 0$, iz česar dobimo $\vec{N} = 0$. Sklepamo, da se masno središče giblje premočrtno s konstantno hitrostjo, in da je vrtilna količina konstanta. Dodatno je konstantna $m\dot{\mathbf{P}}_*$, torej velja izrek o ohranitvi gibalne količine.

2.7 Togo telo

Definicija. Gibanje sistema materialnih točk je TOGO, če gibanje ohranja razdalje med točkami.

Definicija. Telo je TOGO, če nima koles. Alternativno je telo togo, če je edini njegov način gibanja togo gibanje.

Izrek. *Izometrija je afina preslikava.*

Dokaz. Izberimo $\mathbf{P}_0$ in točke $\mathbf{A}, \mathbf{B}, \mathbf{C}$, ki niso kolinearne. Naj bo $\mathbf{P}$ poljubna točka. Za $\vec{r} = \mathbf{P} - \mathbf{P}_0$, $\vec{a} = \mathbf{A} - \mathbf{P}_0$, $\vec{b} = \mathbf{B} - \mathbf{P}_0$ in $\vec{c} = \mathbf{C} - \mathbf{P}_0$ velja

$$\vec{r} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}.$$

Izometrija naj slika $\mathbf{T}$ v $\mathbf{T}'$, poleg tega naj je

$$\vec{r}' = \alpha'\vec{a}' + \beta'\vec{b}' + \gamma'\vec{c}'.$$

Če enačbo za $\vec{r}$ skalarno množimo z $\vec{a}$, $\vec{b}$ oziroma $\vec{c}$, dobimo enačbe

$$\begin{aligned}
\vec{r} \cdot \vec{a} &= \alpha\vec{a} \cdot \vec{a} + \beta\vec{b} \cdot \vec{a} + \gamma\vec{c} \cdot \vec{a} \\
\vec{r} \cdot \vec{b} &= \alpha\vec{a} \cdot \vec{b} + \beta\vec{b} \cdot \vec{b} + \gamma\vec{c} \cdot \vec{b} \\
\vec{r} \cdot \vec{c} &= \alpha\vec{a} \cdot \vec{c} + \beta\vec{b} \cdot \vec{c} + \gamma\vec{c} \cdot \vec{c}
\end{aligned}$$

To zložimo v matriko A, da dobimo

$$\begin{bmatrix} \vec{r}\cdot\vec{a} \\ \vec{r}\cdot\vec{b} \\ \vec{r}\cdot\vec{c} \end{bmatrix} = A \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$$

in podobno za količine s črto. Izometrija ohranja tako razdalje kot kote, torej sta levi strani sistemov in matriki A, A' enaki. Sledi torej $\alpha = \alpha'$, $\beta = \beta'$ in $\gamma = \gamma'$. Dobimo

$$\mathbf{P}' = \mathbf{P}'_0 + \vec{r}\,' = \mathbf{P}'_0 + A\vec{r} = \mathbf{P}'_0 + A(\mathbf{P} - \mathbf{P}_0)$$

in $A \in O(3)$. □

Togo gibanje lahko zapišemo v obliki

$$\mathbf{P}(t) = \mathbf{P}_0(t) + Q(t)(\mathbf{P}(t=0) - \mathbf{P}_0(t=0)).$$

Predstavimo ga kot relativno gibanje, kjer je RKS položaj telesa, kot ga vidimo ob času $t = 0$, AKS pa položaj telesa ob času t. RKS se giblje skupaj s telesom, zato mu pravimo TELESNI KS, absolutnemu pa pravimo PROSTORSKI KS.

Vprašanje 36. Kako opišeš gibanje togega telesa?

Za telo B definiramo volumensko in masno mero in predpostavimo, da za $B' = B(t)$ velja $m(B') = m(B)$. Sledi

$$m(B') = \iiint_{B'} \rho' dV' = \iiint_{B} \rho'(\mathbf{P}'(\mathbf{P})) \left|\det \frac{\partial \mathbf{P}'}{\partial \mathbf{P}}\right| dV.$$

Ker pa je $\mathbf{P}' = \mathbf{P}'_0 + Q(\mathbf{P} - \mathbf{P}_0)$, je torej odvod enak Q in njegova determinanta enaka 1. Torej

$$\iiint_B \rho'(\mathbf{P}'(\mathbf{P}))dV = m(B') = m(B) = \iiint_B \rho(\mathbf{P})dV,$$

torej (ker to velja tudi na delih telesa) $\rho'(\mathbf{P}'(\mathbf{P})) = \rho(\mathbf{P})$. Podobno izpeljemo, da je

$$\mathbf{P}'_* = \frac{1}{m(B)} \iiint_{B'} \mathbf{P}' - \mathbf{P}'_0 dm + \mathbf{P}'_0$$

enak $\mathbf{P}'_* = \mathbf{P}'_0 + Q(\mathbf{P}_* - \mathbf{P}_0)$ za $\mathbf{P}_*$, definiran s podobnim integralom. Podobno kot v diskretnem primeru lahko tudi tu izpeljemo

$$\dot{\mathbf{P}}'_* = \frac{1}{m(B')} \iiint_{B'} \dot{\mathbf{P}}' dm.$$

Vrtilno količino izračunamo kot

$$\begin{aligned}\vec{l}'(\mathbf{P}'_0) &= \iiint_{B'} (\mathbf{P}' - \mathbf{P}'_0) \times \partial_t(\mathbf{P}' - \mathbf{P}'_0) dm' \\ &= \iiint_{B'} \vec{\zeta}' \times (\vec{\omega}' \times \vec{\zeta}') dm' \\ &= \iiint_{B'} \left|\vec{\zeta}'\right|^2 \vec{\omega}' - (\vec{\zeta}' \cdot \vec{\omega}').\vec{\zeta}' dm' \\ &= \iiint_{B'} \left(\left|\vec{\zeta}'\right|^2 I - \vec{\zeta}' \otimes \vec{\zeta}' \right) dm' \cdot \vec{\omega}'.\end{aligned}$$

Iz tega dobimo definicijo prostorskega vztrajnostnega tenzorja

$$J'(\mathbf{P}'_0) = \iiint_{B'} \left(\left|\vec{\zeta}'\right|^2 I - \vec{\zeta}' \otimes \vec{\zeta}' \right) dm' = \iiint_{B'} \left|\mathbf{P}' - \mathbf{P}'_0\right|^2 I - (\mathbf{P}' - \mathbf{P}'_0) \otimes (\mathbf{P}' - \mathbf{P}'_0) dm'.$$

Podobno definiramo telesni vztrajnostni tenzor, izračunamo lahko $J' = QJQ^T$. Vztrajnostni tenzor je simetričen, poleg tega pa ga lahko diagonaliziramo

$$J(\mathbf{P}'_0) = \sum_{i=1}^{3} J_i \vec{w}_i \otimes \vec{w}_i,$$

kjer so J_i GLAVNE (lastne) vrednosti.

Trditev. *Vztrajnostni tenzor je nenegativen. Če B ni tanka palica, je pozitivno definiten.*

Dokaz. Trdimo, da je $\vec{e} \cdot J\vec{e} \geq 0$ za enotski $\vec{e}$. Po Cauchy-Schwarzu velja

$$\iiint_B \left|\vec{\zeta}\right|^2 - (\vec{e} \cdot \vec{\zeta})^2 dm \geq 0,$$

enakost dobimo natanko tedaj, ko je $\vec{e}$ vzporeden $\vec{\zeta}$, kar se zgodi le, če je B tanka palica. □

Vprašanje 37. Definiraj vztrajnostni tenzor in dokaži, da je pozitivno definiten.

Trditev. *Normala na ravnino zrcalne simetrije telesa je glavna smer vztrajnostnega tenzorja s polom na tej ravnini.*

Dokaz. Naj bo $\vec{e}$ ta normala. Trdimo $J(\mathbf{P}_0)\vec{e} = J\vec{e}$ za neko lastno vrednost J. To je ekvivalentno temu, da je deviator

$$D\vec{e} = \vec{\zeta} \otimes \vec{\zeta}\vec{e}$$

vzporeden $\vec{e}$. Naj bo $\vec{f} \perp \vec{e}$ in izračunajmo

$$\vec{f} \cdot D\vec{e} = \iiint_B \vec{f} \cdot \vec{\zeta}.\vec{e} \cdot \vec{\zeta} dm = \iiint_B \vec{f} \cdot \vec{\zeta}'.\vec{e} \cdot \vec{\zeta}' dm$$

za zrcalno sliko $\vec{\zeta}'$. Izračunamo lahko, da velja $\vec{e} \cdot \vec{\zeta}' = -\vec{e} \cdot \vec{\zeta}$ in $\vec{f} \cdot \vec{\zeta}' = \vec{f} \cdot \vec{\zeta}$, torej velja $\vec{f} \cdot D\vec{e} = 0$. □

Trditev. *Naj ima B dve ravnini simetrije z normalama $\vec{e}_1$ in $\vec{e}_2$. Potem ima $J(\mathbf{P}_0)$, kjer je $\mathbf{P}_0$ v preseku obeh ravnin, glavne smeri $\vec{e}_1, \vec{e}_2, \vec{e}_1 \times \vec{e}_2$.*

Vprašanje 38. Kaj lahko poveš o vztrajnostnemu tenzorju, če ima eno ali dve ravnini zrcalne simetrije? Dokaži.

Če razpišemo predpis za tenzorski produkt z uporabo $\mathbf{P} - \mathbf{P}_0 = (\mathbf{P} - \mathbf{P}_*) + (\mathbf{P}_* - \mathbf{P}_0)$, dobimo

$$J(\mathbf{P}_0) = J(\mathbf{P}_*) + m\,|\mathbf{P}_* - \mathbf{P}_0|^2\, I - m(\mathbf{P}_* - \mathbf{P}_0) \otimes (\mathbf{P}_* - \mathbf{P}_0).$$

Tej formuli pravimo STEINERJEV IZREK.

Vprašanje 39. Kaj pravi Steinerjev izrek?

Za kinetično energijo razpišemo

$$T = \frac{1}{2} \iiint_{B'} \left|\dot{\mathbf{P}}'\right|^2 dm',$$

pri čemer upoštevamo $\dot{\mathbf{P}}' = \dot{\mathbf{P}}'_* + \vec{\omega}\,' \times (\mathbf{P}' - \mathbf{P}'_*)$, ker se telo ne giblje v svojem sistemu in je zato $Q\dot{\mathbf{P}} = 0$. Na koncu dobimo

$$T = \frac{1}{2} m \left|\dot{\mathbf{P}}'_*\right|^2 + \frac{1}{2} \vec{\omega} \cdot J(\mathbf{P}_*)\vec{\omega}.$$

Vprašanje 40. Izpelji predpis za kinetično energijo togega telesa.

Pri dinamiki togega telesa imamo tri principe:

- $m\ddot{\mathbf{P}}'_* = \vec{F}'_{\text{zun}}$ (enačba gibanja masnega središča)
- $\partial_t \vec{l}\,'(\mathbf{P}'_*) = \vec{N}\,'(\mathbf{P}'_*)$ (princip o vrtilni količini)
- $\vec{N}\,'(\mathbf{P}'_0) = \vec{N}\,'(\mathbf{P}'_*) + (\mathbf{P}'_* - \mathbf{P}'_0) \times \vec{F}'_{\text{zun}}$

Izrek. *Naj bo $\mathbf{P}'_0$ točka trenutnega ali stalnega mirovanja togega telesa. Potem velja $\partial_t \vec{l}\,'(\mathbf{P}'_0) = \vec{N}\,'(\mathbf{P}'_0)$.*

Dokaz. Računamo

$$\begin{aligned}
\vec{l}\,'(\mathbf{P}'_0) &= Q\vec{l}(\mathbf{P}_0) \\
&= QJ(\mathbf{P}_0)\vec{\omega} \\
&= Q(J(\mathbf{P}_*) + m\,|\mathbf{P}_0 - \mathbf{P}_*|^2\, I - m(\mathbf{P}_0 - \mathbf{P}_*) \otimes (\mathbf{P}_0 - \mathbf{P}_*))\vec{\omega} \\
&= Q\vec{l}(\mathbf{P}_*) + m\,|\mathbf{P}_0 - \mathbf{P}_*|^2\, \vec{\omega} - m((\mathbf{P}_0 - \mathbf{P}_*) \cdot \vec{\omega}).(\mathbf{P}'_0 - \mathbf{P}'_*) \\
&= \vec{l}\,'(\mathbf{P}'_*) + m((\mathbf{P}'_0 - \mathbf{P}'_*) \times \vec{\omega}\,') \times (\mathbf{P}'_0 - \mathbf{P}'_*) \\
&= \vec{l}\,'(\mathbf{P}'_*) + m\dot{\mathbf{P}}'_* \times (\mathbf{P}_0 - \mathbf{P}_*).
\end{aligned}$$

Zadnji sklep uporabi dejstvo $\dot{\mathbf{P}}'_* = \dot{\mathbf{P}}'_0 + \vec{\omega}\,' \times (\mathbf{P}'_* - \mathbf{P}'_0)$, ki ga dobimo z odvajanjem $\mathbf{P}'_* = \mathbf{P}'_0 + Q(\mathbf{P}_* - \mathbf{P}_0)$. Upoštevamo še dejstvo, da je $\dot{\mathbf{P}}'_0 = 0$, ker je to točka trenutnega mirovanja. Zgornjo enakost nato še odvajamo do

$$\begin{aligned}\partial_t \vec{l}\,'(\mathbf{P}'_0) &= \vec{N}\,'(\mathbf{P}'_*) + m\ddot{\mathbf{P}}'_* \times (\mathbf{P}_0 - \mathbf{P}_*) + m\dot{\mathbf{P}}'_* \times (-\dot{\mathbf{P}}'_*) \\ &= \vec{N}\,'(\mathbf{P}'_*) + (\mathbf{P}'_* - \mathbf{P}'_0) \times \vec{F}\,'_{\text{zun}} \\ &= \vec{N}\,'(\mathbf{P}'_0).\end{aligned}$$

□

Navor okoli masnega središča ali navor okoli točke mirovanja lahko razpišemo v

$$\begin{aligned}\vec{N}\,'(\mathbf{P}'_*) &= \partial_t \vec{l}\,'(\mathbf{P}'_*) \\ &= \partial_t (QJ(\mathbf{P}_*)\vec{\omega}) \\ &= \dot{Q}J(\mathbf{P}_*)\vec{\omega} + QJ(\mathbf{P}_*)\dot{\vec{\omega}} \\ &= \vec{\omega}\,' \times QJ(\mathbf{P}_*)\vec{\omega} + QJ(\mathbf{P}_*)\dot{\vec{\omega}} \\ &= Q(\vec{\omega} \times J(\mathbf{P}_*)\vec{\omega} + J(\mathbf{P}_*)\dot{\vec{\omega}}).\end{aligned}$$

Iz tega dobimo Eulerjeve dinamične enačbe

$$\vec{N}(\mathbf{P}_*) = J(\mathbf{P}_*)\dot{\vec{\omega}} + \vec{\omega} \times J(\mathbf{P}_*)\vec{\omega}.$$

Enako dobimo za točko trenutnega mirovanja.

Vprašanje 41. Izpelji Eulerjeve dinamične enačbe. Dokaži, da delujejo tudi v točki mirovanja.

V lastnem KS vztrajnostnega momenta je J diagonalen. Dobimo

$$\begin{aligned}J_1\dot{\omega}_1 - \omega_2\omega_3(J_2 - J_3) &= N_1 \\ J_2\dot{\omega}_2 - \omega_1\omega_3(J_3 - J_1) &= N_2 \\ J_3\dot{\omega}_3 - \omega_1\omega_2(J_1 - J_2) &= N_3\end{aligned}$$

Podobno lahko naredimo za rotacijo okoli stalne osi $\vec{e}$

$$\begin{aligned}J_{13}\ddot{\varphi} - J_{23}\dot{\varphi}^2 &= N_1 \\ J_{23}\ddot{\varphi} + J_{13}\dot{\varphi}^2 &= N_2 \\ J_{33}\ddot{\varphi} &= N_3.\end{aligned}$$

Vprašanje 42. Izpelji Eulerjeve dinamične enačbe v lastnem koordinatnem sistemu in za rotacijo okoli stalne osi.

Poznamo volumenske sile $(B_1, \vec{f}_1)$, za katere velja

$$\vec{F}_1 = \iiint_{B_1} \vec{f}_1 dV.$$

Trditev. *Homogena volumenska sila je ekvivalentna točkovni sili s prijemališčem v masnem središču.*

Dokaz. Za silo je očitno, za navor pa velja

$$\vec{N}(\mathbf{O}) = \iiint_B (\mathbf{P} - \mathbf{O}) \times \vec{f} dm = \left(\iiint_B \mathbf{P} - \mathbf{O} dm \right) \times \vec{f} = (\mathbf{P}_* - \mathbf{O}) \times \vec{F}.$$

□

Za odvod kinetične energije velja

$$\partial_t T = m\dot{\mathbf{P}}'_* \cdot \ddot{\mathbf{P}}'_* + \vec{\omega} \cdot J_* \dot{\vec{\omega}} = \dot{\mathbf{P}}'_* \cdot \vec{F}' + \vec{\omega}' \cdot \vec{N}'_*,$$

kjer smo upoštevali Eulerjeve dinamične enačbe. Recimo, da je $\vec{F}$ volumenska sila, in da velja $\vec{f}' = -\vec{\nabla}.u$. Definiramo

$$U = \iiint_{B'} u(\mathbf{P}') dm'$$

in računamo

$$\partial_t U = \iiint_{B'} \partial_t u(\mathbf{P}') dm' = \iiint_{B'} \vec{\nabla}.u \cdot \dot{\mathbf{P}}' dm = - \iiint_{B'} \vec{f}' \cdot (\dot{\mathbf{P}}'_* + \vec{\omega} \times (\mathbf{P}' - \mathbf{P}'_*)) dm'$$

Sedaj ciklično zamenjamo člene mešanega produkta v drugem členu, do

$$\partial_t U = -\vec{F}' \cdot \dot{\mathbf{P}}'_* - \iiint_{B'} \vec{\omega} \cdot ((\mathbf{P}' - \mathbf{P}'_*) \times \vec{f}') dm = -\vec{F}' \cdot \dot{\mathbf{P}}'_* - \vec{\omega}' \cdot \vec{N}'_*.$$

Torej je $T + U$ konstantna. To je izrek o energiji za togo telo.

Vprašanje 43. Izpelji izrek o energiji za togo telo.

Definicija. Telo B_1 se kotali po B_2, če imata telesi v dotikališču enako hitrost.

2.7.1 Prosta vrtavka

Vrtavka je prosta, če je $\vec{F}' = 0$ in $\vec{N}' = 0$. Eulerjeve dinamične enačbe imajo tedaj obliko

$$\begin{aligned} J_1 \dot{\omega}_1 - \omega_2 \omega_3 (J_2 - J_3) &= 0, \\ J_2 \dot{\omega}_2 - \omega_1 \omega_3 (J_3 - J_1) &= 0, \\ J_3 \dot{\omega}_3 - \omega_1 \omega_2 (J_1 - J_2) &= 0. \end{aligned}$$

Enačba ima več rešitev:

- Trivialna: $\omega_1 = \omega_2 = \omega_3 = 0$

- Enakomerno vrtenje okoli ene osi: $\omega_1 = \omega_2 = 0$, $\omega_3 =$ konst..
- Netrivialne rešitve

Če je vrtavka simetrična, $J_1 = J_2$, je ω_3 konstantna, in velja

$$\dot{\omega}_1 - \frac{\omega_3(J_1 - J_3)}{J_1}\omega_2 = 0$$
$$\dot{\omega}_2 - \frac{\omega_3(J_3 - J_1)}{J_2}\omega_1 = 0$$

Za $\Omega = \omega_3(J_1 - J_3)/J_1$ je rešitev sistema

$$\omega_1 = A\cos(\Omega t - \delta)$$
$$\omega_2 = -A\sin(\Omega t - \delta)$$

Vektor $\vec{\omega}$ torej precesira s kotno hitrostjo Ω okoli tretje osi.

Vprašanje 44. Reši prosto simetrično vrtavko.

Če se preselimo v koordinatni sistem, v katerem je J diagonalen, velja $l_i = J_i\omega_i$. Velja $l^2 = l_1^2 + l_2^2 + l_3^2 =$ konst. in

$$T = \frac{1}{2}J_1\omega_1^2 + \frac{1}{2}J_2\omega_2^2 + \frac{1}{2}J_3\omega_3^2 = \frac{l_1^2}{2J_1} + \frac{l_2^2}{2J_2} + \frac{l_3^2}{2J_3}.$$

Iz tega dobimo

$$1 = \frac{l_1^2}{2TJ_1} + \frac{l_2^2}{2TJ_2} + \frac{l_3^2}{2TJ_3},$$

čemur pravimo BINETOV ELIPSOID. Vektor $\vec{l}$ leži na preseku sfere $\left|\vec{l}\right| = l$ in Binetovega elipsoida. Če uredimo $J_1 \leq J_2 \leq J_3$, dobimo $1 \leq l^2/(2TJ_1)$, torej za polos velja $a_1 \leq l$. Podobno dobimo $a_3 \geq l$. Če je rotacija le okoli največje ali najmanjše osi, se sfera in elipsoid dotikata v točki; za malo zmoten elipsoid imamo v preseku dve krožnici. Za srednjo polos pa majhna perturbacija elipsoida razklene točko v dve krivulji, ki prideta daleč od točke. Enakomerna rotacija okoli srednje osi je torej nestabilna.

Vprašanje 45. Analiziraj stabilnost enakomerne rotacije proste vrtavke.

2.7.2 Eulerjevi koti

Rotacijo predstavimo kot kompozitum rotacij

$$R = R(\vec{k}'', \psi)R(\vec{i}', \theta)R(\vec{k}, \varphi).$$

Lema. *Za* $\vec{f}' = R(\vec{e}, \varphi)\vec{f}$ *velja* $R(\vec{f}', \psi)R(\vec{e}, \varphi) = R(\vec{e}, \varphi)R(\vec{f}, \psi)$.

Trditev. *Velja* $R(\vec{k}'', \psi)R(\vec{i}', \theta)R(\vec{k}, \varphi) = R(\vec{k}, \varphi)R(\vec{i}, \theta)R(\vec{k}, \psi)$.

Dokaz za lemo in trditev je preprost račun.

Trditev. *Vsako rotacijo, ki ni rotacija okoli osi $\vec{k}$, lahko enolično napišemo z Eulerjevo rotacijo s koti $\theta, \varphi, \psi \in [0, 2\pi)$.*

Dokaz. Naj bo $\vec{\varepsilon}_1, \vec{\varepsilon}_2, \vec{\varepsilon}_3$ rotiran koordinatni sistem. Za vozliščnico $\vec{k} \times \vec{\varepsilon}_3 = \vec{e}$ definiramo kote

- φ je kot med $\vec{i}$ in $\vec{e}$,
- ψ je kot med $\vec{e}$ in $\vec{\varepsilon}_1$,
- θ je kot med $\vec{k}$ in $\vec{\varepsilon}_3$.

Za dokaz enoličnosti pogledamo delovanje preslikave

$$R(\vec{i}, \theta_1) = R(\vec{k}, \varphi_2 - \varphi_1) R(\vec{i}, \theta_2) R(\vec{k}, \psi_2 - \psi_1)$$

na vektorjih $\vec{i}$ in $\vec{k}$. □

Vprašanje 46. Dokaži, da se lahko vsako rotacijo zapiše na enoličen način z Eulerjevimi rotacijami.

Trditev. *Vektor kotne hitrosti Eulerjeve rotacije je $\vec{\omega}' = \dot{\varphi}\vec{k} + \dot{\theta}\vec{i}' + \dot{\psi}\vec{k}''$.*

Dokaz. Označimo $R = R(\vec{k}, \varphi) R(\vec{i}, \theta) R(\vec{k}, \psi) = R_1 R_2 R_3$. Po definiciji kotne hitrosti je

$$\begin{aligned}
\dot{R}\vec{a} &= \vec{\omega}' \times R\vec{a} \\
&= (\dot{R}_1 R_2 R_3 + R_1 \dot{R}_2 R_3 + R_1 R_2 \dot{R}_3)\vec{a} \\
&= \vec{\omega}'_1 \times R_1 R_2 R_3 \vec{a} + R_1(\vec{\omega}'_2 \times R_2 R_3 \vec{a}) + R_1 R_2 (\vec{\omega}'_3 \times R_3 \vec{a}) \\
&= (\vec{\omega}'_1 + R_1 \vec{\omega}'_2 + R_1 R_2 \vec{\omega}'_3) \times R\vec{a}.
\end{aligned}$$

Sledi

$$\begin{aligned}
\vec{\omega}' &= \vec{\omega}'_1 + R_1 \vec{\omega}'_2 + R_1 R_2 \vec{\omega}'_3 \\
&= \dot{\varphi}\vec{k} + R_1 \dot{\theta}\vec{i} + R'_2 R_1 \dot{\phi}\vec{k} \\
&= \dot{\varphi}\vec{k} + \dot{\theta}\vec{i}' + \dot{\psi}\vec{k}''.
\end{aligned}$$

□

Vprašanje 47. Kako izraziš vektor kotne hitrost Eulerjeve rotacije? Dokaži.

Če razpišemo $\vec{k}''$, $\vec{i}'$ in $\vec{j}'$, dobimo

$$\vec{\omega}' = (\dot{\theta}\cos\varphi + \dot{\psi}\sin\varphi\sin\theta)\vec{i} + (\dot{\theta}\sin\varphi - \dot{\psi}\cos\varphi\sin\theta)\vec{j} + (\dot{\varphi} + \dot{\psi}\cos\theta)\vec{k},$$

po drugi strani pa je $\vec{\omega}' = Q\vec{\omega}$. Za rotacijo Q velja

$$\vec{\omega}(Q) \times \vec{a} = Q^T \dot{Q}\vec{a} = -\dot{Q}^T Q\vec{a} = -\vec{\omega}'(Q^T) \times Q^T Q\vec{a} = -\vec{\omega}'(Q^T) \times \vec{a},$$

torej $\vec{\omega}(Q) = -\vec{\omega}'(Q^T)$. Sledi

$$\vec{\omega} = (\dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi)\vec{i} + (-\dot{\theta}\sin\psi + \dot{\varphi}\sin\theta\cos\psi)\vec{j} + (\dot{\psi} + \dot{\varphi}\cos\theta)\vec{k}$$

oziroma

$$\vec{\omega}' = (\dot{\theta}\cos\psi + \dot{\varphi}\sin\theta\sin\psi)\vec{\varepsilon}_1 + (-\dot{\theta}\sin\psi + \dot{\varphi}\sin\theta\cos\psi)\vec{\varepsilon}_2 + (\dot{\psi} + \dot{\varphi}\cos\theta)\vec{\varepsilon}_3,$$

kjer je $\vec{\varepsilon}_1 = \vec{i}'''$ in podobno za drugi komponenti.

Vprašanje 48. Izpelji predpis za $\vec{\omega}'$ v rotirani bazi.

Trditev. *Ne obstaja parametrizacija $SO(3)$, da je $\vec{\omega}' = \dot{\vec{x}}$.*

Dokaz. Če bi obstajala taka parametrizacija $\vec{x} = \vec{x}(\vec{y}) = \vec{x}(\varphi, \theta, \psi)$, potem velja

$$\vec{\omega}' = \dot{\vec{x}} = \frac{\partial \vec{x}}{\partial \vec{y}}\dot{\vec{y}} = \begin{bmatrix} \sin\theta\sin\psi & \cos\psi & 0 \\ \sin\theta\cos\psi & -\sin\psi & 0 \\ \cos\theta & 0 & 1 \end{bmatrix} \dot{\vec{y}}.$$

Računamo lahko

$$\frac{\partial^2 x_1}{\partial y_1 \partial y_2} = \cos\theta\sin\psi \qquad\qquad \frac{\partial^2 x_1}{\partial y_2 \partial y_1} = 0.$$

To ne mora biti res. $\square$

3 Uvod v numerične metode

3.1 Računske napake

Kadar z numerično metodo nekaj izračunamo, ne dobimo točne vrednosti, vendar nek približek. ABSOLUTNO NAPAKO definiramo kot razliko med približkom in točno vrednostjo:

$$d_a = \hat{x} - x.$$

Po drugi strani je RELATIVNA NAPAKA kvocient

$$d_r = \frac{\hat{x} - x}{x}.$$

Približek lahko izrazimo kot $\hat{x} = x(1 + d_r)$.

Vprašanje 1. Definiraj absolutno in relativno napako.

Števila predstavljamo s plavajočo vejico, ki je pravzaprav eksponentni zapis

$$x = \pm m \cdot b^e,$$

kjer je m MANTISA, zapisana kot $m = 0.c_1c_2 \dots c_t$ za $c_i \in \{0, \dots, b-1\}$, število b je BAZA zapisa, e pa EKSPONENT v mejah $L \leq e \leq U$. Števila običajno zapišemo NORMALIZIRANA, torej s $c_1 \neq 0$. V primeru najnižje možne potence dovoljujemo tudi SUBNORMALIZIRANA števila, kjer je $c_1 = 0$. Predstavljiva števila v takšnem zapisu označujemo s $P(b, t, L, U)$.

V standardu IEEE imamo dve števili:

- Enojni zapis: $P(2, 24, -125, 128)$,
- Dvojni zapis: $P(2, 53, -1021, 1023)$.

Vprašanje 2. Kaj je $P(b, t, L, U)$? Kakšne vrednosti imata `float` in `double`?

Pri zaokoževanju številu odrežemo decimalke za neko vrednostjo, in po potrebi prištejemo b^{-t}. Boljšega od teh približkov označimo s $\mathrm{fl}(x)$.

Izrek. *Če za x velja, da $|x|$ leži na intervalu med najmanjšim in največjim pozitivnim predstavljivim normaliziranim številom, potem velja*

$$\frac{|\mathrm{fl}(x) - x|}{|x|} \leq u,$$

za OSNOVNO ZAOKROŽITVENO NAPAKO $u = \frac{1}{2} b^{1-t}$.

Vprašanje 3. Kaj je osnovna zaokrožitvena napaka? Povej izrek.

Standard IEEE zagotovlja omejeno napako tudi pri osnovnih operacijah:

- $\mathrm{fl}(x \oplus y) = (x \oplus y)(1 + \delta)$ za $|\delta| \leq u$ za osnovne operacije $+, -, \cdot, /$,
- $\mathrm{fl}(\sqrt{x}) = \sqrt{x}(1 + \delta)$ za $|\delta| \leq u$.

Drug vir napak je občutljivost problema, ki ni povezana z numeriko. Obravanavamo vprašanje, kako se pri majhni spremembi v vhodnih podatkih spremeni pravilni odgovor. Za zvezno odvedljivo f lahko absolutno občutljivost merimo z odvodom

$$|f(x+\delta x)-f(x)| \approx \left|f'(x)\right| |\delta x| .$$

Poznamo tri vrste napak, ki skupaj sestavljajo celotno napako:

- Neodstranljiva napaka: napaka zaradi zaokroževanja podatkov
- Napaka metode: nenatančnost metode
- Zaokrožitvena napaka: napaka zaradi zaokroževanja znotraj metode

Vprašanje 4. Katere vrste napak poznamo?

3.2 Nelinearne enačbe

Iščemo rešitve enačbe $f(x) = 0$ za nek $f : \mathbb{R} \to \mathbb{R}$. Pri tem lahko pridemo do več različnih situacij glede obstoja in enoličnosti rešitve; možnost je, da rešitev obstaja in je ena sama, da je rešitev več, ampak končno mnogo, da jih je neskončno mnogo, ali pa da rešitve sploh ni.

Naj bo α ničla za zvezno odvedljivo f. Ničla je enostavna natanko tedaj, ko je $f'(\alpha) \neq 0$. Če ni enostavna, je m-kratna natanko tedaj, ko je prvi neničelni odvod v α reda m. V primeru enostavne ničle lokalno obstaja inverzna funkcija, da je $\alpha = f^{-1}(0)$. Absolutna občutljivost problema je tedaj enaka $\left|(f^{-1})'(0)\right| = \frac{1}{|f'(\alpha)|}$. Če je α dvojna ničla, uporabimo Taylorjev približek druge stopnje

$$f(x) \approx \underbrace{f(\alpha)}_{=0} + \underbrace{f'(\alpha)}_{=0}(x-\alpha) + \frac{1}{2}f''(\alpha)(x-\alpha)^2,$$

torej za $|f(x)| \leq \varepsilon$ velja

$$|x-\alpha| \leq \sqrt{\frac{2\varepsilon}{|f''(\alpha)|}}.$$

Višje kot so ničle, bolj občutljiv je problem iskanja. Za večkratno ničlo v splošnem velja

$$|x-\alpha| \leq \left(\frac{m!\varepsilon}{\left|f^{(m)}(\alpha)\right|}\right)^{1/m}$$

Vprašanje 5. Analiziraj problem iskanja ničel.

3.2.1 Bisekcija

Pri implementaciji bisekcije si hranimo zaporedja $(a_n)_n$ levih mej, $(b_n)_n$ desnih mej, $(c_n)_n$ sredinskih približkov, in $(e_n)_n$ polovičnih velikosti intervala. Nove člene izračunamo po predpisih

$$\begin{aligned} e_{n+1} &= e_n/2, \\ a_n &= a_{n-1} \text{ ali } c_{n-1}, \\ b_n &= b_{n-1} \text{ ali } c_{n-1}, \\ c_n &= a_n + e_n. \end{aligned}$$

Pri tem zmanjšamo število računskih operacij, se izognemo problemom glede možnih nepredvidenih zaokrožitev, skokov izven območja ali računskih napak. Bisekcija nam lahko poišče liho ničlo, ne pa sode, prav tako lahko poišče lih pol (ne pa sodega).

Vprašanje 6. Opiši delovanje bisekcije.

3.2.2 Navadna iteracija

Pri navadni iteraciji iskanje rešitve enačbe $f(x) = 0$ prevedemo na iskanje rešitve enačbe $x = g(x)$ za ustrezno izbrano funkcijo g. Splošna primerna izbira je recimo $g(x) = x - f(x)$, ali pa $g(x) = x - h(x)f(x)$ za neničelno funkcijo h. Da postopek $x_{r+1} = g(x_r)$ konvergira, mora biti g v okolici α skrčitev.

Izrek. *Naj bo $\alpha = g(\alpha)$ in naj g na intervalu $I = [\alpha - \delta, \alpha + \delta]$ za nek $\delta > 0$ zadošča Lipschitzovem pogoju $|g(x) - g(y)| \leq m\,|x - y|$ za nek $m \in [0, 1)$, in poljubna $x, y \in I$. Potem za vsak $x_0 \in I$ zaporedje $x_{r+1} = g(x_r)$ konvergira k α, in velja*

- $|x_r - \alpha| \leq m^r\,|x_0 - \alpha|$,
- $|x_{r+1} - \alpha| \leq \frac{m}{1-m}\,|x_{r+1} - x_r|$.

Dokaz. Dokažimo prvo, da zaporedje ne zapusti intervala I; če velja $x_r \in I$, je $|x_r - \alpha| \leq \delta$, torej

$$|x_{r+1} - \alpha| = |g(x_r) - g(\alpha)| \leq m\,|x_r - \alpha| < |x_r - \alpha| \leq \delta,$$

torej je tudi $x_{r+1} \in I$. S ponavljanjem tega postopka tudi dokažemo prvo točko, za drugo točko pa ocenimo

$$\begin{aligned} |x_{r+1} - \alpha| &= |x_{r+1} - x_{r+2} + x_{r+2} - x_{r+3} + x_{r+3} - \ldots - \alpha| \\ &\leq |x_{r+1} - x_{r+2}| + |x_{r+2} - x_{r+3}| + \ldots \\ &\leq (m + m^2 + m^3 + \ldots)\,|x_{r+1} - x_r| \\ &= \frac{m}{1-m}\,|x_{r+1} - x_r|\,. \end{aligned}$$

□

Posledica. *Če je $\alpha = g(\alpha)$, če je g zvezno odvedljiva in če velja $|g'(\alpha)| < 1$, potem obstaja nek $\delta > 0$, da za vsak x_0, $|x_0 - \alpha| < \delta$, zaporedje $x_{r+1} = g(x_r)$ konvergira k α.*

Vprašanje 7. Povej in dokaži izrek o navadni iteraciji.

Definicija. Naj bo $\lim x_r = \alpha$. Pravimo, da $(x_r)_r$ KONVERGIRA K α Z REDOM p, če velja

$$\lim_{r\to\infty} \frac{|x_{r+1} - \alpha|}{|x_r - \alpha|^p} = C$$

za neko konstanto $C > 0$.

Izrek. *Naj bo $\alpha = g(\alpha)$ za p-krat zvezno odvedljivo funkcijo g in naj velja $g'(\alpha) = \ldots = g^{(p-1)}(\alpha) = 0$ ter $g^{(p)}(\alpha) \neq 0$. Tedaj v bližini α zaporedje $x_{r+1} = g(x_r)$ konvergira k α z redom p.*

Dokaz. Izraz $x_{r+1} = g(x_r)$ v okolici α razvijemo v Taylorjevo vrsto:

$$x_{r+1} = g(\alpha) + g'(\alpha)(x_r - \alpha) + \ldots + \frac{g^{(p-1)}(\alpha)}{(p-1)!}(x_r - \alpha)^{p-1} + \frac{g^{(p)}(\xi)}{p!}(x_r - \alpha)^p$$

za nek ξ v bližini α. Sledi

$$\frac{x_{r+1} - \alpha}{(x_r - \alpha)^p} = \frac{g^{(p)}(\xi)}{p!}.$$

□

Vprašanje 8. Kaj je red konvergence zaporedja? Kako ga poiščeš z odvodom?

3.2.3 Tangentna metoda

Če vzamemo $\alpha = x_r + \Delta x_r$, in razvijemo dva člena Taylorjeve vrste

$$0 = f(x_r + \Delta x_r) = f(x_r) + f'(x_r)\Delta x_r + \frac{1}{2} f''(\xi_r)\Delta x_r^2,$$

ter nato zanemarimo zadnji člen, dobimo

$$\Delta x_r = \frac{-f(x_r)}{f'(x_r)},$$

s čimer smo izpeljali tangentno metodo, ki je pravzaprav poseben primer naravne iteracije. Če je α enostavna ničla, lahko izračunamo, da ima $g(x) = x - f(x)/f'(x)$ ničelni odvod v α (ob predpostavki $f \in \mathcal{C}^2$), in je torej konvergenca vsaj kvadratična. V primeru m-kratne ničle za $m \geq 2$ pa po dolgopisni izpeljavi dobimo, da je konvergenca zagotovljena, a linearna. Za dvakrat zvezno odvedljive funkcije je vsaka ničla f torej privlačna negibna točka.

Vprašanje 9. Izpelji tangentno metodo in pokaži, kakšen red konvergence ima.

Izrek. *Naj bo f na $I = [0, \infty)$ dvakrat zvezno odvedljiva, strogo naraščajoča, konveksna in naj ima na I ničlo α. Potem za vsak $x_0 \in I$ tangenta metoda konvergira k α.*

Dokaz. Velja $f'(x) > 0$ in $f''(x) > 0$. S Taylorjevim razvojem $f(\alpha)$ okoli točke x_0 dobimo oceno

$$x_1 - \alpha = \frac{f''(\xi)}{2f'(x_0)}(x_0 - \alpha)^2 \geq 0$$

pokažemo, da je ne glede na x_0 točka x_1 vedno desno od α. Dokažimo še, da je x_{r+1} nujno med α in x_r:

$$x_{r+1} = x_r - \frac{f(x_r)}{f'(x_r)} < x_r,$$

vedno pa velja $x_r > \alpha$. To je torej strogo padajoče omejeno zaporedje, ki mora nekam konvergirati; to bo seveda α. □

Vprašanje 10. Za kakšne funkcije lahko globalno zagotoviš konvergenco tangentne metode? Dokaži.

3.2.4 Sekantna metoda

Če je izračun odvoda zahteven, ga lahko aproksimiramo z diferenčnim kvocientom. Namesto tangente na f v točki x_r tako uporabimo sekanto skozi točki $(x_r, f(x_r))$ in $(x_{r-1}, f(x_{r-1}))$. Dobljena metoda tehnično ni navadna iteracija, ker uporablja zadnja dva približka, vendar se obnaša sorodno. Naslednji približek izračunamo s predpisom

$$x_{r+1} = x_r - \frac{f(x_r)(x_r - x_{r-1})}{f(x_r) - f(x_{r-1})}.$$

Analiza sekantne metode je težja kot analiza metod navadne iteracije. Izkaže se, da velja

$$|e_{r+1}| \approx c\,|e_r|\,|e_{r-1}|$$

za neko konstanto c, ki je pravzaprav enaka

$$c = \frac{|f''(\alpha)|}{2\,|f'(\alpha)|}.$$

Označimo red sekantne metode s p. Obstaja konstanta $D > 0$, da je $|e_{r+1}| \approx D\,|e_r|^p$, torej

$$|e_{r+1}| \approx CD\,|e_{r-1}|^{p+1} = D^{p+1}\,|e_{r-1}|^{p^2},$$

iz česar sledi $p^2 = p + 1$ oziroma $p = \phi$, torej je konvergenca superlinearna.

Vprašanje 11. Razloži sekantno metodo in izpelji njen red konvergence.

3.2.5 Ostale metode

Pri *Mullerjevi metodi* uporabimo tri približke x_r, x_{r-1} in x_{r-2}, ter skozi točke $(x_r, f(x_r))$, $(x_{r-1}, f(x_{r-1}))$, $(x_{r-2}, f(x_{r-2}))$ potegnemo polinom stopnje 2. Za naslednji približek vzamemo tisto izmed dveh ničel polinoma, ki je bližnja x_r. Ena od prednosti te metode je, da lahko išče kompleksne ničle tudi z realnimi začetnimi približki. Izkaže se, da je red konvergence približno 1.84.

Vprašanje 12. Razloži Mullerjevo metodo.

Če zamenjamo vlogi x in y, in najdemo polinom $p(y)$, ki poteka skozi točke $(f(x_r), x_r)$, $(f(x_{r-1}), x_{r-1})$, $(f(x_{r-2}), x_{r-2})$, dobimo približek za f^{-1}, in lahko za naslednji približek vzamemo $x_{r+1} = p(0)$. Metodo imenujemo *inverzna interpolacija*, ima pa isti red konvergence kot Mullerjeva metoda.

Vprašanje 13. Razloži metodo inverzne interpolacije.

3.2.6 Ničle polinomov

Ničle polinomov lahko iščemo na več načinov:

- Poiščeš eno ničlo in reduciraš polinom.
- Računaš vse ničle hkrati.
- Prevedeš problem na problem iskanja lastnih vrednosti.

Za redukcijo na problem lastnih vrednosti uporabljamo *spremljevalno matriko* polinoma $p(x) = a_0x^n + \ldots + a_n$

$$C_p = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -\frac{a_n}{a_0} & -\frac{a_{n-1}}{a_0} & \ldots & -\frac{a_2}{a_0} & -\frac{a_1}{a_0} \end{bmatrix}$$

Vprašanje 14. Kako izgleda spremljevalna matrika polinoma?

Ena od metod, ki računa vse ničle hkrati, je *Laguerrova metoda.* Za polinom $p(x) =$

$a_0x^n + \ldots + a_n$ z ničlami $\alpha_1, \ldots, \alpha_n$ definiramo

$$\begin{aligned} S_1(x) &= \sum_{i=1}^{n} \frac{1}{x-\alpha_i} = \frac{p'(x)}{p(x)}, \\ S_2(x) &= \sum_{i=1}^{n} \frac{1}{(x-\alpha_i)^2} = -S_1'(x) = \frac{(p'(x))^2 - p(x)p''(x)}{p^2(x)}, \\ a(x) &= \frac{1}{x-\alpha_n}, \\ b(x) &= \frac{1}{n-1}\sum_{i=1}^{n-1} \frac{1}{x-\alpha_i}, \end{aligned}$$

da velja $S_1(x) = a(x) + (n-1)b(x)$. Tedaj za

$$\begin{aligned} d_i(x) &= \frac{1}{x-\alpha_i} - b(x), \\ d(x) &= \sum_{i=1}^{n-1} d_i^2(x) \end{aligned}$$

dobimo $S_2 = a^2 + (n-1)b^2 + d$, ker je $\sum_i d_i = 0$. Dobili smo sistem enačb v spremenljivkah a, b, ki ga lahko rešimo in dobimo

$$a_{1,2} = \frac{1}{n}\left(S_1 \pm \sqrt{(n-1)(nS_2 - S_1^2 - nd)}\right).$$

Če x obravnavamo kot približek za ničlo α_n, bo člen nd v bližini α_n majhen, zato ga zanemarimo. Iz tega izrazimo

$$\alpha_n = x - \frac{n}{S_1 \pm \sqrt{(n-1)(nS_2 - S_1^2)}}.$$

Laguerrova metoda nam torej da postopek za izračun približka ničle

$$x_{r+1} = x_r - \frac{np(x_r)}{p'(x_r) \pm \sqrt{(n-1)\left[(n-1)(p'(x_r))^2 - np(x_r)p''(x_r)\right]}}.$$

Za odločitev, kateri predznak pripišemo korenu v imenovalcu, imamo tri možnosti:

- Vedno izberemo plus,
- Vedno izberemo minus,
- *Stabilna varianta*: izbereš tistega, ki ti v imenovalcu da večjo absolutno vrednost.

V prvih dveh primerih fiksno iščemo v eni smeri od začetnega približka, v tretjem pa to ni zagotovljeno.

Izrek. *Če ima polinom p same realne ničle, potem za vsak začetni približek x_0 stabilna verzija Laguerrove metode konvergira proti najbližji desni oz. levi ničli, pri čemer si mislimo, da sta kraka realne osi pri $+\infty$ in $-\infty$ združena. Konvergenca v bližini enostavne ničle je kubična.*

Če ima polinom kompleksne ničle, metoda konvergira za skoraj vse začetne približke.

Vprašanje 15. Izpelji Laguerrovo metodo in razloži, kako delujejo vse možnosti za izbiro naslednjega približka.

3.2.7 Durand-Kernerjeva metoda

Izberimo približke $x_1, \ldots, x_n$ za ničle polinoma $p(x)$ z vodilnih koeficientom 1. Iščemo popravke $\Delta x_1, \ldots, \Delta x_n$, da bodo $x_i + \Delta x_i$ točne ničle. Velja

$$\begin{aligned} p(x) &= (x - (x_1 + \Delta x_1))(x - (x_2 + \Delta x_2)) \ldots (x - (x_n + \Delta x_n)) \\ &= \prod_{i=1}^{n} (x - x_i) - \sum_{j=1}^{n} \Delta x_j \prod_{i \neq j} (x - x_i) + \ldots , \end{aligned}$$

člene drugega in večjega reda pa zanemarimo (torej vse, kar je v tropičju). Če s $q(x)$ označimo nezanemarjen del, velja

$$q(x_l) = -\Delta x_l \prod_{i \neq l} (x_l - x_i),$$

iz česar lahko izračunamo Δx_l.

Vprašanje 16. Razloži Durand-Kernerjevo metodo.

3.3 Sistemi linearnih enačb

3.3.1 Matrične norme

Definicija. Preslikava $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$ je VEKTORSKA NORMA, če velja

- $\|x\| \geq 0$ za vse x, in $\|x\| = 0 \Leftrightarrow x = 0$,
- $\|\alpha x\| = |\alpha| \, \|x\|$,
- $\|x + y\| \leq \|x\| + \|y\|$.

Vse vektorske norme so ekvivalentne, za poljubni normi $\|x\|_A$ in $\|x\|_B$ obstajata konstanti C_1, C_2, da velja

$$C_1 \|x\|_A \leq \|x\|_B \leq C_2 \|x\|_A .$$

Konkretno za 1-normo, 2-normo in supremum normo veljajo ocene

$$\begin{aligned}
\|x\|_2 &\le \|x\|_1 \le \sqrt{n}\,\|x\|_2 \\
\|x\|_\infty &\le \|x\|_1 \le \sqrt{n}\,\|x\|_\infty \\
\|x\|_\infty &\le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty
\end{aligned}$$

Vprašanje 17. Definiraj vektorsko normo. V kakšnem razmerju so znane vektorske norme?

Definicija. Preslikava $\|\cdot\| : \mathbb{C}^{n\times n} \to \mathbb{R}$ je MATRIČNA NORMA, če velja

- $\|A\| \ge 0$ in $\|A\| = 0 \Leftrightarrow A = 0$,
- $\|\alpha A\| = |\alpha|\,\|A\|$,
- $\|A + B\| \le \|A\| + \|B\|$,
- $\|AB\| \le \|A\|\,\|B\|$ (submultiplikativnost).

Matrika je tudi vektor, na njej so tudi definirane običajne vektorske norme. Definiramo funkcije

$$\begin{aligned}
N_1(A) &= \sum_{i,j} |a_{ij}|\,, \\
N_2(A) &= \sqrt{\sum_{i,j} |a_{ij}|^2}, \\
N_\infty(A) &= \max_{i,j} |a_{ij}|\,.
\end{aligned}$$

Te funkcije ustrezajo prvim trem točkam definicije matrične norme, to pa še ni dovolj. Za matriki

$$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

velja $N_\infty(A) = N_\infty(B) = 1$, ampak $N_\infty(AB) = 2$, torej N_∞ ni matrična norma. V nasprotju N_1 in N_2 dejansko sta matrični normi. Funkcijo N_2 imenujemo FROBENIUSOVA NORMA in označimo z $N_2(A) = \|A\|_F$.

Vprašanje 18. Dokaži, da N_∞ ni matrična norma. Kaj je Frobeniusova norma?

Izrek. *Naj bo $\|\cdot\|_v$ vektorska norma na $\mathbb{C}^n$. Potem je*

$$\|A\|_m = \max_{x\neq 0} \frac{\|Ax\|_v}{\|x\|_v}$$

matrična norma. Taki normi pravimo OPERATORSKA NORMA.

Dokaz. Prve tri točke so očitne, preverimo samo submultiplikativnost. Za vsak $x \neq 0$ velja $\|Ax\|_v \leq \|A\|_m \|x\|_v$ po definiciji, torej

$$\|AB\| = \max_{x\neq 0} \frac{\|ABx\|_v}{\|x\|_v} \leq \max_{x\neq 0} \frac{\|A\|_m \|Bx\|_v}{\|x\|_v} = \|A\|_m \|B\|_m .$$

□

Vprašanje 19. Dokaži, da so operatorske norme res matrične norme.

Za matrično normo $\|\cdot\|_m$ in vektorsko normo $\|\cdot\|_v$ pravimo, da sta USKLAJENI, če za vsako matriko A in vektor x velja

$$\|Ax\|_v \leq \|A\|_m \|x\|_v .$$

Lema. *Za vsako matrično normo obstaja usklajena vektorska norma.*

Dokaz. Za vektor x definiramo

$$\|x\|_v = \left\| \begin{bmatrix} x & 0 & \dots & 0 \end{bmatrix} \right\|_m ,$$

kjer smo vektor dopolnili do kvadratne matrike. To je očitno vektorska norma, normi sta očitno usklajeni. □

Vprašanje 20. Dokaži, da ima vsaka matrična norma usklajeno vektorsko normo.

Posledica. *Za vsako matrično normo in poljubno lastno vrednost λ matrike A velja $|\lambda| \leq \|A\|$.*

Dokaz. Naj bo $Ax = \lambda x$ za nek $x \neq 0$. Velja

$$|\lambda| \|x\|_v = \|\lambda x\|_v = \|Ax\|_v \leq \|A\| \|x\|_v .$$

□

Vprašanje 21. Kakšna je povezava med lastnimi vrednostmi in matrično normo? Dokaži.

Lema. *Norma $\|A\|_1$ je enaka največji 1-normi stolpca matrike A.*

Dokaz. Naj bo x vektor. Velja

$$\|Ax\|_1 = \left\| \sum_i x_i a_i \right\|_1 \leq \sum_i |x_i| \|a_i\|_1 \leq \max_{j=1,\dots,n} \sum_i |x_i| \|a_j\|_1 = \max_j \|x\|_1 \|a_j\|_1 .$$

Sledi

$$\frac{\|Ax\|_1}{\|x\|_1} \leq \max_j \|a_j\|_1 .$$

Enakost dobimo, če za x vzamemo e_k, kjer je k stolpec z največjo 1-normo. □

Podobno pokažemo, da je $\|A\|_\infty$ enaka največji 1-normi vrstice.

Vprašanje 22. Čemu sta enaki normi $\|A\|_1$ in $\|A\|_\infty$? Dokaži za eno.

Lema. *Velja* $\|A\|_2 = \max_j \sqrt{\lambda_j}$*, kjer so* λ_j *lastne vrednosti matrike* $A^H A$.

Dokaz. Matrika $A^H A$ je hermitska, torej so vse njene lastne vrednosti realne. Velja

$$x^H A^H A x = (Ax)^H A x = \|Ax\|_2 \geq 0,$$

torej je $A^H A$ pozitivno semidefinitna, in ima nenegativne lastne vrednosti. Izraz je torej res dobro definiran.

Naj bodo $\sigma_1^2 \leq \sigma_2^2 \leq \ldots \leq \sigma_n^2$ singularne vrednosti matrike A (lastne vrednosti $A^H A$). Ker je $A^H A$ hermitska, lahko lastne vektorje izberemo tako, da so ortonormirani. Naj za v_i torej velja $A^H A v_i = \sigma_i^2 v_i$. Za poljuben x velja

$$x = \sum_i \alpha_i v_i,$$

torej

$$\|Ax\|_2^2 = (Ax)^H (Ax) = x^H A^H A x = \sum_i |\alpha_i|^2 \sigma_i^2 \leq \sigma_n^2 \sum_i |\alpha_i|^2 = \sigma_n^2 \|x\|_2^2 .$$

Sledi neenakost

$$\frac{\|Ax\|_2}{\|x\|_2} \leq \sigma_n,$$

kjer dobimo enačaj, če vzamemo $x = v_n$. $\square$

Vprašanje 23. Čemu je enaka 2-norma matrike? Dokaži.

Frobeniusova norma ni operatorska, ker za vse operatorske norme velja $\|I\| = 1$, ampak $\|I\|_F = \sqrt{n}$. Kljub temu pa so vse matrične norme ekvivalentne. Za konkretne primere velja

$$\frac{1}{\sqrt{n}} \|A\|_F \leq \|A\|_2 \leq \|A\|_F ,$$
$$\frac{1}{\sqrt{n}} \|A\|_1 \leq \|A\|_2 \leq \sqrt{n} \|A\|_1 ,$$
$$\frac{1}{\sqrt{n}} \|A\|_\infty \leq \|A\|_2 \leq \sqrt{n} \|A\|_\infty .$$

Velja tudi $\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$ in $N_\infty(A) \leq \|A\|_2 \leq n N_\infty(A)$.

Vprašanje 24. Kako oceniš 2-normo matrike?

Lema. *Normi* $\|\cdot\|_2$ *in* $\|\cdot\|_F$ *sta invariantni na množenje z unitarno matriko.*

Dokaz. Za $x \in \mathbb{C}$ velja $\|Ux\|_2 = \|x\|_2$. Pri 2-normi torej

$$\|UA\|_2 = \max_x \frac{\|UAx\|_2}{\|x\|_2} = \max_x \frac{\|Ax\|_2}{\|x\|_2} = \|A\|_2\,,$$

za Frobeniusovo normo po

$$\|UA\|_F^2 = \|U[a_1 \dots a_n]\|_F^2 = \sum_i \|Ua_i\|_2^2 = \sum_i \|a_i\|_2^2 = \|A\|_F^2\,.$$

V drugo smer uporabimo dejstvi $\left\|A^H\right\|_2 = \|A\|_2$ in $\left\|A^H\right\|_F = \|A\|_F$. □

Vprašanje 25. Dokaži, da sta 2-norma in Frobeniusova norma invariantni na množenje z unitarno matriko.

Lema. *Naj bo $\|X\| < 1$. Potem je $I - X$ nesingularna, inverz je enak*

$$(I - X)^{-1} = \sum_{k=0}^{\infty} X^k,$$

in če je $\|I\| = 1$, velja

$$\left\|(I - X)^{-1}\right\| \leq \frac{1}{1 - \|X\|}.$$

Dokaz. Recimo, da je $I - X$ singularna. Tedaj obstaja vektor $w \neq 0$, da je $(I - X)w = 0$, torej $Xw = w$ in je w lastni vektor za lastno vrednost 1. To je protislovje, ker mora biti norma večja od lastne vrednosti.

Računamo

$$(I - X)\sum_{k=0}^{m} X^k = I - X^{m+1} \xrightarrow[m\to\infty]{} I,$$

ker je $\|X\| < 1$ in $\|X\|^{m+1} \geq \left\|X^{m+1}\right\|$. Dodatno velja

$$\left\|\sum_{k=0}^{\infty} X^k\right\| \leq \|I\| + \|X\| + \|X\|^2 + \ldots = \frac{1}{1 - \|X\|}.$$

□

Vprašanje 26. Povej predpostavke in dokaži formulo za inverz matrike $I - X$.

3.3.2 Občutljivost sistema linearnih enačb

Imejmo sistem $Ax = b$, kjer je A nesingularna matrika. Denimo, da A in b zmotimo v $A + \Delta A$ ni $b + \Delta b$, kjer je $A + \Delta A$ še vedno nesingularna. Nov sistem ima potem obliko

$$(A + \Delta A)(x + \Delta x) = b + \Delta b.$$

Lema. *Če je A nesingularna in $\|\Delta A\| < \frac{1}{\|A^{-1}\|}$, je $A + \Delta A$ nesingularna.*

Dokaz. Računamo

$$A + \Delta A = A(I + A^{-1}\Delta A),$$

in ocenimo normo

$$\left\|A^{-1}\Delta A\right\| \leq \left\|A^{-1}\right\| \|\Delta A\| < 1,$$

torej velja po prejšnji lemi. □

Naj torej velja ta pogoj za eno izmed znanih operatorskih norm. Računamo

$$\begin{gathered}
(A + \Delta A)\Delta x = \Delta b - \Delta A x, \\
\Delta x = (I + A^{-1}\Delta A)^{-1}(\Delta b - \Delta A x), \\
\frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A\| \left\|A^{-1}\right\|}{1 - \|A^{-1}\| \|\Delta A\|} \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right),
\end{gathered}$$

kjer smo v zadnjem delu uporabili sklep $\|b\| \leq \|A\| \|x\|$. Kvaliteta ocene je odvisna od vrednosti

$$\kappa(A) = \left\|A^{-1}\right\| \|A\|,$$

ki ji pravimo OBČUTLJIVOST MATRIKE ali POGOJENOSTNO ŠTEVILO.

Za 2-normo velja

$$\left\|A^{-1}\right\|_2 = \max_{x \neq 0} \frac{\left\|A^{-1}x\right\|_2}{\|x\|_2} = \max_{y \neq 0} \frac{\|y\|_2}{\|Ay\|_2} = \left(\min_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2} \right)^{-1} = \frac{1}{\sigma_n(A)},$$

kjer je $\sigma_n(A)$ najmanjša singularna vrednost. Dobimo torej

$$\kappa_2(A) = \frac{\sigma_1(A)}{\sigma_n(A)}.$$

Vprašanje 27. Kaj je občutljivost matrike? Kako jo uporabiš za oceno občutljivosti sistema linearnih enačb?

3.3.3 LU razcep

Naj bo dan vektor $w \in \mathbb{R}^n$ in naj velja $w_k \neq 0$. Definiramo ELIMINACIJSKO MATRIKO

$$L_k = \begin{bmatrix}
1 & & & & & & \\
& 1 & & & & & \\
& & \ddots & & & & \\
& & & 1 & & & \\
& & & -l_{k+1,k} & 1 & & \\
& & & \vdots & & \ddots & \\
& & & -l_{n,k} & & & 1
\end{bmatrix}$$

za $l_{i,k} = \frac{w_i}{w_k}$. Velja

$$L_k w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Če je $L_k = I - l_k e_k^T$, lahko izračunamo $L_k^{-1} = I + l_k e_k^T$. Če množimo matriko A z leve z L_1, uničimo prvi stolpec, razen prvega elementa. Če to matriko množimo z L_2 z leve, uničimo poddiagonalne elemente v drugem stolpcu (L_2 gradimo iz elementom te druge matrike). Tako lahko nadaljujemo in pridemo do

$$L_{n-1}L_{n-2}\dots L_2 L_1 A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n-1)} \end{bmatrix} = U$$

Za matriko $L = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1}$ tedaj velja $A = LU$. Izračunamo lahko

$$L = I + \sum_k l_k^T e_k,$$

iz česar vidimo, da je L spodnje trikotna z enicami na diagonali. Da LU razcep lahko naredimo, morajo biti elementi $a_{11}, a_{22}^{(1)}, \dots, a_{nn}^{(n-1)}$ na diagonali neničelni. Pravimo jim PIVOTI.

Vprašanje 28. Izpelji in razloži osnovni LU razcep. Kaj so pivoti?

Algorithm 1 LU razcep brez pivotiranja

```
for j = 1, ..., n − 1 do
    for i = j + 1, ..., n do
        l_ij = a_ij / a_jj
        for k = j + 1, ..., n do
            a_ik = a_ik − l_ij a_jk
        end for
    end for
end for
```

Osnovni postopek za izračun LU razcepa je prikazan v algoritmu 1. S tem dobimo elemente matrike L, razen enic na diagonali, ter elemente matrike U nad diagonalo. Algoritem deluje v $O(n^3)$ s predfaktorjem $2/3$.

Sistem $Ax = b$ rešimo tako, da prvo razcepimo $A = LU$, nato s premo substitucijo rešimo trikotni sistem $Ly = b$, in nazadnje z obratno substitucijo rešimo še $Ux = y$.

Vprašanje 29. Zapiši osnovni algoritem za LU razcep. Kakšno časovno zahtevnost ima?

Izrek. *Za matriko A je ekvivalentno*

- *Obstaja enoličen LU razcep $A = LU$, kjer je L spodnje trikotna z enicami na diagonali, in U zgornje trikotna nesingularna.*
- *Vse vodilne podmatrike A so nesingularne.*

Dokaz. V desno: A_k je produkt ustreznih vodilnih podmatrik L_k in U_k. V levo: Indukcija na n. Za matrike velikosti 1 ni nič za dokazati. Naj sedaj velja za n, dokažimo da velja tudi za $n+1$. Definiramo

$$\tilde{A} = \begin{bmatrix} A & b \\ c^T & \delta \end{bmatrix}.$$

Po indukcijski predpostavki lahko matriko A razcepimo v $A = LU$. Določimo lahko torej $u = L^{-1}b$, $l = U^{-T}c$ in $\xi = \delta - l^T u$, da dobimo

$$\tilde{A} = \begin{bmatrix} L & 0 \\ l^T & 1 \end{bmatrix} \cdot \begin{bmatrix} U & u \\ 0^T & \xi \end{bmatrix}.$$

Po predpostavki sta ξ in $\det U$ različni od 0, torej je $\det \tilde{U} = \xi \det U \neq 0$. □

Vprašanje 30. Kdaj je LU razcep enoličen? Dokaži.

Ničle in majhni elementi na diagonali so problem za LU razcep. Rešimo ga tako, da uvedemo delno oz. kompletno pivotiranje. Pri delnem pivotiranju na vsakem koraku namesto elementa v diagonali vzamemo element pod diagonalo, ki je največji po absolutni vrednosti, ter ga z menjavo vrstic postavimo na pivotno mesto. Rezultat tega je razcep $PA = LU$, kjer je P permutacijska matrika, ki ustreza menjavi vrstic.

Lema. *Če je A nesingularna, obstaja taka permutacijska matrika P, da za PA obstaja LU razcep brez pivotiranja.*

Dokaz. Vmesne matrike so oblike $L_{j-1}P_{j-1}\dots L_2P_2L_1P_1A$. Ker so vse naštete matrike nesingularne, mora biti tudi produkt nesingularen, torej obstaja neničelni element v ostanku stolpca. □

Vprašanje 31. Kako deluje LU razcep z delnim pivotiranjem? Pod katerim pogojem ga lahko naredimo?

Pri kompletnem pivotiranju poiščemo največji element v neobdelani podmatriki, in ga postavimo na pivotno mesto z zamenjavo stolpcev in vrstic. Dobimo razcep $PAQ = LU$, kjer P permutira vrstice in Q stolpce. V algoritmu dobimo $O(n^3)$ dodatnih primerjanj.

Vprašanje 32. Opiši LU razcep s kompletnim pivotiranjem.

Sistem $Ax = b$ rešimo numerično z eno izmed variant LU razcepa, in dobimo rešitev $\hat{x}$. Zanima nas ocena obratne stabilnosti $\|\Delta A\|$, kjer je $(A + \Delta A)\hat{x} = b$. Za analizo si mislimo, da smo pivotiranje že naredili, tako da lahko analiziramo le osnovni LU razcep.

Lema. *Naj bo $L \in \mathbb{R}^{n\times n}$ nesingularna spodnje trikotna matrika. Če sistem $Ly = b$ numerično rešimo s premo substitucijo, potem za izračunani $\hat{y}$ velja $(L+\Delta L)\hat{y} = b$, kjer je $|\Delta L| \le nu\,|L|$ po elementih.*

Dokaz. V i-tem koraku nastavimo

$$\hat{y}_i = \frac{1}{l_{ii}(1+\alpha_i)(1+\beta_i)}\left(b_i - \sum_{k=1}^{i-1} l_{ik}\hat{y}_k(1+\gamma_{ik})\right),$$

kjer so $\alpha_i, \beta_i, \gamma_{ik}$ manjše od osnovne računske napake u. Za $\gamma_{ii} = (1+\alpha_i)(1+\beta_i)$ velja $|\gamma_{ii}| \le 2u$, in $|\gamma_{ik}| \le (i-1)u$, torej za poljubna j, l velja $|\gamma_{jl}| \le nu$. □

Lema. *Če je U nesingularna zgornje trikotna matrika, in če sistem $Ux = y$ numerično rešimo z obratno substitucijo, izračunani $\hat{x}$ zadošča enačbi $(U + \Delta U)\hat{x} = y$, kjer je $|\Delta U| \le nu\,|U|$.*

Lema. *Naj bo A taka matrika, da se izvede LU razcep brez pivotiranja. Za izračunani matriki $\hat{L}$ in $\hat{U}$ tedaj velja $\hat{L}\hat{U} = A + E$, kjer je $|E| \le nu\left|\hat{L}\right|\left|\hat{U}\right|$.*

Izrek. *Če sistem $Ax = b$ rešimo z LU razcepom, potem za izračunani $\hat{x}$ velja $(A+\Delta A)\hat{x} = b$, kjer je $|\Delta A| \le 3nu\,|L|\,|U| + O(u^2)$.*

Posledica tega je, da velja $\|\Delta A\|_\infty \le 3nu\,\|L\|_\infty\,\|U\|_\infty$, to pa nam ne pomaga zares, ker je lahko produkt norm L in U poljubno velik v primerjavi z normo A.

Če ne pivotiramo, postopek ni obratno stabilen, če pa pivotiramo, pa je $|l_{ij}| \le 1$, torej je $\|L\|_\infty \le n$. Če vpeljemo PIVOTNO RAST

$$g = \frac{\max_{ij} |u|_{ij}}{\max_{ij} |a_{ij}|},$$

velja $\|U\|_\infty \le ng\,\|A\|_\infty$, in $\|\Delta A\|_\infty \le 3n^3gu\,\|A\|_\infty$. Pri delnem pivotiranju velja $g \le 2^{n-1}$, kar se v splošnem tudi lahko kdaj zgodi, recimo za matriko

$$\begin{bmatrix} 1 & & & & & 1 \\ -1 & 1 & & & & 1 \\ -1 & -1 & 1 & & & 1 \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ -1 & -1 & -1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & \cdots & -1 & 1 \end{bmatrix},$$

torej LU razcep z delnim pivotiranjem tudi ni obratno stabilen, v praksi pa se to redko zgodi.

Točne ocene za LU razcep s kompletnim pivotiranjem ne poznamo, smatramo pa, da je obratno stabilno.

Vprašanje 33. Analiziraj obratno stabilnost LU razcepa. Kaj je pivotna rast?

3.3.4 Razcep Choleskega

Izrek. *Veljajo naslednje točke:*

- *Če je A simetrična pozitivno definitna matrika, je vsaka vodilna podmatrika simetrično pozitivno definitna.*
- *Če je A simetrična pozitivno definitna, obstaja enoličen razcep $A = LU$, kjer je L spodnje trikotna matrika z enicami na diagonali in U zgornje trikotna matrika s pozitivnimi diagonalnimi elementi.*
- *A je simetrična pozitivno definitna natanko tedaj, ko je $A = VV^T$ za neko spodnje trikotno matriko V, ki ima pozitivne diagonalne elemente.*

Dokaz. Prva točka: Velja $x^T A_k x = \tilde{x}^T A \tilde{x}$, kjer je $\tilde{x}$ vektor x, dopolnjen z ničlami.

Druga točka: Vse vodilne podmatrike so simetrične pozitivno definitne, torej nesingularne in obstaja enoličen LU razcep. Če je $A = LU$, je $\det A_k = u_{11}u_{22}\dots u_{kk}$, torej so $u_{ii} > 0$.

Tretja točka: Razcepimo $A = LU$ in to dodatno razcepimo v $A = LDW$, kjer je D diagonalna matrika z elementi $u_{11}, \dots, u_{nn}$, in $W = D^{-1}U$ zgornje trikotna matrika z enicami na diagonali.

Ker je $A = A^T$, je $W^T DL^T$ LU razcep matrike A (če združimo desni dve matriki) in velja $W^T = L$ ter $DL^T = U$. Torej je $A = LDL^T$, in lahko definiramo $V = L\sqrt{D}$. □

Vprašanje 34. Karakteriziraj pozitivno definitnost matrike z razcepom Choleskega in dokaži karakterizacijo.

Če imamo izračunan razcep Choleskega, za $j < k$ velja

$$a_{jk} = \sum_{i=1}^{k-1} v_{ji}v_{ki} + v_{jk}v_{kk},$$

pri $j = k$ pa dobimo

$$a_{kk} = \sum_{i=1}^{k-1} v_{ki}^2 + v_{kk}^2.$$

Vidimo, da če poznamo vse elemente V pred v_{jk}, ga lahko direktno izračunamo. Iz tega dobimo algoritem 2. Ta porabi $\frac{1}{3}n^3$ operacij za račun razcepa, pri čemer pa računamo n korenov.

Algorithm 2 Razcep Choleskega

---

for $k = 1, \ldots, n$ **do**
  Nastavi
$$v_{kk} = \sqrt{a_{kk} - \sum_{i=1}^{k-1} v_{ki}^2}$$
  for $j = k+1, \ldots, n$ **do**
    Nastavi
$$v_{jk} = \frac{1}{v_{kk}} \left(a_{jk} - \sum_{i=1}^{k-1} v_{ji} v_{ki} \right)$$
  end for
end for

Vprašanje 35. Zapiši algoritem za izračun razcepa Choleskega.

Če rešujemo sistem z razcepom Choleskega, izračunamo rešitev $\hat{x}$. Tedaj vemo $(A + \Delta A)\hat{x} = b$, kjer velja $|\Delta A| \leq 3nu\,|V|\,|V^T|$. Ocenimo lahko

$$[|V|\,|V^T|]_{jk} = \sum_{i=1}^{\min(j,k)} |v_{ji}|\,|v_{ki}| \leq \sqrt{\sum_{i=1}^{j} |v_{ji}|^2} \sqrt{\sum_{i=1}^{k} |v_{ki}|^2}$$

po Cauchy-Schwarzu. To je nadalje enako

$$\leq \sqrt{a_{jj}}\sqrt{a_{kk}} \leq \|A\|_\infty .$$

Reševanje sistema z razcepom Choleskega je torej obratno stabilno, in velja

$$\|\Delta A\|_\infty \leq 3n^2 u\,\|A\|_\infty ,$$

kjer dodaten n na desni pride od tega, da smo na levi vzeli normo namesto absolutne vrednosti.

Vprašanje 36. Analiziraj stabilnost razcepa Choleskega.

3.4 Sistemi nelineranih enačb

Rešujemo sistem

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \ldots, x_n) &= 0 \end{aligned}$$

za $f_i : \mathbb{R} \to \mathbb{R}$ ali $\mathbb{C} \to \mathbb{C}$. Ekvivalentno $F(x) = 0$ za $F : \mathbb{R}^n \to \mathbb{R}^n$ (ali $\mathbb{C}^n \to \mathbb{C}^n$).

Prvi možni pristop reševanja je navadna iteracija. Sistem $F(x) = 0$ zapišemo v ekvivalenti obliki $x = G(x)$, izberemo $x^{(0)}$ ter iteriramo.

Izrek. *Naj bo $G : \mathbb{R}^n \to \mathbb{R}^n$ zvezno odvedljiva na zaprti množici $\Omega \subseteq \mathbb{R}^n$. Če za $x \in \Omega$ velja*

- $G(x) \in \Omega$,
- $\rho(JG(x)) \leq m < 1$,

kjer je JG Jacobijeva matrika, ρ pa spektralni radij (po absolutni vrednosti največja lastna vrednost), potem ima G na Ω natanko eno negibno točko α, in za vsak $x^{(0)} \in \Omega$ zaporedje $x^{(r+1)} = G(x^{(r)})$ konvergira k α.

Zadosten pogoj za konvergenco je že, da je $\|JG(\alpha)\| < 1$ v neki matrični normi. Za kvadratično konvergenco mora biti $JG(\alpha) = 0$ po komponentah.

Vprašanje 37. Kako poiščeš rešitev sistema nelinearnih enačb z navadno iteracijo? Povej izrek.

Podobno kot v enodimenzionalnem primeru lahko uporabimo razvoj v Taylorjevo vrsto in zanemarimo višje člene. Dobimo izraz za popravek

$$x^{(r+1)} = x^{(r)} - (JF(x^{(r)}))^{-1}F(x^{(r)}).$$

V praksi raje uporabimo algoritem 3.

Algorithm 3 Newtonova metoda

Izberi $x^{(0)}$.
for $r = 0, 1, 2, \ldots$ **do**
  Reši sistem $JF(x^{(r)})\Delta x^{(r)} = -F(x^{(r)})$.
  $x^{(r+1)} = x^{(r)} + \Delta x^{(r)}$.
end for

Vprašanje 38. Razloži Newtonovo metodo za rešitev sistema nelinearnih enačb.

Ker je računanje Jacobijeve matrike zahtevno, se lahko poslužimo kakšne kvazi-Newtonove metode. Pri taki metodi na različne načine aproksimiramo Jacobijevo matriko in zmanjšamo zahtevnost enega koraka. S tem običajno pade red konvergence na superlinearno. Najbolj znana kvazi-Newtonova metoda je Broydnova metoda, kjer približek Jacobijeve matrike B_{r+1} določimo kot najbližjo matriko, ki zadošča t.i. *sekantnemu pogoju*

$$B_{r+1}(x^{(r+1)} - x^{(r)}) = F(x^{(r+1)}) - F(x^{(r)}).$$

Ker je $B_r\Delta x^{(r)} = -F(x^{(r)})$, mora torej veljati

$$\Delta B_r \Delta x^{(r)} = F(x^{(r+1)}),$$

matrika ΔB_r pa je taka, da je $\|\Delta B_r\|_2$ minimalna.

Lema. *Dana sta neničelna vektorja x, y. Matrika A z minimalno normo, ki preslika x v y, je*

$$A = \frac{yx^T}{\|x\|_2^2}.$$

Dokaz. Očitno je $Ax = y$. Če za matriko B velja $Bx = y$, je $\|y\|_2 = \|Bx\|_2 \leq \|B\|_2 \|x\|_2$, torej

$$\|B\|_2 \geq \frac{\|y\|_2}{\|x\|_2}.$$

Po drugi strani se da preveriti, da za matrike ranga 1 velja $\left\|yx^T\right\|_2 = \|y\|_2 \|x\|_2$. □

Algorithm 4 Broydnova metoda

Določi $x^{(0)}$ in B_0.
for $r = 0, 1, \ldots$ **do**
  Reši $B_r \Delta x^{(r)} = -F(x^{(r)})$.
  Izračunaj

$$B_{r+1} = B_r + \frac{F(x^{(r+1)})(\Delta x^{(r)})^T}{\left\|\Delta x^{(r)}\right\|_2^2}.$$

end for

Vprašanje 39. Izpelji Broydnovo metodo.

3.5 Linearni problemi najmanjših kvadratov

3.5.1 Normalni sistem

Dana je matrika $A \in \mathbb{R}^{m \times n}$ za $m > n$ in vektor $b \in \mathbb{R}^m$. Iščemo $x \in \mathbb{R}^n$, ki minimizira napako $\|Ax - b\|_2$. Ta napaka bo minimalna, ko bo Ax pravokotna projekcija b na sliko $\operatorname{im} A$. Velja $Ax - b \perp Az$ za vse $z \in \mathbb{R}^n$ natanko tedaj, ko je za vsak z

$$z^T A^T (Ax - b) = 0.$$

Sledi $A^T(Ax - b) = 0$ oziroma $A^T Ax = A^T b$, čemur pravimo NORMALNI SISTEM. Pri tem smo tiho predpostavili, da je $\operatorname{rang} A = n$, sicer sistem nebi bil enolično rešljiv.

Velja $w^T A^T Aw = \|Aw\|_2^2 > 0$ za $w \neq 0$, torej je GRAMOVA MATRIKA $A^T A$ simetrična pozitivno definitna, in zanjo obstaja razcep Choleskega. Pri reševanju sistema seveda uporabimo ta razcep.

Vprašanje 40. Izpelji normalni sistem. Kaj je Gramova matrika?

3.5.2 QR razcep

Izrek. *Naj bo $A \in \mathbb{R}^{m\times n}$ za $m \geq n$ polnega ranga. Potem obstaja enoličen razcep $A = QR$, kjer je $Q \in \mathbb{R}^{m\times n}$ z ortonormiranimi stolpci ($Q^TQ = I_n$) in $R \in \mathbb{R}^{n\times n}$ zgornje trikotna s pozitivnimi diagonalnimi elementi.*

Dokaz. Če bi veljalo $A = QR$, je $A^TA = R^TQ^TQR = R^TR$. Matrika A^TA je simetrična pozitivno definitna, torej je R^TR njen razcep Choleskega, in velja $R = V^T$. Iz $A = QR$ sledi $Q = AR^{-1}$. □

Vprašanje 41. Povej in dokaži izrek o obstoju in enoličnosti QR razcepa.

Če poznamo $A = QR$, je $\operatorname{im} A = \operatorname{im} Q$. V drugem primeru bo vsakršno delo stabilnejše, ker so stolpci ortonormirani. Normalni sistem se tedaj prevede na $Rx = Q^Tb$, ki ga lahko rešimo s premo substitucijo.

3.5.3 Gram-Schmittova ortogonalizacija

Poznamo tri načine za izračun QR razcepa. Najenostavnejši pristop je Gram-Schmittova ortogonalizacija. Velja

$$a_k = \sum_{i=1}^{k-1} r_{ik}q_i + r_{kk}q_k.$$

Če to enačbo množimo z leve s q_j^T, nam ostane

$$r_{jk} = q_j^T a_k.$$

Celoten postopek je prikazan v algoritmu 5, ki ima računsko zahtevnost $2mn^2$.

Algorithm 5 QR razcep s klasičnim Gram-Schmittovim postopkom

for $k = 1, \ldots, n$ **do**
  $q_k = a_k$
  for $i = 1, \ldots, k-1$ **do**
    $r_{ik} = q_i^T a_k$
    $q_k = q_k - r_{ik}q_i$
  end for
  $r_{kk} = \|q_k\|_2$
  $q_k = \frac{1}{r_{kk}} q_k$
end for

Vprašanje 42. Izpelji in zapiši klasični Gram-Schmittov postopek za izračun QR razcepa.

V algoritmu naredimo še popravek, ki bo povečal stabilnost; pri računanju r_{ik} namesto formule $r_{ik} = q_i^T a_k$ uporabimo $r_{ik} = q_i^T q_k$. Novemu postopku pravimo MODIFICIRAN

Gram-Schmitt, in je v teoriji ekvivalenten klasičnemu. Modificiran postopek moramo tudi bolj pametno uporabiti; izračunamo QR razcep razširjene matrike

$$\begin{bmatrix} Q & q_{n+1} \end{bmatrix} \cdot \begin{bmatrix} R & z \\ 0 & \rho \end{bmatrix} = \begin{bmatrix} A & b \end{bmatrix},$$

in dobimo

$$Ax - b = \begin{bmatrix} Q & q_{n+1} \end{bmatrix} \cdot \begin{bmatrix} R & z \\ 0 & \rho \end{bmatrix} \cdot \begin{bmatrix} x \\ -1 \end{bmatrix} = \begin{bmatrix} Q & q_{n+1} \end{bmatrix} \cdot \begin{bmatrix} Rx - z \\ -\rho \end{bmatrix}.$$

Najboljšo rešitev dobimo, ko je $Rx = z$. Dejansko je $z = Q^T b$, le da smo ga v tem postopku izračunali z modificiranem Gram-Schmittovem postopkom, kar je numerično bolje.

Vprašanje 43. Kaj je modificiran Gram-Schmittov postopek? Kako ga pravilno uporabiš za reševanje sistema najmanjših kvadratov?

3.5.4 Givensove rotacije

Če je $c = \cos\varphi$ in $s = \sin\varphi$, matrika

$$R_{ik}^T(\varphi) = \begin{bmatrix} 1 &&&&&&&&& \\ & \ddots &&&&&&&& \\ && 1 &&&&&&& \\ &&& c &&&& s &&\\ &&&& 1 &&&&& \\ &&&&& \ddots &&&& \\ &&&&&& 1 &&& \\ &&& -s &&&& c && \\ &&&&&&&& 1 & \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{bmatrix},$$

ki ima elemente na diagonali in v stolpcih in vrsticah i, k, predstavlja rotacijo za φ v ravnini, ki jo razpenjata e_i in e_k v $\mathbb{R}^m$. Z ustrezno izbiro c in s lahko slikamo (x_i, x_k) v $(y_i, 0)$. Če je $r = \sqrt{x_i^2 + x_k^2}$, je taka izbira $c = x_i/r$ in $s = x_k/r$. Če te rotacije ustrezno kombiniramo, dobimo QR razcep, kakor je prikazano v algoritmu 6, ki ima zahtevnost $3mn^2 - n^3$ če ne računamo Q, za računanje Q pa porabimo še $6m^2n - 3mn^2$ operacij.

Vprašanje 44. Izpelji QR razcep z Givensonovimi rotacijami. Kakšna je njegova časovna zahtevnost?

Algorithm 6 QR razcep z Givensonovimi rotacijami

$Q = I_m$
for $i = 1, \dots, n$ **do**
  for $k = i+1, \dots, m$ **do**
    $r = \sqrt{a_{ii}^2 + a_{ki}^2}$
    $c = {}^{a_{ii}}/_{r}$
    $s = {}^{a_{ik}}/_{r}$
    Izračunaj

$$A([i,k], i:n) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \cdot A([i,k], i:n)$$
$$b([i,k]) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \cdot b([i,k])$$
$$Q([i,k], :) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \cdot Q([i,k], :)$$

  end for
end for
$Q = Q^T$

3.5.5 Householderjeva zrcaljenja

Vzemimo $w \in \mathbb{R}^m$, ki je različen od 0, in definirajmo

$$P = I - \frac{2}{w^T w} w w^T.$$

Velja $P = P^T$ in $P^2 = I$, poleg tega pa je w lastni vektor za P z lastno vrednostjo -1. Če je $u \perp w$, je $Pu = u$. Preslikavo lahko torej obravnavamo kot zrcaljenje čez ravnino, katere normala je w.

Če imamo dana dva enako dolga vektorja x, y, lahko z izbiro $w = x - y$ dobimo $y = Px$. Z izbiro $w = x \mp \|x\|_2 e_1$ se x preslika v

$$P \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \pm \|x\|_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Za numerično stabilnost si izberemo, da prištevamo, če je x_1 pozitiven, sicer odštevamo; $w = x + \operatorname{sgn} x_1 \|x\|_2 e_1$, kjer je $\operatorname{sgn} 0 \neq 0$. Z zrcaljenjem na enem koraku uničimo celoten stolpec matrike. Postopek izračuna QR razcepa je prikazan v algoritmu 7. Algoritem ima časovno zahtevnost $2mn^2 - \frac{2}{3}n^3$, če nas ne zanima Q.

Vprašanje 45. Izpelji QR razcep s Householderjevimi zrcaljenji.

Algorithm 7 QR razcep s Householderjevimi zrcaljenji

$Q = I_m$
for $i = 1, \ldots, n$ **do**
  Določi $w_i \in \mathbb{R}^{m-i+1}$ iz $A(i:m, i)$
  $A(i:m, i:n) = P_i A(i:m, i:n)$
  $b(i:m) = P_i b(i:m)$
  $Q(i:m, :) = P_i Q(i:m, i)$
end for
$Q = Q^T$

Občutljivost predoločenega sistema $Ax = b$ je odvisna od $\kappa_2(A) + \|r\|_2 \kappa_2^2(A)$ za $r = b - Ax$. Če uporabimo normalni sistem $A^T Ax = A^T b$, rešujemo z občutljivostjo $\kappa_2(A^T A) = \frac{\sigma_1^2(A)}{\sigma_n^2(A)} = \kappa_2^2(A)$, če pa uporabimo QR razcep, pa velja $\kappa_2(R) = \kappa_2(A)$.

Vprašanje 46. Kakšna je občutljivost reševanja predoločenega sistema?

3.6 Lastne vrednosti

Dana je matrika $A \in \mathbb{R}^{n \times n}$. Iščemo njene lastne vrednosti, in morda lastne vektorje. Iščemo lahko tudi leve lastne vektorje $y^H A = \lambda y^H$.

Lema. *Če je x desni lastni vektor A za lastno vrednost λ in je y levi lastni vektor za lastno vrednost $\mu \neq \lambda$, potem je $y^H x = 0$.*

Dokaz. Velja $\lambda y^H x = y^H Ax = \mu y^H x$, torej $y^H x = 0$. □

Posledica. *Če je A simetrična in sta x, y lastna vektorja za različni lastni vrednosti, je $y^T x = 0$.*

Vprašanje 47. Kakšni so lastni vektorji simetrične matrike? Dokaži.

Vprašanje 48. Zakaj Jordanova forma ni primerna za numerično računanje?

Odgovor: Poglejmo si primer

$$A(\varepsilon) = \begin{bmatrix} 0 & 1 \\ \varepsilon & 0 \end{bmatrix}.$$

Matrika $A(0)$ je že svoja Jordanova kletka, ki pa ni blizu Jordanove forme za $\varepsilon \neq 0$, ki je

$$J = \begin{bmatrix} \sqrt{\varepsilon} & \\ & -\sqrt{\varepsilon} \end{bmatrix}.$$

⊠

Izrek. *Za vsako matriko $A \in \mathbb{R}^{n \times n}$ obstaja Schurova forma $A = USU^H$, kjer je U unitarna in S zgornje trikotna.*

Dokaz. Dokažemo z indukcijo na n. Za $n = 1$ je primeren razcep $A = [1]A[1]$. Za splošen n pa: Vsaka matrika ima vsaj en lastni vektor, torej velja $Ax = \lambda x$ za nek x velikosti 1, in nek λ. Za U_1 izberemo tako unitarno matriko, da je $U_1 e_1 = x$, in definiramo $B = U_1^H A U_1$. Velja $Be_1 = \lambda e_1$, torej je B oblike

$$B = \begin{bmatrix} \lambda & y^T \\ 0 & C \end{bmatrix}.$$

Za matriko C velja indukcijska predpostavka: obstajata U_2 in S_2, da je $C = U_2 S_2 U_2^H$, in je U_2 unitarna, S_2 pa zgornje trikotna. Definiramo

$$U = U_1 \begin{bmatrix} 1 & \\ & U_2 \end{bmatrix},$$
$$S = U^H A U.$$

□

Vprašanje 49. Dokaži, da za vsako matriko obstaja Schurova forma.

Izrek. *Če je $A \in \mathbb{R}^{n \times n}$, obstajata ortogonalna Q in kvazi zgornje trikotna T, obe realni, da je $A = QTQ^T$.*

Kvazi zgornje trikotna matrika je zgornje trikotna, razen da na diagonali dopuščamo bloke 2x2. V vsakem takem bloku se nahajajo konjugirani pari kompleksnih lastnih vrednosti. Razcepu $A = QTQ^T$ pravimo REALNA SCHUROVA FORMA.

Vprašanje 50. Kaj je realna Schurova forma?

3.6.1 Potenčna metoda

Algorithm 8 Potenčna metoda

```
Izberemo nek z_0 ≠ 0.
for k = 1, 2, ... do
    y_k = A z_{k-1}
    z_k = y_k / ||y_k||
end for
```

Izrek. *Naj bo λ_1 dominantna lastna vrednost matrike A (največja po absolutni vrednosti in strogo večja od druge največje). Za naključno izbrani začetni vektor z_0 vektorji z_k iz potenčne metode po smeri konvergirajo k lastnemu vektorju za λ_1.*

Dokaz. Dokažemo le za primer, ko je A diagonalizabilna. Naj bo $A = XDX^{-1}$ za $X = [x_1, \ldots, x_n]$ in $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Začetni vektor z_0 lahko razvijemo v bazi lastnih vektorjev

$$z_0 = \sum_{i=1}^{n} \alpha_i x_i.$$

Za vsak k velja

$$z_k = \frac{A^k z_0}{\|A^k z_0\|} = \frac{\sum_i \alpha_i \lambda_i^k x_i}{\left\|\sum_i \alpha_i \lambda_i^k x_i\right\|}.$$

Ulomek na obeh straneh delimo z λ_1^k:

$$z_k = \frac{\sum_i \alpha_i \lambda_i^k \lambda_1^{-k} x_i}{\left\|\sum_i \alpha_i \lambda_i^k \lambda_1^{-k} x_i\right\|}$$

Če je $\alpha_1 \neq 0$, bo to po smeri konvergiralo k x_1. □

Vprašanje 51. Zapiši algoritem potenčne metode in dokaži pravilnost za digonalizabilne matrike.

Konvergenca metode je linearna in odvisna od razmerja $|\lambda_2/\lambda_1|$. Če je z približek za lastni vektor, lahko približek za pripadajočo lastno vrednost dobimo z Rayleighovim kvocientom

$$\lambda = \frac{z^H A z}{z^H z} = \rho(z, A).$$

Pri tem dejansko rešimo predoločen sistem $Az = \lambda z$, kjer je z matrika, λ spremenljivka in Az desna stran. Za kvocient veljata naslednji lastnosti:

- $\rho(\alpha z, A) = \rho(z, A)$ za $\alpha \neq 0$
- $\rho(x_i, A) = \lambda_i$

To nam poda ustavitveni kriterij za algoritem; izračunamo $\rho(z_k, A)$ in primerjamo $\|Az_k - \rho(z_k, A)z_k\| < \varepsilon$.

Vprašanje 52. Kaj je Rayleighov kvocient?

Denimo, da smo našli λ_1 in x_1. Kako sedaj poiščemo λ_2 in x_2? Podobno kot pri dokazu obstoja Schurove forme poiščemo unitarno matriko U_1, da je $U_1 e_1 = x$ (npr. s Householderjevim zrcaljenjem). Vzamemo

$$B = U^H A U = \begin{bmatrix} \lambda_1 & \alpha^T \\ 0 & C \end{bmatrix},$$

in nadaljujemo s potenčno metodo na C, ki pa je seveda ne izrazimo eksplicitno. Množenje izvedemo tako, da množimo z B, in iz produkta vzamemo spodnjih $n-1$ elementov.

Vprašanje 53. Kako s potenčno metodo poiščeš vse lastne vrednosti?

3.6.2 Inverzna iteracija

Naj bo σ zelo dober približek za eno izmed lastnih vrednosti matrike A. Matrika $(A - \sigma I)^H$ ima lastne vrednosti $\frac{1}{\lambda_i - \sigma}$, in velja

$$\frac{1}{|\lambda_j - \sigma|} \gg \frac{1}{|\lambda_i - \sigma|},$$

kjer je λ_j tista lastna vrednost, za katero je σ dober približek. To dejstvo izrabimo v algoritmu 9.

Algorithm 9 Inverzna iteracija

Izberi naključen $z_0 \neq 0$.
for $k = 1, 2, \ldots$ **do**
  Reši $z_{k-1} = (A - \sigma I)y_k$
  $z_k = {}^{y_k}/_{\|y_k\|}$
end for

Vprašanje 54. Opiši delovanje inverzne iteracije.

3.6.3 Ortogonalna iteracija

Definicija. Prostor N je INVARIANTNI PODPROSTOR za matriko A, če je za vsak $x \in N$ tudi $Ax \in N$.

Izrek. *Naj bo $S = [S_1 S_2]$ nesingularna matrika. Naj bo $B = S^{-1}AS$ oblike*

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Potem stolpci S_1 razpenjajo invariantni podprostor za A natanko tedaj, ko je $B_{21} = 0$.

Dokaz. Ker je $SB = AS$, velja $S_1 B_{11} + S_2 B_{21} = AS_1$. □

Recimo, da je res $B_{21} = 0$. Naj za lastne vrednosti matrike A velja $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_p| > |\lambda_{p+1}| \geq \ldots$, kjer je p število stolpcev v S_1. V tem primeru je invariantni podprostor velikosti p, ki mu pripadajo največje lastne vrednosti, dominanten.

Algorithm 10 Ortogonalna iteracija

Izberi naključno matriko $Z_0 \in \mathbb{R}^{n \times p}$.
for $k = 1, 2, \ldots$ **do**
  $Y_k = AZ_{k-1}$
  Za Z_k vzemi Q_k iz QR razcepa $Y_k = Q_k S_k$.
end for

Izrek. *Če za lastne vrednosti A velja $|\lambda_1| \geq \ldots \geq |\lambda_p| > |\lambda_{p+1}|$, potem za naključno izbrano matriko $Z_0 \in \mathbb{R}^{n-p}$ matrika Z_k konvergira proti ortonormirani bazi za dominantni invariantni podprostor dimenzije p.*

Dokaz. Dokažemo samo za primer, kjer lahko matriko A diagonaliziramo v $A = XDX^{-1}$. Označimo $D = \operatorname{diag}(D_1, D_2)$ in $X = [X_1, X_2]$, kjer sta D_1 in X_1 dimenzije p. Ker je $|\lambda_p| > 0$, velja $\det D_1 \neq 0$. Če izrazimo Z_0 v bazi lastnih vektorjev kot

$$Z_0 = X \begin{bmatrix} W_1 \\ W_2 \end{bmatrix},$$

potem lahko za naključen Z_0 predpostavimo, da je W_1 nesingularna. Pokazati moramo, da $\operatorname{im} Z_k \to \operatorname{im} X_1$ za $k \to \infty$. Velja

$$\operatorname{im} Z_k = \operatorname{im} AZ_{k-1} = \cdots = \operatorname{im} A^k Z_0 = \operatorname{im} XD^k X^{-1} X \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} = \operatorname{im} X \begin{bmatrix} D_1^k W_1 \\ D_2^k W_2 \end{bmatrix}$$

Ker je W_1 obrnljiva, je to enako

$$\operatorname{im} Z_k = \operatorname{im} X \begin{bmatrix} I \\ D_2^k W_2 W_1^{-1} D_1^{-k} \end{bmatrix} D_1^k W_1 = \operatorname{im} X \begin{bmatrix} I \\ D_2^k W_2 W_1^{-1} D_1^{-k} \end{bmatrix}.$$

Elemente spodnje matrike lahko razpišemo kot

$$\left[D_2^k W_2 W_1^{-1} D_1^{-k} \right]_{ij} = \left[W_2 W_1^{-1} \right]_{ij} \frac{\lambda_{p+i}^k}{\lambda_j^k}$$

Ker so vsi elementi D_1 večji od vseh v D_2, ulomek na desni konvergira k 0 za $k \to \infty$, torej $\operatorname{im} Z_k \to \operatorname{im} X_1$. □

Vprašanje 55. Pod katerim pogojem deluje ortogonalna iteracija? Dokaži za diagonalizabilne matrike.

Posledica. *Če za matriko A velja $|\lambda_1| > \cdots > |\lambda_n|$, potem za naključno izbrano matriko Z_0 matrika $Z_k^T A Z_k$ konvergira proti Schurovi formi.*

Dokaz. Vzemimo $p \in \{1, \ldots, n\}$ in razcepimo $Z_k = [Z_{k1} Z_{k2}]$. Po izreku Z_{k1} konvergira proti ONB za dominantni invariantni podprostor dimenzije p. Matrika $Z_{k2}^T A Z_{k1}$ torej konvergira k 0 (po izreku od prej), to pa velja za vsak p, torej smo dobili zgornje trikotno matriko, kjer so elementi na diagonali lastne vrednosti, urejene po absolutni vrednosti. □

Vprašanje 56. Kako z ortogonalno iteracijo poiščeš Schurovo formo? Dokaži.

3.6.4 QR iteracija

Algorithm 11 QR iteracija

$A_0 = A$
for $k = 0, 1, \ldots$ **do**
    Izračunaj QR razcep $A_k = Q_k R_k$
    $A_{k+1} = R_k Q_k$
end for

Če za lastne vrednosti velja $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n|$, potem A_k konvergira k Schurovi formi. Velja $A_{k+1} = Q_k^T A_k Q_k$, torej so si vse A_k ortogonalno podobne.

Izrek. *Za matriko A_k iz QR iteracije velja $A_k = Z_k^T A Z_k$, kjer je Z_k matrika iz ortogonalne iteracije z začetkom $Z_0 = I$.*

Dokaz. Dokažemo z indukcijo na k, kjer je baza indukcije trivialna. Če velja za k, računamo $AZ_k = Z_{k+1}S_{k+1}$, zato $Z_k^T A Z_k = Z_k^T Z_{k+1} S_{k+1}$. Matrika $Z_k^T Z_{k+1}$ je ortogonalna, S_{k+1} pa zgornje trikotna, torej je to QR razcep matrike $Z_k^T A Z_k = A_k$ po indukcijski predpostavki. Velja torej

$$A_{k+1} = R_k Q_k = S_{k+1} Z_k^T Z_{k+1} = Z_{k+1}^T A Z_k Z_k^T Z_{k+1} = Z_{k+1}^T A Z_{k+1}.$$

□

Vprašanje 57. Zapiši algoritem QR iteracije. Kako deluje? Dokaži.

Definicija. Matrika H je ZGORNJE HESSENBERGOVA, če je $h_{ij} = 0$ za $i > j + 1$.

Lema. *Če je matrika A zgornje Hessenbergova, se oblika med QR iteracijo ohranja.*

Dokaz. Če je A zgornja Hessenbergova, je Q zgornja Hessenbergova. □

Algorithm 12 Redukcija matrike v zgornje Hessenbergovo obliko

```
Q = I
for k = 1, ..., n − 2 do
    Določi w_k ∈ ℝ^{n−k} za Householderjevo zrcaljenje, ki slika A(k + 1 : n, k) v ±xe_1
    A(k + 1 : n, k : n) = P_k A(k + 1 : n, k : n)
    A(:, k + 1 : n) = A(:, k + 1 : n)P_k
    Q(:, k + 1 : n) = Q(:, k + 1 : n)P_k
end for
```

Lema nam pove, da si lahko prihranimo veliko računanja med iteracijo. Prvo lahko s Householderjevimi zrcaljenji pretvorimo A v zgornjo Hessenbergovo obliko, kot v algoritmu 12 Izračun QR razcepa take matrike lahko naredimo z $n - 1$ Givensonovimi rotacijami, torej v $O(n^2)$. Če je R_{ij} rotacija, velja

$$R_{n-1,n}^T \dots R_{23}^T R_{12}^T A_k = R_k$$

oziroma

$$A_{k+1} = R_k Q_k = R_k R_{12} R_{23} \dots R_{n-1,n}.$$

Vprašanje 58. Kako optimiziraš korak QR iteracije na $O(n^2)$?

Definicija. Zgornje Hessenbergova matrika H je NERAZCEPNA, če so vsi elementi v spodnji diagonali neničelni. V nasprotnem primeru je RAZCEPNA.

Algorithm 13 QR iteracija z enojnim premikom

Naredi redukcijo na Hessenbergovo obliko $A_0 = Q^T AQ$
for $k = 0, 1, \ldots$ **do**
  Izberi premik σ_k
  Naredi QR razcep $A_k - \sigma_k I = Q_k R_k$
  $A_{k+1} = R_k Q_k + \sigma_k I$
end for

Če je matrika razcepna, lahko problem razdelimo na dva manjša. Predpostavimo lahko torej, da je začetna Hessenbergova matrika nerazcepna. Pri numeričnem reševanju element $h_{i+1,i}$ proglasimo za 0, če je $|h_{i+1,i}| \leq \varepsilon(|h_{i,i}| + |h_{i+1,i+1}|)$.

Število iteracij lahko zmanjšamo z uporabo premikov, kot je prikazano v algoritmu 13.

Matriki A_k in A_{k+1} sta še vedno ortogonalno podobni, ker velja

$$A_{k+1} = R_k Q_k + \sigma_k I = Q_k^T (A_k - \sigma_k I) Q_k + \sigma_k I = Q_k^T A_k Q_k.$$

Lema. *Če je A nerazcepna zgornje Hessenbergova matrika in za premik σ izberemo lastno vrednost A, potem za matriko $B = RQ + \sigma I$, kjer je $A - \sigma I = QR$, velja $b_{n,n-1} = 0$ in $b_{n,n} = \sigma$.*

Dokaz. Matrika $A - \sigma I$ je singularna, zaradi nerazcepnosti pa je prvih $n-1$ stolpcev linearno nedovisnih, torej mora veljati $r_{n,n} = 0$. Torej je zadnja vrstica R enaka 0, in je B predpisane oblike. □

Vprašanje 59. Kako deluje QR iteracija s premikom? Kaj je njena prednost? Dokaži.

Idealno je za premik torej izbrati čim boljši približek za lastno vrednost. Poznamo dve pogosti izbiri premika;

- Enojni premik: izberemo $\sigma_k = (A_k)_{nn}$. Deluje dobro za matrike s samimi realnimi lastnimi vrednostmi.
- Dvojni premik: za σ_{k1} in σ_{k2} izberemo lastni vrednosti matrike $A(n-1:n, n-1:n)$, in v okviru ene iteracije naredimo dva premika

$$\begin{aligned} A_k - \sigma_{k1} I &= Q_k R_k & \qquad A_{k+1/2} - \sigma_{k2} &= \tilde{Q}_k \tilde{R}_k \\ A_{k+1/2} &= R_k Q_k + \sigma_{k1} I & \qquad A_{k+1} &= \tilde{R}_k \tilde{Q}_k + \sigma_{k2} I \end{aligned}$$

  Pokažemo lahko, da je za realno A_k tudi A_{k+1} realna.

Vprašanje 60. Kako izberemo premik za QR iteracijo?

3.7 Polinomska interpolacija

3.7.1 Lagrangeova oblika

Problem z interpolacijo v standardni bazi je, da je Vandermondova matrika zelo občutljiva.

Izrek. *Za paroma različne točke $x_0, \dots, x_n$ in vrednosti $y_0, \dots, y_n$ obstaja natanko en polinom stopnje $\leq n$, da je $p(x_i) = y_i$.*

Dokaz. Obstoj dokažemo s konstrukcijo, kjer uporabimo Lagrangeove bazne polinome

$$l_{n,i}(x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Če take polinome lahko konstruiramo (kjer je $l_{n.i}$ stopnje $\leq n$), bo veljalo

$$p = \sum_i y_i l_{n,i}.$$

Če definiramo

$$l_{n,i}(x) = \frac{(x - x_0) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)},$$

potem $l_{0,n}, \dots, l_{n,n}$ sestavljajo bazo za polinome stopnje $\leq n$.

Dokazati moramo še enoličnost: denimo, da je $\tilde{p}$ tudi polinom stopnje $\leq n$, ki zadošča $\tilde{p}(x_i) = y_i$. Tedaj je $q = p - \tilde{p}$ polinom stopnje $\leq n$, za katerega velja $q(x_i) = 0$. Polinom stopnje $\leq n$, ki ima vsaj $n + 1$ ničel, je enak $q = 0$. □

Vprašanje 61. Kako interpoliraš polinom z Lagrangeovo bazo? Dokaži, da je interpolacija enolična.

Izrek. *Če je f $(n + 1)$-krat zvezno odvedljiva na $[a, b]$, ki vsebuje paroma različne vozle $x_0, \dots, x_n$ ter x, in je p interpolacijski polinom za f, potem velja*

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n + 1)!} \omega(x)$$

za $\omega(x) = (x - x_0) \dots (x - x_n)$ in $\min(x_0, \dots, x_n, x) \leq \xi \leq \max(x_0, \dots, x_n, x)$.

Dokaz. Predpostavimo lahko, da je x različen od $x_0, \dots, x_n$. Definiramo $g(z) = f(z) - p(z) - C\omega(z)$, kjer nastavimo C tako, da ima g pri x ničlo. Poleg tega ima g tudi ničle v vozlih $x_0, \dots, x_n$. Funkcija g je $(n + 1)$-krat zvezno odvedljiva in ima $n + 2$ različnih ničel. Po Rollovem izreku ima g' vsaj $n + 1$ ničel. Če postopek nadaljujemo, dobimo, da ima $g^{(n+1)}$ ničlo ξ. Če upoštevamo, da sta p in ω polinoma, dobimo

$$C = \frac{f^{(n+1)}(\xi)}{(n + 1)!}.$$

□

Vprašanje 62. Kako izraziš napako interpolacije? Dokaži.

3.7.2 Deljene diference

Definicija. Za paroma različne točke $x_0, \dots, x_k$ je DELJENA DIFERENCA $[x_0, \dots, x_k]_f$ vodilni koeficient (pri x^k) interpolacijskega polinoma za f v točkah $x_0, \dots, x_k$.

Izrek. *Za paroma različne točke $x_0, \dots, x_n$ lahko interpolacijski polinom za f na $x_0, \dots, x_n$ zapišemo kot*

$$p(x) = [x_0]_f + [x_0, x_1]_f(x - x_0) + \dots + [x_0, \dots, x_n]_f(x - x_0) \dots (x - x_{n-1}).$$

Dokaz. Dokažemo z indukcijo na n. Baza indukcije je očitna, dokažimo korak. Naj bo $p_n(x)$ polinom stopnje $\leq n$, ki se z f ujema na $x_0, \dots, x_n$. Dodamo še točko x_{n+1} in iščemo p_{n+1}. Za

$$p_{n+1}(x) = p_n(x) + C(x - x_0) \dots (x - x_n)$$

in ustrezen C velja $p_{n+1}(x_i) = f(x_i)$. Izračunamo torej

$$C = \frac{f(x_{n+1}) - p_n(x_{n+1})}{(x_{n+1} - x_0) \dots (x_{n+1} - x_n)}.$$

□

Vprašanje 63. Kakšna je Newtonova oblika interpolacije? Dokaži.

Izrek. *Za deljene diference velja*

- $[x_0, \dots, x_n]_f$ *je simetrična funkcija argumentov*
- $[x_0, \dots, x_n]_f$ *je linearen funkcional v* f
- *Velja rekurzivna formula*

$$[x_0, \dots, x_k]_f = \frac{[x_1, \dots, x_k]_f - [x_0, \dots, x_{k-1}]_f}{x_k - x_0}.$$

Dokaz. Prvi točki sta očitni. Za tretjo: Naj p_a interpolira f v točkah $x_0, \dots, x_{k-1}$, in p_b v točkah $x_1, \dots, x_k$. Hitro lahko preverimo, da je ustrezen interpolacijski polinom

$$p(x) = \frac{x - x_k}{x_0 - x_k} p_a(x) + \frac{x - x_0}{x_k - x_0} p_b(x),$$

ker je stopnja p manjša ali enaka k in $p(x_0) = p_a(x_0) = f(x_0)$ ter $p(x_k) = p_b(x_k) = f(x_k)$. Za ostale točke tudi velja $p(x_i) = f(x_i)$. □

Vprašanje 64. Povej in dokaži rekurzivno formulo za deljene diference.

Formula nam pove, da je smiselna definicija, če se točke x_i ponavljajo, naslednja:

$$[x_0, \dots, x_k]_f = \begin{cases} \frac{f^{(k)}(x_0)}{k!} & x_0 = x_1 = \cdots = x_k \\ \frac{[x_1, \dots, x_k]_f - [x_0, \dots, x_{k-1}]_f}{x_k - x_0} & \text{sicer} \end{cases}$$

kjer pri drugi definiciji poskrbimo, da $x_0 \neq x_k$ (vrstni red točk je nepomemben).

Izrek. *Za k-krat zvezno odvedljivo f velja*

$$[x_0, \dots, x_k]_f = \int_0^1 dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{k-1}} f^{(k)}(\xi_k) dt_k,$$

kjer je $\xi_k = t_k(x_k - x_{k-1}) + \cdots + t_1(x_1 - x_0) + x_0$.

Posledica. *Za k-krat zvezno odvedljivo f velja*

$$[x_0, \dots, x_k]_f = \frac{f^{(k)}(\xi)}{k!},$$

kjer je $\min(x_0, \dots, x_k) \leq \xi \leq \max(x_0, \dots, x_k)$.

Dokaz. Uporabimo zadnji izrek in izrek o povprečni vrednosti. Upoštevamo, da je volumen simpleksa enak $k!^{-1}$. □

Izrek. *Če je p interpolacijski polinom za f v točkah $x_0, \dots, x_n$, potem velja*

$$f(x) - p(x) = [x_0, \dots, x_n, x]_f (x - x_0) \dots (x - x_n).$$

Dokaz. Če desni strani prištejemo $p(x)$, dobimo interpolacijski polinom za f točkah $x_0, \dots, x_n$ in x. □

Posledica tega je, da je ocena za napako enaka kot prej, tudi če uporabljamo ponovljene točke.

Vprašanje 65. Kako izraziš oceno za napako interpolacije z deljenimi diferencami?

3.8 Numerično integriranje

Želimo izračunati integral

$$I(f) = \int_a^b f(x) dx.$$

Ideja je, da funkcijo aproksimiramo z interpolacijskim polinomom

$$\int_a^b f(x) dx = \int_a^b \sum_{i=0}^n f(x_i) l_{n,i}(x) dx + R(f) = \sum_{i=0}^n f(x_i) \int_a^b l_{n,i}(x) dx + R(f),$$

kjer integrale interpolacijskih polinomov imenujemo UTEŽI ali KOEFICIENTI α_i, napaka $R(f)$ pa je oblike

$$R(f) = \int_a^b \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega(x)dx.$$

Tako dobljene kvadraturne formule so določene z izbiro vozlov. Napaka pri računanju se razdeli na dve komponenti; napaka metode $R(f)$ ter neodstranljiva napaka, ki jo dobimo, ker ne poznamo točnih vrednosti f v vozlih.

Vprašanje 66. Izpelji obliko kvadraturnih formul.

3.8.1 Newton-Colesove formule

Pri NC formulah vozle izberemo enakomerno, $x_i = a + ih$. Ločimo zaprti tip formul, kjer uporabimo vse vozle, in odprti tip, kjer izpustimo vozla v krajiščih.

Najenostavnejša kvadraturna formula je TRAPEZNA FORMULA, t.j. formula zaprtega tipa za $n = 1$

$$\int_{x_0}^{x_1} f(x)dx = \frac{h}{2}(f(x_0) + f(x1)) + \int_{x_0}^{x_1} \frac{1}{2}f''(\xi_x)(x - x_0)(x - x_1)dx,$$

kjer za napako po izreku o povprečni vrednosti velja

$$R(f) = \frac{1}{2}f''(\xi)\int_{x_0}^{x_1}(x - x_0)(x - x_1)dx = -\frac{h^3}{12}f''(\xi).$$

Vprašanje 67. Opiši trapezno formulo.

Če namesto tega vzamemo $n = 2$, dobimo SIMPSONOVO FORMULO

$$\int_{x_0}^{x_2} f(x)dx = \frac{h}{3}(f(x_0) + 4f(x_1) + f(x_2)) + R(f)$$

za napako

$$R(f) = \int_{x_0}^{x_2} \frac{1}{6}f'''(\xi_x)(x - x_0)(x - x_1)(x - x_2)dx.$$

Če je f polinom stopnje 3, je f''' konstantna, in velja $R(f) = 0$. Podobno se zgodi tudi pri vseh NC formulah za sode n. Oblika napake za vse formule je vedno $R(f) = Cf^{(m)}(\xi)$, kjer je m stopnja najnižjega polinoma, za katerega formula ni točna. Za Simpsonovo formulo je to $m = 4$, če vstavimo $f(x) = (x - x_0)^4$ dobimo $C = -h^5/90$.

Vprašanje 68. Povej Simpsonovo formulo. Za katero stopnjo polinomov je točna? Izpelji predpis za napako.

Druga vrsta so Newton-Colesove formule zaprtega tipa, ki so smiselne za $n \geq 2$. Pri $n = 2$ dobimo SREDINSKO FORMULO

$$\int_{x_0}^{x_2} f(x)dx = 2hf(x_1) + \frac{h^3}{3}f''(\xi),$$

za $n = 4$ pa MILNEOVO FORMULO

$$\int_{x_0}^{x_4} f(x)dx = \frac{4h}{3}(2f(x_1) - f(x_2) + 2f(x_3)) + \frac{28h^5}{90} f^{(4)}(\xi),$$

ki NC formula z najmanj vozli in negativno utežjo.

Vprašanje 69. Kaj sta sredinska in Milneova formula? Zakaj je druga omembe vredna?

3.8.2 Napake pri numeričnem integriranju

Velja

$$\int_a^b f(x)dx = \sum_{i=0}^n \alpha_i f(x_i) + D_m,$$

kjer za napako metode D_m ne bo veljalo nujno $D_m \to 0$ za $n \to \infty$. Če npr. integriramo $f(x) = (1 + x^2)^{-1}$ na $[-5, 5]$ z enakomerno razporejenimi točkami, bo veljalo

$$\max_{x \in [-5,5]} |f(x) - p_n(x)| \to \infty,$$

kjer je p_n interpoliran polinom za n točk. Pri izračunu vsote se pojavita še neodstranljiva D_n in zaokrožitvena napaka. Recimo, da velja $|f(x_i) - f_i| \leq \varepsilon$, kjer je f_i izračunan približek in $f(x_i)$ točna vrednost. Potem lahko ocenimo

$$|D_n| = \left| \sum_{i=0}^n \alpha_i f(x_i) - \sum_{i=0}^n \alpha_i f_i \right| \leq \sum_{i=0}^n |\alpha_i| \, \varepsilon.$$

Če je $\alpha_i \geq 0$ za vse i, lahko to nadalje ocenimo z $|D_n| \leq (b - a)\varepsilon$, ker velja

$$\sum_{i=0}^n \alpha_i = \sum_{i=0}^n \int_a^b l_{i,n}(x)dx = \int_a^b 1dx = b - a.$$

Če to ne velja, pa imamo težave. Pri NC formulah vsota absolutnih vrednosti α_i hitro divergira za $n \to \infty$.

Vprašanje 70. Analiziraj neodstranljivo napako pri numeričnem integriranju.

Temu problemu se lahko ognemo tako, da razdelimo interval na manjše kose in integriramo na vsakem posebej. Temu pravimo SESTAVLJENO PRAVILO, sedaj pa imamo težavo, da moramo računati veliko vrednosti, tudi če tega ne potrebujemo na celotni domeni. Rešitev so adaptivne metode.

Oglejmo si adaptivno Simpsonovo metodo. Velja $I(f) = S_h(f) + R_h(f) = S_{h/2}(f) + R_{h/2}(f)$, s konkretnima napakama

$$R_h(f) = \frac{h^4(b-a)}{180} f^{(4)}(\xi_1), \qquad R_{h/2}(f) = \frac{h^4(b-a)}{16 \cdot 180} f^{(4)}(\xi_2).$$

Pri predpostavki, da je $f^{(4)}(\xi_1) \approx f^{(4)}(\xi_2)$, dobimo

$$R_h(f) \approx 16 R_{h/2}(f),$$

iz česar izpeljemo oceno za napako

$$R_{h/2}(f) \approx \frac{S_{h/2}(f) - S_h(f)}{15}.$$

Potem lahko dobimo ekstrapoliran približek

$$I(f) = S_{h/2}(f) + R_{h/2}(f) \approx \frac{16 S_{h/2}(f) - S_h(f)}{15}.$$

Če želimo izračunati integral funkcije f med a in b, po (sestavljeni) Simpsonovi formuli izračunamo S_h in $S_{h/2}$ ter preverimo, če je ocena za napako manjša od ε. Če je, končamo, sicer pa interval razdelimo na dva, in rekurzivno izračunamo integrala f na teh intervalih, pri čemer zahtevamo, da je napaka manjša od $\varepsilon/2$. Postopek se bo končal, ker se mera za napako zmanjšuje s faktorjem 16, zahtevana natančnost pa s faktorjem 2.

Vprašanje 71. Razloži adaptivno Simpsonovo metodo.

3.8.3 Gaussove kvadraturne formule

Imejmo

$$I(f) = \int_a^b f(x)\rho(x)dx,$$

kjer je ρ nenegativna utež. Kvadraturna formula ima obliko

$$I(f) = \sum_{i=0}^{n} \alpha_i f(x_i) + R(f),$$

kjer je

$$\alpha_i = \int_a^b l_{i,n}(x)\rho(x)dx.$$

Formula je vedno točna za polinome stopnje $\leq n$ za poljubno izbiro vozlov $x_0, \ldots, x_n$. Če definiramo skalarni produkt

$$\langle f, g \rangle = \int_a^b f(x)g(x)\rho(x)dx$$

in uporabimo Gram-Schmittovo ortogonalizacijo na standardni polinomski bazi, dobimo polinome $\varphi_0, \varphi_1, \ldots$, za katere velja st $\varphi_i = i$ in $\langle \varphi_i, \varphi_j \rangle = \delta_{ij}$.

Izrek. *Če so $\varphi_0, \varphi_1, \ldots$ ortogonalni polinomi na $[a, b]$ z utežjo ρ, ima φ_k same realne enostavne ničle, ki vse ležijo v (a, b).*

Dokaz. Naj ima φ_k v (a,b) l različnih ničel, kjer je $l < k$. Označimo jih z $z_1, \ldots, z_l$, in definiramo

$$g(x) = (x - z_1)^{j_1} \cdots (x - z_l)^{j_l},$$

kjer je j_i enak 1, če je z_i liha ničla, in 0 sicer. Ker je integrand pozitiven, velja

$$\int_a^b g(x)\varphi_k(x)\rho(x)dx > 0.$$

Polinom g je manjše stopnje kot k, ker pa je φ_k pravokoten na vse polinome stopnje manjše od k, bi moral biti integral enak 0. To je protislovje. □

Če za vozle vzamemo ničle polinoma φ_{n+1}, bo veljalo $\omega(x) = c\varphi_{n+1}(x)$, torej je ω pravokoten na vse polinome stopnje $\leq n$. Naj bo f polinom stopnje $\leq 2n+1$. Zapišemo ga lahko kot $f(x) = h(x)\omega(x) + g(x)$, kjer sta stopnji h in g manjši od n. Računamo

$$\int_a^b f(x)\rho(x)dx = \underbrace{\int_a^b h(x)\omega(x)\rho(x)dx}_{\langle h,\omega\rangle=0} + \int_a^b g(x)\rho(x)dx = \sum_{i=0}^n \alpha_i g(x_i) = \sum_{i=0}^n \alpha_i f(x_i).$$

Formula je natančna za vse polinome stopnje $\leq n$, torej tudi za g; sledi, da je natančna tudi za f. Izkaže se tudi, da so vse uteži pri Gaussovih formulah pozitivne, tako da ni problemov s stabilnostjo.

Vprašanje 72. Povej idejo za Gaussovimi kvadraturnimi formulami in dokaži, da so natančne za polinome stopnje $\leq 2n+1$.

3.9 Diferencialne enačbe

Numerična rešitev diferencialne enačbe je sestavljena iz zaporedja $x_0, x_1, \ldots$ in pripadajočih približkov $y_0, y_1, \ldots$. Metode delimo na enokoračne, kjer y_{n+1} izračunamo iz y_n, ter večkoračne, kjer y_{n+1} izračunamo iz prejšnjih nekaj približkov. Poleg tega ločimo metode na eksplicitne, kjer imamo formulo za y_{n+1}, in implicitne, kjer y_{n+1} izračunamo z reševanjem nelinearne enačbe.

Če je $f = f(x,y)$ Lipschitzova v y s konstanto L, in je $y(x)$ točna rešitev začetnega problema $y' = f(x,y), y(x_0) = y_0$, ter $\tilde{y}(x)$ rešitev začetnega problema z zmotenim začetnim pogojem $y(x_0) = \tilde{y}(x_0)$, potem za poljuben $x \geq x_0$ velja

$$|\tilde{y}(x) - y(x)| \leq e^{L(x-x_0)} |\tilde{y}_0 - y_0| .$$

Podobno slabo mejo dobimo tudi, če zmotimo še f.

Najenostavnejša eksplicitna metoda je EKSPLICITNA EULERJEVA METODA

$$\begin{aligned} y_{n+1} &= y_n + hf(x_n, y_n), \\ x_{n+1} &= x_n + h. \end{aligned}$$

Poznamo tudi implicitno obliko

$$y_{n+1} = y_n + hf(x_{n+1}, y_{n+1}).$$

3.9.1 Runge-Kutta metode

Pri Runge-Kutta metodah najprej izračunamo koeficiente

$$k_i = hf\left(x_n + \alpha_i h, y_n + \sum_{j=1}^{m} \beta_{ij} k_j\right)$$

za $i = 1, \ldots, m$, nato pa

$$y_{n+1} = y_n + \sum_{i=1}^{m} \gamma_i k_i.$$

Pri tem je m stopnja metode, konstante $\alpha_i, \beta_{ij}, \gamma_i$ pa določimo tako, da se y_{n+1} čim bolj ujema z razvojem $y(x_n + h)$ v Taylorjevo vrsto. Veljati mora

$$\alpha_i = \sum_{j=1}^{m} \beta_{ij}.$$

Metoda je eksplicitna, če za $i \leq j$ velja $\beta_{ij} = 0$, sicer pa je implicitna. Stopnja metode m je različna od REDA metode k, ki je enak eksponentu v zadnjem členu Taylorjeve vrste, s katerim se metoda natančno ujema.

Vprašanje 73. Razloži Runge-Kutta metode za reševanje diferencialnih enačb.

4 Verjetnost

Komentar za učenje: poglej si tudi vserazne primere v zvezku, in jih poračunaj za vajo.

4.1 Izidi, dogodki, verjetnosti

Vprašanje 1. Kaj je množica Ω vseh možnih izidov? Povej nekaj primerov.

Odgovor: To je množica, ki hrani vse možne rezultate nekega poskusa. Pri mešanju kupa n kart velja $\Omega = S_n$, pri n-kratnem metu kovanca je to $\Omega = \{G, S\}^n$, itd. ⊠

Definicija. Družina $\mathcal{F}$ podmnožic množice Ω je σ-ALGEBRA, če velja:

- $\Omega \in \mathcal{F}$,
- $A \in \mathcal{F} \implies A^{\mathsf{c}} \in \mathcal{F}$,
- $A_1, A_2, \ldots \in \mathcal{F} \implies \bigcup_i A_i \in \mathcal{F}$.

Definicija. Naj bo Ω množica možnih izidov, in $\mathcal{F}$ σ-algebra nad Ω. VERJETNOST je preslikava $P : \mathcal{F} \to [0,1]$, za katero velja $P(\Omega) = 1$, in kjer za disjunktne dogodke $A_1, A_2, \ldots \in \mathcal{F}$ velja $P(\bigcup_i A_i) = \sum_i P(A_i)$.

Opomba. To sta *aksioma Kolmogorova.*

Vprašanje 2. Kaj je verjetnost?

Izrek (Formula za vključitve in izključitve)**.** *Naj bodo $A_1, \ldots, A_n$ dogodki. Potem velja*

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i \cap A_j) + \ldots + (-1)^{n-1} P(\bigcap_{i=1}^{n} A_i).$$

Dokaz. Definirajmo dogodke

$$B_r = \{\omega \in \Omega \,|\, \omega \text{ je vsebovan v natanko } r \text{ množicah } A_i\}.$$

To so disjunktni dogodki, za katere velja $\bigcup_i A_i = \bigcup_r B_r$. Sledi

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{r=1}^{n} P(B_r).$$

Poglejmo si, kolikokrat smo v formuli v izreku šteli vsako izmed množic B_r. Ta množica je vsebovana v preseku do r dogodkov, torej se v prvem členu pojavi r-krat, v drugem $\binom{r}{2}$, v tretjem $\binom{r}{3}$, itd. Vsota je tedaj

$$\binom{r}{1} - \binom{r}{2} + \binom{r}{3} - \ldots + (-1)^r \binom{r}{r} = 1,$$

kar lahko izpeljemo iz razvoja izraza $0 = (1-1)^r$. □

Vprašanje 3. Povej formulo za izključitve in izključitve. Kaj je ideja dokaza?

Lema. *Naj bodo $A_1, A_2, \ldots$ dogodki. Če je $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$, je verjetnost unije*

$$P(\bigcup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n).$$

Če namesto tega velja $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$, je

$$P(\bigcap_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n).$$

Dokaz. Druga formula sledi iz De Morganovih pravil, dokažemo samo prvo. Zapišemo

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup (A_2 \backslash A_1) \cup (A_3 \backslash (A_1 \cup A_2)) \cup \ldots$$

To so disjunktni dogodki, torej zanje velja

$$\begin{aligned}
P(\bigcup_{i=1}^{\infty} A_i) &= P(A_1) + \sum_{k=2}^{\infty} P(A_k \backslash (A_1 \cup \ldots \cup A_{k-1})) \\
&= \lim_{n\to\infty} (P(A_1) + \sum_{k=2}^{n} P(A_k \backslash (A_1 \cup \ldots \cup A_{k-1}))) \\
&= \lim_{n\to\infty} P(\bigcup_{k=1}^{n} A_k \backslash (A_1 \cup \ldots \cup A_{k-1})) \\
&= \lim_{n\to\infty} P(A_n).
\end{aligned}$$

□

Lema (Prva Borel-Cantorjeva lema)**.** *Naj bodo $A_1, A_2, \ldots$ dogodki, za katere velja $\sum_i P(A_i) < \infty$. Definiramo $\overline{A} = \{\omega \in \Omega \mid \omega$ je vsebovan v neskončno mnogo A_k $\}$. Tedaj velja $P(\overline{A}) = 0$.*

Dokaz. Prepričamo se lahko, da velja $\overline{A} = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$. Te unije so padajoče za $n \to \infty$, zatorej po prejšnji lemi velja

$$P\left(\overline{A}\right) = \lim_{n\to\infty} P\left(\bigcup_{m=n}^{\infty} A_m\right).$$

Iz dokaza prešnje leme vidimo, da velja sklep

$$P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P(A_k) \implies P\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} P(A_k).$$

Torej velja

$$P\left(\overline{A}\right) \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k).$$

Izraz na desni pa je rep konvergenčne vrste, torej je limita enaka 0. □

Vprašanje 4. Povej in dokaži prvo Borel-Cantorjevo lemo.

4.1.1 Pogojna verjetnost in neodvisnost

Definicija. Naj bo B dogodek s $P(B) > 0$. POGOJNA VERJETNOST dogodka A glede na B je

$$P(A \,|\, B) = \frac{P(A \cap B)}{P(B)}.$$

Vprašanje 5. Kaj je pogojna verjetnost?

Primer (Bertrandov paradoks). Imamo tri škatle. V prvi sta dva zlatnika, v drugi zlatnik in srebrnik, in v zadnji dva srebrnika. Izberemo eno škatlo tako, da ima vsaka verjetnost $1/3$. Iz izbrane škatle tedaj naključno izberemo kovanec. Definiramo dogodka A, drugi kovanec v škatli je zlatnik, in B, izbrani kovanec je zlatnik. Z izpisom izidov izračunamo $P(A \,|\, B) = 2/3$.

Definicija. Družina dogodkov $\{H_1, \ldots, H_n, \ldots\}$ je PARTICIJA Ω, če je njihova unija enaka Ω in če so paroma disjunktni.

Vprašanje 6. Kaj je particija? Izpelji formulo za popolno verjetnost.

Odgovor: Za definicijo glej zgoraj. Naj bo A dogodek. Računamo

$$\begin{aligned}
P(A) &= P(A \cap \Omega) \\
&= P(A \cap \bigcup_i H_i) \\
&= P(\bigcup_i A \cap H_i) \\
&= \sum_i P(A \cap H_i) \\
&= \sum_i \frac{P(A \cap H_i)}{P(H_i)} P(H_i) \\
&= \sum_i P(A \,|\, H_i) P(H_i).
\end{aligned}$$

Če je $P(H_i) = 0$, lahko člen izpustimo. ⊠

4.1.2 Neodvisnost dogodkov

Definicija. Dogodki $\{A_i\}_{i\in I}$ so NEODVISNI, če za vsako končno poddružino $A_1, A_2, \ldots, A_n$ velja

$$P(A_1 \cap \ldots \cap A_n) = P(A_1) \ldots P(A_n).$$

Vprašanje 7. Kdaj so dogodki neodvisni?

Definicija. Družina dogodkov $\mathcal{P} = \{A_1, \ldots, A_n\}$ je π-SISTEM, če za vsaka $A_i, A_j \in \mathcal{P}$ velja $A_i \cap A_j \in \mathcal{P}$.

Opomba. Če π-sistemu dodamo $\varnothing$ in Ω, spet dobimo π-sistem.

Izrek. *Če je $\mathcal{P} = \{B_1, \ldots, B_n\}$ π-sistem in je A neodvisen od vseh B_k, je A neodvisen od vseh dogodkov, ki jih lahko sestavimo iz dogodkov v $\mathcal{P}$ s komplementiranjem, preseki in unijami.*

Dokaz. S preprostim izračunom lahko pokažemo, da če je A neodvisen od dogodkov $C_1, \ldots, C_m$, ki so vsi disjunktni od A, je A neodvisen tudi od njihove unije. Poleg tega opazimo, da so vsi dogodki, ki jih sestavimo v izreku, končne unije dogodkov $B_1^* \cap \ldots \cap B_m^*$, kjer je B_i^* bodisi enak B_i bodisi B_i^{c}.

V luči teh ugotovitev je dovolj dokazati, da je A neodvisen od vsakega dogodka $B_1^* \cap \ldots \cap B_m^*$. Če izberemo vse dogodke, kjer ni komplementa, je presek v $\mathcal{P}$, zato jih lahko nadomestimo z enim samim. Brez škode za splošnost se torej omejimo na dogodke oblike $B_1^{\mathsf{c}} \cap \ldots \cap B_m^{\mathsf{c}} \cap B_{m+1}$. Velja

$$P\left(A \cap \left(\bigcup_i B_i\right)^{\mathsf{c}} \cap B_{m+1}\right) = P(A \cap B_{m+1}) - P\left(\left(\bigcup_i B_i\right) \cap A \cap B_{m+1}\right),$$

kjer smo uporabili pomožni sklep $P(A \cap B^{\mathsf{c}}) = P(A) - P(A \cap B)$, ki ga izpeljemo iz dejstva $P(A) = P(A \cap B) + P(A \cap B^{\mathsf{c}})$. Zgornji izraz je nadalje enak

$$P(A)P(B_{m+1}) - P\left(\bigcup_i A \cap B_i \cap B_{m+1}\right),$$

ker sta A in B_{m+1} neodvisna. Drugi člen razvijemo po formuli za vključitve in izključitve in dobimo

$$\begin{aligned}P(A)P(B_{m+1}) - \sum_i P(A \cap B_i \cap B_{m+1}) + \sum_{i,j} P(A \cap B_i \cap B_j \cap B_{m+1})&\\ - \ldots + (-1)^m P(A \cap B_1 \cap \ldots \cap B_{m+1}).&\end{aligned}$$

V vseh členih dobimo presek A z dogodkom v $\mathcal{P}$, torej lahko izpostavimo $P(A)$;

$$P(A)\left(P(B_{m+1}) - \sum_i P(B_i \cap B_{m+1}) + \ldots\right).$$

V drugem členu produkta smo dobili razvoj dogodka po formuli za vključitve in izključitve, ki ga lahko skrčimo v

$$P(A)\left(P(B_{m+1}) - P\left(\bigcup_i B_i \cap B_{m+1}\right)\right).$$

Nazadnje še uporabimo zgornji sklep v drugo smer in dobimo

$$P(A)P\left(B_{m+1}\left(\bigcup_i B_i\right)^{\mathsf{c}}\right),$$

kar zaključi dokaz. □

4.2 Slučajne spremenljivke in porazdelitve

Definicija. SLUČAJNA SPREMENLJIVKA X je funkcija $\Omega \to \mathbb{R}$, da je za $a < b$ množica $X^{-1}((a, b])$ dogodek v σ-algebri dogodkov $\mathcal{F}$.

Opomba. Ekvivalentno definicijo dobimo, če namesto polodprtih intervalov vzamemo odprte ali zaprte. Izbiro je predpisal ISO standard.

Opomba. Funkcija sama po sebi je popolnoma deterministična, naključna je izbira argumenta.

Definicija. Slučajna spremenljivka je DISKRETNA, če je njena zaloga vrednosti števna ali končna množica.

Definicija. PORAZDELITEV diskretne slučajne spremenljivke X z vrednostmi $(x_i)_i$ je dana z verjetnostmi $P(X^{-1}(x_i))$.

Vprašanje 8. Definiraj slučajne spremenljivke. Kdaj je slučajna spremenljivka diskretna? Kaj je porazdelitev?

Obstaja nekaj standardnih diskretnih porazdelitev.

Primer (Hipergeometrijska porazdelitev). Imamo posodo z B belimi in R rdečimi kroglicami. Označimo $N = B + R$ in naključno izberemo $n \leq N$ kroglic tako, da so vse podmnožice enako verjetne. Če z X označimo število izbranih belih kroglic, dobimo slučajno spremenljivko. Za $\max\{0, n - R\} \leq k \leq \min\{n, B\}$ je

$$P(X = k) = \frac{\binom{B}{k}\binom{R}{n-k}}{\binom{N}{n}}.$$

Na kratko označimo $X \sim \mathrm{HiperGeom}(n, B, N)$.

Vprašanje 9. Opiši hipergeometrijsko porazdelitev.

Primer (Binomska porazdelitev). Kovanec z maso m vržemo n-krat zaporedoma, pri čemer so vsi meti medsebojno neodvisni, verjetnost grba pa je $p \in (0,1)$. Naj bo X število grbov v teh n metih. Tedaj za $k = 0, \dots, n$ velja

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Označimo $X \sim \mathrm{Bin}(n, p)$.

Vprašanje 10. Opiši binomsko porazdelitev.

Primer (Geometrijska porazdelitev). Naj bo X število metov kovanca, potrebnih, da pade prvi grb. Pri tem so meti neodvisni, kovanec pade na grb z verjetnostjo p. Možne vrednosti za X so vsi $k \in \mathbb{N}$, velja

$$P(X = k) = (1-p)^{k-1} p.$$

Na kratko označimo $X \sim \mathrm{Geom}(p)$.

Vprašanje 11. Opiši geometrijsko porazdelitev.

Primer (Negativna binomska porazdelitev). Mečemo kovanec in čakamo na m grbov; naj bo X število potrebnih metov. Možne vrednosti X so tedaj $k = m, m+1, \dots$, pri čemer velja

$$P(X = k) = \binom{k-1}{m-1} p^m (1-p)^{k-m}.$$

Oznaka je $X \sim \mathrm{NegBin}(m, p)$.

Vprašanje 12. Opiši negativno binomsko porazdelitev.

Definicija. Pochhammerjev simbol $(a)_n$ je definiran kot

$$(a)_n = a(a+1) \dots (a+n-1).$$

Opomba. Izračunamo ga lahko tudi kot

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}.$$

Primer (Poissonova porazdelitev). Oglejmo si dogajanje binomske porazdelitve, ko velja $np = \lambda$ konstanta, in ko $n \to \infty$. Tedaj

$$\begin{aligned}
\lim_{n \to \infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} &= \lim_{n \to \infty} \frac{\lambda^k}{k!} \frac{n(n-1)\dots(n-k+1)}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k}{k!} e^{-\lambda}.
\end{aligned}$$

Če je za $k = 0, 1, \ldots$ verjetnost

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!},$$

pravimo, da ima X Poissonovo porazdelitev, in označimo $X \sim \mathrm{Po}(\lambda)$.

Vprašanje 13. Opiši Poissonovo porazdelitev.

Definicija. Porazdelitev zvezne slučajne spremenljivke je podana z verjetnostmi $P(X \in (a, b])$ za $a < b$.

Opomba. Pogosto želimo izračunati $P(X \in A)$, kjer $A \subseteq \mathbb{R}$ ni interval. V tem primeru lahko verjetnost izračunamo, če je A sestavljena iz števnih unij, števnih presekov in komplementov polodprtih intervalov. Taki družini pravimo BORELOVE MNOŽICE, in jo označimo z $B(\mathbb{R})$. Tehnično so to najmanjša σ-algebra na $\mathbb{R}$, ki vsebuje vse polodprte intervale.

Definicija. Slučajna spremenljivka X ima ZVEZNO PORAZDELITEV, če obstaja nenegativna funkcija $f_X : \mathbb{R} \to [0, \infty)$, da je

$$P(X \in (a, b]) = \int_a^b f_X(x)dx.$$

Funkciji f_X pravimo GOSTOTA PORAZDELITVE.

Vprašanje 14. Definiraj zvezno porazdelitev in gostoto porazdelitve.

Primer (Normalna porazdelitev). Pravimo, da ima X normalno porazdelitev s parametroma μ in σ^2, če je gostota enaka

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Pri tem je σ razdalja od μ do prevoja, μ pa središče porazdelitve.

Vprašanje 15. Opiši normalno porazdelitev.

Primer (Eksponentna porazdelitev). Pravimo, da ima X eksponentno porazdelitev s parametrom λ, če velja

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Označimo z $X \sim \exp(\lambda)$.

Primer (Gama porazdelitev). Slučajna spremenljivka X ima gama porazdelitev s parametroma $a, \lambda > 0$, če je gostota enaka

$$f_X(x) = \begin{cases} \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

Oznaka: $X \sim \Gamma(a, \lambda)$.

Primer (Enakomerna porazdelitev). Predvidljivo je

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b, \\ 0 & \text{sicer.} \end{cases}$$

Oznaka: $X \sim \mathrm{U}(a, b)$.

Primer (Beta porazdelitev). Spremenljivka X ima beta porazdelitev, če je

$$f_X(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} & 0 < x < 1, \\ 0 & \text{sicer.} \end{cases}$$

Oznaka: $X \sim \mathrm{Beta}(a, b)$.

Vprašanje 16. Opiši eksponentno, gama, enakomerno in beta porazdelitev.

Če je X slučajna spremenljivka, kakšna mora biti funkcija $f : \mathbb{R} \to \mathbb{R}$, da bo $Y = f(X)$ tudi slučajna spremenljivka? Za poljubna $a < b$ mora biti $X^{-1}(f^{-1}((a, b)))$ dogodek. Če je funkcija (odsekoma) zvezna, že zadošča; potreben in zadosten pogoj pa je, da je f merljiva, torej da je $f^{-1}(A)$ dogodek za vse $A \in B(\mathbb{R})$.

Definicija. Porazdelitvena funkcija slučajne spremenljivke X je funkcija $F_X : \mathbb{R} \to \mathbb{R}$, podana z $F_X(x) = P(X \leq x)$.

Če ima X gostoto f_X, je

$$F_X(x) = \int_{-\infty}^{x} f_X(u) du.$$

Izrek. *Naj bo F_X porazdelitvena funkcija slučajne spremenljivke X. Tedaj velja*

- *F_X je nepadajoča,*
- $\lim_{x\to\infty} F_X(x) = 1$ *in* $\lim_{x\to-\infty} F_X(x) = 0$*,*
- *F_X je desno zvezna.*

Dokaz. Prva točka: za $x < y$ velja $F_X(y) - F_X(x) = P(X \in (x, y]) \geq 0$.

Druga točka: Definiramo $A_n = \{X \leq n\}$. Tedaj velja

$$\bigcup_{n=1}^{\infty} A_n = \Omega,$$

in ti dogodki so naraščajoči. Potem je

$$1 = P(\Omega) = P\left(\bigcup_n A_n\right) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} F_x(n).$$

Ker je F_X nepadajoča, velja tudi $\lim_{x\to\infty} F_X(x) = 1$ zvezno. Za drugo formulo podobno definiramo $B_n = \{X \leq -n\}$, kar je padajoče zaporedje dogodkov s praznim presekom. Za limito velja podoben sklep kot prej.

Tretja točka: Naj bo $x_n \downarrow x$. Definiramo $C_n = \{X \leq x_n\}$, velja

$$\bigcap_{n=1}^{\infty} C_n = \{X \leq x\}.$$

Ker so C_n padajoči, je

$$F_X(x) = P\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} P(X \leq x_n) = \lim_{n\to\infty} F_X(x_n),$$

torej je F_X res desno zvezna. □

Vprašanje 17. Definiraj porazdelitveno funkcijo slučajne spremenljivke. Kakšne lastnosti ima?

Vprašanje 18. Naj velja $X \sim N(\mu, \sigma^2)$ in $Y = aX + b$. Kakšna je gostota Y?

Odgovor: Računamo

$$F_Y(y) = P(Y \leq y) = P(X \leq \frac{y-b}{a}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{(y-b)/a} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) du.$$

Ker je F_X zvezno odvedljiva, je

$$f_Y(y) = F_Y'(y) = \frac{1}{\sqrt{2\pi}\sigma a} \exp\left(-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}\right),$$

torej $Y \sim N(a\mu + b, a^2\sigma^2)$. ⊠

Definicija. Če je $Z \sim N(0,1)$, rečemo, da ima Z STANDARDIZIRANO NORMALNO PORAZDELITEV. Porazdelitveno funkcijo Z označimo s ϕ, torej

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{1}{2}u^2\right) du.$$

Če je $X \sim N(\mu, \sigma^2)$, je torej

$$F_X(x) = \phi\left(\frac{x-\mu}{\sigma}\right).$$

Vprašanje 19. Kaj je standardizirana normalna porazdelitev?

Vprašanje 20. Kaj je verjetnostna transformacija?

Odgovor: Naj bo X zvezno porazdeljena s porazdelitveno funkcijo F_X, za katero predpostavimo, da je zvezna. Definiramo $Y = F_X(X)$ in računamo za $y \in (0,1)$

$$P(Y \leq y) = P(X \in F^{-1}((-\infty, y])) = F_X(\sup\{x \mid F_X(x) \leq y\}) = y,$$

torej $Y \sim U(0,1)$. ⊠

Definicija. Naj bo $p \in (0,1)$. Vsakemu številu x_p, za katerega je $P(X \leq x_p) = p$, rečemo p-TI KVANTIL porazdelitve slučajne spremenljivke X.

Opomba. p-ti kvantil <u>ni</u> enolično določen.

Vprašanje 21. Kaj je p-ti kvantil porazdelitve slučajne spremenljivke?

4.2.1 Slučajni vektorji

Primer (Multinomska porazdelitev). Imamo r škatel, vanje mečemo n kroglic. Meti so neodvisni, škatlo k zadenemo z verjetnostjo p_k. Velja $\sum_k p_k = 1$. V vsaki škatli je pristalo slučajno število kroglic $X_k \sim \mathrm{Bin}(n, p_k)$. Te spremenljivke zložimo v vektor $\underline{X} = (X_1, \ldots, X_r)$, ki ima porazdelitev

$$P(X_1 = k_1, \ldots, X_r = k_r) = p_1^{k_1} \ldots p_r^{k_r} \binom{n}{k_1, \ldots, k_r}.$$

Pravimo, da ima $\underline{X}$ multinomsko porazdelitev s parametroma n in $\underline{p}$, ter označimo $\underline{X} \sim \mathrm{Multinom}(n, \underline{p})$.

Vprašanje 22. Kaj je multinomska porazdelitev?

Definicija. SLUČAJNI VEKTOR $\underline{X}$ s komponentami $X_1, \ldots, X_r$ je funkcija $\underline{X} : \Omega \to \mathbb{R}^r$, da je

$$\underline{X}^{-1}\left(\prod_{k=1}^{r} (a_k, b_k]\right)$$

dogodek za vse $a_k < b_k$.

Definicija. Slučajni vektor je DISKRETEN, če ima vrednosti v končni ali števni množici.

Vprašanje 23. Definiraj slučajne vektorje.

Izrek. *Naj bosta $\underline{X}$ in $\underline{Y}$ diskretna slučajna vektorja. Za vse možne vrednosti $\underline{x}$ vektorja $\underline{X}$ velja*

$$P(\underline{X} = \underline{x}) = \sum_{\underline{y}} P(\underline{X} = \underline{x}, \underline{Y} = \underline{y}).$$

Formuli pravimo FORMULA ZA ROBNO PORAZDELITEV.

Vprašanje 24. Povej formulo za robno porazdelitev.

4.2.2 Neodvisnost slučajnih spremenljivk

Definicija. Slučajni spremenljivki X in Y sta NEODVISNI, če za vsaki Borelovi A in B velja $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$.

Definicija. Slučajne spremenljivke $X_1, \ldots, X_n$ so NEODVISNE, če za vsak nabor Borelovih množic $A_1, \ldots, A_n$ velja

$$P(\forall i \,.\, X_i \in A_i) = \prod_j P(X_j \in A_j).$$

Definicija. Slučajne spremenljivke $\{X_i\}_{i \in I}$ so NEODVISNE, če so neodvisne vse končne poddružine.

Vprašanje 25. Definiraj neodvisnost slučajnih spremenljivk.

Izrek. *Naj za diskretni slučajni spremenljivki X in Y velja $P(X = x, Y = y) = f(x)g(y)$ za funkciji $f : R(X) \to \mathbb{R}$ in $g : R(Y) \to \mathbb{R}$ (R je zaloga vrednosti). Potem sta X in Y neodvisni.*

Dokaz. Po formuli za robne porazdelitve je

$$P(X = x) = f(x) \sum_y g(y) = f(x) C_1.$$

Podobno $P(Y = y) = C_2 g(y)$. Predpišemo

$$P(X = x, Y = y) = P(X = x) P(Y = y) C_1^{-1} C_2^{-1}.$$

Dokazati moramo še, da velja $C_1 C_2 = 1$. Seštejmo

$$\begin{aligned} 1 &= \sum_{x,y} P(X = x, Y = y) \\ &= \frac{1}{C_1 C_2} \sum_{x,y} P(X = x) P(Y = y) \\ &= \frac{1}{C_1 C_2} \left(\sum_x P(X = x) \right) \left(\sum_y P(Y = y) \right) \\ &= \frac{1}{C_1 C_2}. \end{aligned}$$

□

Vprašanje 26. Kako še lahko določiš, da sta slučajni spremenljivki neodvisni? Dokaži.

4.2.3 Pričakovana vrednost diskretnih spremenljivk

Definicija. Naj bo X diskretna slučajna spremenljivka z vrednostmi $x_1, x_2, \ldots$. PRIČAKOVANA VREDNOST $E(X)$ je število, dano z

$$E(X) = \sum_i x_i P(X = x_i).$$

Če slučajno spremenljivko X vstavimo v funkcijo, spet dobimo slučajno spremenljivko $Y = f(X)$. Če je X diskretna, je taka tudi Y, torej

$$E(Y) = \sum_y yP(Y = y) = \sum_x f(x)P(X = x).$$

Vprašanje 27. Definiraj pričakovano vrednost diskretne spremenljivke. Kako se presli-ka s funkcijo?

Izrek. *Naj bodo $X_1, \ldots, X_n$ slučajne spremenljivke in $\alpha_1, \ldots, \alpha_n$ konstante. Če obstaja $E(X_i)$ za $i = 1, \ldots, n$, obstaja tudi*

$$E\left(\sum_i \alpha_i X_i\right) = \sum_i \alpha_i E(X_i).$$

Definicija. Slučajna spremenljivka I ima Bernoullijevo porazdelitev, če je njena zaloga vrednosti enaka $\{0, 1\}$. Če označimo $p = P(I = 1)$, pišemo $I \sim \text{Bernoulli}(p)$.

Vprašanje 28. Kaj je Bernoullijeva porazdelitev?

4.2.4 Večrazsežne zvezne porazdelitve

Definicija. Slučajni vektor $\underline{X}$ ima zvezno porazdelitev, če obstaja nenegativna funkcija $f_{\underline{X}}(\underline{x})$, da za $A \subseteq \mathbb{R}^n$ velja

$$P(\underline{X} \in A) = \int_A f_{\underline{X}}(\underline{x}) d\underline{x}.$$

Funkciji $f_{\underline{X}}$ pravimo gostota.

Vprašanje 29. Definiraj gostoto porazdelitve slučajnega vektorja.

Izrek. *Naj bo $f_{\underline{X}}(\underline{x})$ gostota vektorja $\underline{X}$ in $m < n$. Privzemimo, da je funkcija*

$$(x_1, \ldots, x_n) \mapsto \int_{\mathbb{R}^{n-m}} f_{\underline{X}}(x_1, \ldots, x_n) dx_{m+1} \ldots d_{x_n}$$

Riemannovo integrabilna (lahko tudi v izlimitiranem smislu) po vseh Jordanovo izmer-ljivih množicah. Potem je to funkcija gostote vektorja $\underline{X}' = (x_1, \ldots, x_m)$.

Izrek. *Slučajna vektorja $\underline{X}, \underline{Y}$ sta neodvisna natanko tedaj, ko je*

$$f_{\underline{X},\underline{Y}}(\underline{x}, \underline{y}) = f_{\underline{X}}(\underline{x}) f_{\underline{Y}}(\underline{y})$$

skoraj povsod.

Dokaz. V desno: Velja

$$P(X \in A, Y \in B) = \int_{A\times B} f_{\underline{X},\underline{Y}}(\underline{x},\underline{y})d\underline{x}d\underline{y},$$
$$P(X \in A)P(Y \in B) = \int_A f_{\underline{X}}(\underline{x})d\underline{x}\int_B f_{\underline{Y}}(\underline{y})d\underline{y} = \int_{A\times B} f_{\underline{X}}f_{\underline{Y}}.$$

Ker sta vektorja neodvisna, sta ti količini enaki, torej sta integrirani funkciji enaki skoraj povsod.

V levo: Velja

$$\begin{aligned} P(\underline{X} \in A, \underline{Y} \in B) &= \int_{A\times B} f_{\underline{X}}(\underline{x})f_{\underline{Y}}(\underline{y})d\underline{x}d\underline{y} \\ &= \int_A f_{\underline{X}}(\underline{x})d\underline{x}\int_B f_{\underline{Y}}(\underline{y})d\underline{y} \\ &= P(\underline{X} \in A)P(\underline{Y} \in B). \end{aligned}$$

□

Vprašanje 30. Karakteriziraj neodvisnost zvezno porazdeljenih slučajnih vektorjev in dokaži karakterizacijo.

Izrek. *Naj bo $f_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) = f(\underline{x})g(\underline{y})$ za nenegativni f, g. Potem sta $\underline{X}$ in $\underline{Y}$ neodvisni.*

Dokaz je praktično enak kot pri podobnem izreku za diskretne spremenljivke, le da pišemo integrale namesto vsot.

Izrek. *Naj bo $\underline{X}$ slučajni vektor z gostoto $f_{\underline{X}}(\underline{x})$. Predpostavimo $P(\underline{X} \in U) = 1$ za neko odprto množico $U \subseteq \mathbb{R}^n$. Naj bo $\phi : U \to V$ difeomorfizem in $\underline{Y} = \phi(\underline{X})$. Velja $P(\underline{Y} \in V) = 1$ in*

$$f_{\underline{Y}}(\underline{y}) = \begin{cases} f_{\underline{X}}\left(\phi^{-1}(\underline{y})\right)\left|\det D\phi^{-1}(\underline{y})\right| & \underline{y} \in V \\ 0 & \underline{y} \notin V \end{cases}$$

Vprašanje 31. Povej transformacijsko formulo.

Definicija. Naj bo X zvezno porazdeljena slučajna spremenljivka z gostoto $f_X(x)$. Definiramo

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)dx,$$
$$E(g(x)) = \int_{-\infty}^{\infty} g(x)f_X(x)dx.$$

Vprašanje 32. Kako je definirana pričakovana vrednost zvezne slučajne spremenljivke?

Definicija. Za slučajno spremenljivko X imenujemo količino $E(X^m)$ m-TI MOMENT.

Definicija. Za slučajno spremenljivko X imenujemo količino $E((X - E(x))^m)$ m-TI CENTRALNI MOMENT.

Vprašanje 33. Kaj je moment in kaj centralni moment slučajne spremenljivke?

Definicija. Če imamo množico števil $x_1, \ldots, x_n$, njihov RAZTROS definiramo kot

$$\sigma^2 = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2,$$

kjer je $\bar{x}$ povprečje.

Definicija. VARIANCA slučajne spremenljivke je količina

$$\operatorname{var}(X) = E(X^2) - (E(X))^2.$$

Pravimo, da varianca obstaja, če obstajata obe pričakovani vrednosti.

Definicija. Naj bo X slučajna spremenljivka, za katero $\operatorname{var}(X)$ obstaja. Količini $\sqrt{\operatorname{var}(X)}$ rečemo STANDARDNI ODKLON slučajne spremenljivke X in jo označimo s $\mathrm{SD}(X)$.

Definicija. KOVARIANCA dveh slučajnih spremenljivk je $\operatorname{cov}(X, Y) = E(XY) - E(X)E(Y)$. Rečemo, da obstaja, če obstajajo vse pričakovane vrednosti.

Vprašanje 34. Definiraj raztros, varianco, standardni odklon in kovarianco.

Izrek. *Naj bodo $X_1, \ldots, X_m$ in $Y_1, \ldots, Y_n$ slučajne spremenljivke in $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ konstante. Velja*

$$\operatorname{cov}\left(\sum_{k=1}^{m} \alpha_k X_k, \sum_{l=1}^{n} \beta_l Y_l \right) = \sum_{k,l} \alpha_k \beta_l \operatorname{cov}(X_k, Y_l).$$

Vprašanje 35. Povej nekaj lastnosti kovariance.

Odgovor:

- bilinearnost,
- $\operatorname{cov}(X, X) = \operatorname{var}(X)$,
- $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$,
- če sta X in Y neodvisni, je $\operatorname{cov}(X, Y) = 0$.

⊠

Definicija. Količina

$$\rho = \frac{\operatorname{cov}(X, Y)}{\sqrt{\operatorname{var}(X) \operatorname{var}(Y)}}$$

se imenuje KORELACIJSKI KOEFICIENT.

Vprašanje 36. Definiraj korelacijski koeficient.

4.2.5 Pogojne pričakovane vrednosti

Definicija. POGOJNA PORAZDELITEV diskretne slučajne spremenljivke X glede na dogodek B je dana z verjetnostjo $P(X = x \mid B)$. POGOJNA PRIČAKOVANA VREDNOST slučajne spremenljivke X glede na dogodek B je dana z

$$E(X \mid B) = \sum_x xP(X = x \mid B).$$

Alternativno lahko izračunamo

$$E(X \cdot I_B) = \sum_x xP(X \cdot I_B = x) = P(B) \sum_x xP(X = x \mid B) = P(B)E(X \mid B),$$

torej $E(X \mid B) = E(X \cdot I_B)/P(B)$. Iz te formule sledi linearnost pogojne pričakovane vrednosti.

Vprašanje 37. Definiraj pogojno pričakovano vrednost in izpelji alternativno izražavo.

Izrek. *Naj bo $\{H_1, H_2, \ldots\}$ particija. Naj bo X diskretna slučajna spremenljivka. Velja*

$$E(X) = \sum_k E(X \mid H_k)P(H_k).$$

Dokaz. Računamo

$$\sum_k E(X \mid H_k)P(H_k) = \sum_k \sum_x xP(X = x \mid H_k)P(H_k) = \sum_x P(X = x) = E(X).$$

□

Vprašanje 38. Dokaži formulo za popolno pričakovano vrednost.

Definicija. Naj bosta $\underline{X}, \underline{Y}$ slučajna vektorja z gostoto $f_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})$. Za $\underline{x}$, za katere je $f_{\underline{X}}(\underline{x}) > 0$, definiramo POGOJNO GOSTOTO $\underline{Y}$ glede na $\{\underline{X} = \underline{x}\}$ kot

$$f_{\underline{Y} \mid \underline{X}=\underline{x}}(\underline{y}) = \frac{f_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})}{f_{\underline{X}}(\underline{x})}.$$

Vprašanje 39. Definiraj pogojno gostoto.

Izrek. *Naj bosta $\underline{X}, \underline{Y}$ slučajna vektorja. Za $p = \dim \underline{X}$ velja*

$$E(g(\underline{Y})) = \int_{\mathbb{R}^p} E(g(\underline{Y}) \mid \underline{X} = \underline{x}) f_{\underline{X}}(\underline{x}) d\underline{x}.$$

Dokaz. Računamo

$$\begin{aligned}\int_{\mathbb{R}^p} E(g(\underline{Y}) \mid \underline{X} = \underline{x}) f_{\underline{X}}(\underline{x}) d\underline{x} &= \int_{\mathbb{R}^p} f_{\underline{X}}(\underline{x}) d\underline{x} \int_{\mathbb{R}^q} g(\underline{y}) f_{\underline{Y} \mid \underline{X}=\underline{x}}(\underline{y}) d\underline{y} \\ &= \int_{\mathbb{R}^{p+q}} g(\underline{y}) f_{\underline{X}}(\underline{x}) f_{\underline{Y} \mid \underline{X}=\underline{x}}(\underline{y}) d\underline{x} d\underline{y} \\ &= \int_{\mathbb{R}^{p+q}} g(\underline{y}) f_{\underline{X},\underline{Y}}(\underline{x}, \underline{y}) d\underline{x} d\underline{y} \\ &= E(g(\underline{Y})).\end{aligned}$$

□

Vprašanje 40. Dokaži zvezno verzijo formule za popolno pričakovano vrednost.

4.3 Rodovne funkcije

Definicija. RODOVNA FUNKCIJA nenegativne celoštevilske slučajne spremenljivke X je dana z $G_X(s) = \sum_{k=0}^{\infty} s^k P(X = k)$.

Potenčna vrsta konvergira enakomerno na $[-1, 1]$, ker je tam dominirana z vrsto $\sum_{k=0}^{\infty} P(X = k) = 1$. Torej je $G_X(s)$ zvezna na $[-1, 1]$. Na intervalu $(-1, 1)$ je vrsta neskončnokrat odvedljiva. Iz formule $E(f(x)) = \sum_k f(k)P(X = k)$ za $f(x) = s^x$ dobimo

$$E(s^X) = \sum_{k=0}^{\infty} s^k P(X = k) = G_X(s).$$

Izrek. *Naj bosta X, Y neodvisni spremenljivki. Velja $G_{X+Y}(s) = G_X(s)G_Y(s)$.*

Dokaz. Sledi iz dejstva $E(s^{X+Y}) = E(s^X s^Y)$. □

Vprašanje 41. Povej nekaj lastnosti rodovnih funkcij.

Izrek. *Naj bo X slučajna spremenljivka. Potem velja*

- $E(X) = \lim_{s\uparrow 1} G'_X(s)$
- $E(X(X-1)(X-2)\dots(X-k+1)) = \lim_{s\uparrow 1} G_X^{(k)}(s)$

Dokaz. Samo za prvo točko. Opazimo, da ima G_X pozitivne koeficiente, torej imajo vsi odvodi pozitivne koeficiente. Sledi, da so odvodi na $(0, 1)$ povsod naraščajoči.

Za začetek predpostavimo $E(X) < \infty$. Za $s \in (0, 1)$ velja

$$\sum_{k=1}^{N} kP(X = k)s^{k-1} \leq G'_X(s) \leq E(X),$$

kjer prva neenakost velja, ker je vrsta na levi glava vrste za G'_X. Funkcija G'_X je na $(0,1)$ naraščajoča in omejena, torej limita $\lim_{s\uparrow 1} G'_X(s)$ obstaja. Levi člen v zgornji neenakosti konvergira k $E(X)$ za $N \to \infty$, torej trditev velja po izreku o sendviču.

Če je $E(X) = \infty$, potem levi člen v zgornji neenačbi divergira za $N \to \infty$, in je posledično tudi

$$\lim_{s\uparrow 1} G'_X(s) = \infty.$$

□

Vprašanje 42. Kako z odvodom rodovne funkcije izračunaš pričakovano vrednost? Dokaži.

Izrek. *Naj bodo $N, X_1, \ldots$ neodvisne celoštevilske slučajne spremenljivke in naj bodo $X_1, X_2, \ldots$ enako porazdeljene. Naj bo $X = X_1 + \ldots + X_N$. Potem velja $G_X(s) = G_N(G_{X_i}(s))$.*

Dokaz. Računamo

$$\begin{aligned} G_X(s) &= E(s^X) \\ &= \sum_{n=0}^{\infty} E(s^X \mid N = n)P(N = n) \\ &= \sum_{n=0}^{\infty} E(s^{X_1+\ldots+X_n} \mid N = n)P(N = n). \end{aligned}$$

Na tej točki upoštevamo neodvisnost spremenljivk in dobimo

$$\begin{aligned} G_X(s) &= \sum_{n=0}^{\infty} E(s^{X_1+\ldots+X_n})P(N = n) \\ &= \sum_{n=0}^{\infty} E(s^{X_1}) \cdots E(s^{X_n})P(N = n) \\ &= \sum_{n=0}^{\infty} (G_{X_1}(s))^n P(N = n) \\ &= G_N(G_{X_1}(s)). \end{aligned}$$

□

Vprašanje 43. Kako izračunaš porazdelitev vsote slučajno mnogo slučajnih spremenljivk? Dokaži.

4.3.1 Procesi razvejanja

Predpostavimo, da so porazdelitve števila sinov za vsakega posameznika enake, da so generacije sočasne in da so števila sinov medsebojno neodvisne. Kakšna je verjetnost, da kraljeva rodbina izumre?

Označimo z Z_n število posameznikov v n-ti generaciji. Naj bodo $\xi_{n,k}$ vse neodvisne z rodovno funkcijo G. Definiramo rekurzivno $Z_0 = 1$ in

$$Z_{n+1} = \xi_{n+1,1} + \ldots + \xi_{n+1,Z_n}.$$

Zaradi predpostavke o neodvisnosti so izpolnjene vse predpostavke zadnjega izreka in lahko izračunamo za $G_n = G_{Z_n}$

$$G_{n+1}(s) = G_n(G(s)).$$

Ker je $G_1 = G$, dobimo $G_k = G \circ G \circ \ldots \circ G$. Velja

$$\{\text{rodbina izumre}\} = \bigcup_{n=1}^{\infty} \{Z_n = 0\}$$

in ti dogodki so naraščajoči, torej je

$$\eta = P(\text{rodbina izumre}) = \lim_{n\to\infty} P(Z_n = 0).$$

Ker je $P(Z_n = 0) = G_n(0)$, dobimo

$$\eta = \lim_{n\to\infty} G_n(0) = \lim_{n\to\infty} G_{n+1}(0) = G(\lim_{n\to\infty} G_n(s)) = G(\eta).$$

Ker je $s = 1$ vedno fiksna točka G na $[0, 1]$, ni pa nujno edina. Naj bo $\bar{\eta}$ poljubna fiksna točka na $[0, 1]$. Funkcija G je nepadajoča na $[0, 1]$, torej $G(0) \leq G(\bar{\eta}) = \bar{\eta}$. Z večkratno uporabo G na tej neenakosti dobimo $G_n(0) \leq \bar{\eta}$, iz česar sledi $\eta = \lim_{n\to\infty} G_n(0) \leq \bar{\eta}$ in je η najmanjša fiksna točka.

Vprašanje 44. Kako dobiš verjetnost izumrtja procesa razvejanja? Dokaži.

4.3.2 Panjerjeva rekurzija

Definicija. Celoštevilska slučajna spremenljivka N ima porazdelitev Panjerjevega razreda, če obstajata konstanti a, b, da je

$$P(N = n) = (a + \frac{b}{n})P(N = n - 1).$$

Binomska, Poissonova in negativna binomska so vse v tem razredu. Zanima nas vsota $X = X_1 + \ldots + X_N$. Po dolgem računanju pridemo do formule

$$P(X = n + 1) = \frac{1}{1 - aP(X_1 = 0)} \sum_{k=1}^{n+1} \left(a + \frac{bk}{n+1}\right) P(X_1 = k)P(X = n - k + 1).$$

Vprašanje 45. Povej formulo za Panjerjevo rekurzijo.

4.4 Tabele

Porazdelitev X	X^2
$N(0,1)$	$\Gamma\left(\frac{1}{2},\frac{1}{2}\right)$

Tabela 4.1: Porazdelitve

Porazdelitev X	$E(X)$	$E(X^2)$	Varianca
$\mathrm{Bin}(n,p)$	np	$n^2p^2+np(1-p)$	npq
$\mathrm{NegBin}(n,p)$	$\frac{n}{p}$	$\frac{m(m+1)}{p^2}-\frac{m}{p}$	$\frac{mq}{p^2}$
$N(\mu,\sigma^2)$	μ	$\sigma^2+\mu^2$	σ^2
$\Gamma(a,\lambda)$	$\frac{a}{\lambda}$	$\frac{a(a+1)}{\lambda^2}$	$\frac{a}{\lambda^2}$

Tabela 4.2: Pričakovane vrednosti

Porazdelitev X	Porazdelitev Y	Porazdelitev $X+Y$
$\mathrm{Po}(\lambda)$	$\mathrm{Po}(\mu)$	$\mathrm{Po}(\lambda+\mu)$
$\mathrm{Bin}(m,p)$	$\mathrm{Bin}(n,p)$	$\mathrm{Bin}(m+n,p)$
$\mathrm{Polya}(a,\beta)$	$\mathrm{Polya}(b,\beta)$	$\mathrm{Polya}(a+b,\beta)$
$N(\mu,\sigma^2)$	$N(\nu,\tau^2)$	$N(\mu+\nu,\sigma^2+\tau^2)$
$\Gamma(a,\lambda)$	$\Gamma(b,\lambda)$	$\Gamma(a+b,\lambda)$

Tabela 4.3: Vsote

Porazdelitev $\mathrm{Polya}(a,\beta)$ je dana z

$$P(X=k)=\frac{\beta^a(a)_k}{k!(1+\beta)^{a+k}}.$$

5 Algebra 3

5.1 Reševanje polinomskih enačb

Vemo, da za polinomsko enačbo $x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$ obstaja razširitev polja $\mathbb{F}$, v kateri je enačba razcepna. Rešitev polinomske enačbe se izraža z radikali, če se da zapisati rešitve enačbe s pomočjo računskih operacij in korenov.

Definicija. Naj bo K polje. Razširitvi oblike $K(\sqrt[n]{a})/K$ za $a \in K$ pravimo RADIKALSKA RAZŠIRITEV polja K.

Polinomska enačba $p(X) = 0$ je rešljiva z radikali natanko tedaj, ko rešitve živijo v neki razširitvi E/F, pri čemer obstaja končna veriga razširitev $F \subseteq E_0 \subseteq \ldots \subseteq E_k = E$, kjer so zaporedne razširitve radikalske. Problem se prevede na vprašanje iz teorije grup. Od tu naprej predpostavimo, da so vse razširitve končne.

Definicija. Naj bo E/F končna razširitev. To je NORMALNA RAZŠIRITEV, če za vsak nerazcepen polinom $p(X) \in F[X]$ velja ena od naslednjih možnosti:

- p nima ničle v E
- p ima vse ničle v E

Ekvivalentno: vsak nerazcepen polinom $p(X) \in F[X]$ z vsaj eno ničlo v E razpade v linearne faktorje v $E[X]$.

Primer. $\mathbb{Q}(\sqrt[3]{2})$ nad $\mathbb{Q}$ ni normalna. Polinom $p(X) = X^3 - 2$ ne razpade.

Primer. $\mathbb{Q}(\sqrt{2})$ je normalna.

Vprašanje 1. Definiraj radikalske in normalne razširitve. Povej primer normalne razširitve in razširitve, ki ni normalna.

Izrek. *Naj bo E/F končna razširitev. Potem je ta razširitev normalna natanko tedaj, ko je E razpadno polje nekega polinoma s koeficienti iz F.*

Dokaz. V desno: Recimo, da je E normalna. Naj bo $p_i(X)$ minimalni polinom a_i. Definiramo $p(X) = p_1(X) \ldots p_k(X)$. Zaradi normalnosti so vse ničle polinoma $p_i(X)$ v E, torej so vse ničle polinoma $p(X)$ v E. Velja $F(p) \subseteq E$, ampak $a_1, \ldots, a_k \in F(p)$, torej $F(p) \supseteq F(a_1, \ldots, a_k) = E$.

V levo: Recimo, da je $E = F(p)$. Vzemimo poljuben nerazcepen $q(X) \in F[X]$, ki ima neko ničlo $a \in E$. Naj bo b poljubna ničla q. Ničli a in b imata isti minimalni polinom, q, torej je $F(a) \cong F(b)$. Izomorfizem δ lahko razširimo na izomorfizem razpadnih polj τ. Predpostavimo lahko, da je $\delta(a) = b$, torej tudi $\tau(a) = b$. Ničla a je racionalen izraz, odvisen od $a_1, \ldots, a_k$, torej je tudi b racionalen izraz, odvisen od $a_1, \ldots, a_k$, in posledično $b \in E$. □

Vprašanje 2. Karakteriziraj normalnost za končne razširitve. Dokaži karakterizacijo.

Definicija. Polinom $p(X) \in F[X]$ je SEPARABILEN, če ima same enostavne ničle. Končna razširitev E/F je SEPARABILNA, če je minimalni polinom vsakega elementa $a \in E$ separabilen.

Opomba. V karakteristiki 0 je vsaka končna razširitev separabilna.

Izrek (Primitivni element)**.** *Vsaka separabilna razširitev je enostavna.*

Dokaz. Če je F končno polje, je tudi E končno in je $(E^*, \cdot)$ ciklična grupa, generirana z a. Velja $E = F(a_1, \ldots, a_n)$, brez škode za splošnost $a_i \neq 0$, torej obstajajo n_i, da je $a_i = a^{n_i}$. Torej je $E = F(a)$.

Če pa je F neskončno polje, zapišimo $E = F(a_1, \ldots, a_n)$, ker je razširitev končna. Poleg tega so a_i algebraični. Ker velja $F(a_1, \ldots, a_n) = F(a_1, \ldots, a_{n-2})(a_{n-1}, a_n)$, izrek zadošča dokazati za $n = 2$.

Naj bo torej $E = F(b, c)$ ter $p(X)$ in $q(X)$ minimalna polinoma b in c. Označimo z E_1 razpadno polje polinoma $p(X)q(X)$ in z $b = b_1, \ldots, b_r$, $c = c_1, \ldots, c_s$ njune ničle v tem polju. Izberimo $\lambda \in F$ tako, da $\lambda \neq (b_j - b)(c - c_k)^{-1}$ za vse j in k. Tak λ res obstaja, ker je polje neskončno. Sedaj definiramo $a = b + \lambda c$. Očitno je $F(a) \subseteq F(b, c)$; trdimo, da velja tudi vsebovanost v drugi smeri. Vpeljimo $f(X) = p(a - \lambda X) \in F(a)[X]$. Velja $f(c) = p(a - \lambda c) = p(b) = 0$, torej je c hkrati ničla $f(X)$ in $q(X) \in F[X] \subseteq F(a)[X]$. Naj bo $\tilde{q}(X) \in F(a)[X]$ minimalni polinom elementa c nad poljem $F(a)$. Velja $\tilde{q}|f$ in $\tilde{q}|q$.

Možne ničle polinoma $\tilde{q}$ so skupne ničle polinomov q in f. Vemo, da imata skupno ničlo c; katere od c_i pa so še ničle f? Če je $f(c_j) = p(a - \lambda c_j) = 0$, potem je $a - \lambda c_j = b_k$ za nek k, torej $b + \lambda c - \lambda c_j = b_k$ in posledično $\lambda = (b_j - b)(c - c_k)^{-1}$, kar je po predpostavki nemogoče. Torej je c edina ničla $\tilde{q}$. Ker je to minimalni polinom in polje F separabilno, je ničla enostavna in zato $\tilde{q}(X) = X - c \in F(a)[X]$, torej $c \in F(a)$. □

Primer. Razširitev $E = F_2(\sqrt{t})$ je algebraična razširitev, ki ni separabilna. Minimalni polinom za $\sqrt{t}$ je $x^2 - t = (x + \sqrt{t})^2$.

Vprašanje 3. Kdaj je razširitev separabilna? Podaj primer končne razširitve, ki ni algebraična. Povej izrek o primitivnem elementu in ga dokaži za končna polja.

Definicija. GALOISOVA GRUPA RAZŠIRITVE E/F je grupa

$$\mathrm{Gal}(E/F) = \{\sigma \in \mathrm{Aut}(E) \mid \sigma|_F = \mathrm{id}_F\}.$$

Lema. *Naj bo E/F razširitev, $a \in E$ ničla $p(X) \in F[X]$ ter $\sigma \in \mathrm{Gal}(E/F)$. Potem je $\sigma(a)$ tudi ničla $p(X)$.*

Dokaz. Uporabimo σ na zapisu $p(X) = a_n X^n + \cdots + a_0$. □

Definicija. Naj bo $p(X) \in F[X]$ polinom. GALOISOVA GRUPA polinoma p je $\mathrm{Gal}(p) = \mathrm{Gal}(F(p)/F)$.

Vprašanje 4. Definiraj Galoisovo grupo razširitve in polinoma.

Definicija. Normalnim separabilnim razširitvam pravimo Galoisove razširitve.

Trditev. *Naj bo E/F končna Galoisova razširitev. Potem je* $|\mathrm{Gal}(E/F)| = [E : F]$.

Dokaz. Po izreku o primitivnem elementu obstaja $a \in E$, da je $E = F(a)$. Naj bo $p(X) \in F[X]$ minimalni polinom a. Zaradi normalnosti so vse ničle p v E, torej je za $\sigma \in \mathrm{Gal}(E/F)$ slika $\sigma(a)$ lahko katerakoli ničla p. Ničel je ravno toliko kot je stopnja razširitve, ker ima p paroma različne ničle. Torej je za σ natanko toliko možnosti. □

Opomba. Če je E/F končna separabilna razširitev, je $|\mathrm{Gal}(E/F)| \leq [E : F]$.

Vprašanje 5. Kolikšen je red Galoisove grupe končne Galoisove razširitve? Dokaži.

Lema. *Naj bo $F \subseteq L \subseteq E$.*

- *Če je E/F končna, je tudi E/F končna.*
- *Če je E/F normalna, je tudi E/L normalna.*
- *Če je E/F separabilna, je tudi E/L separabilna.*

Dokaz. Prva točka je enostavna, dokaz tretje pa je podoben dokazu druge, torej dokažemo le to. Naj bo $p(X) \in L[X]$ nerazcepen z ničlo $a \in E$ in naj bo $q(X) \in F[X]$ minimalen polinom za a nad F. Trdimo, da $p(X)$ deli $q(X)$. Vzemimo največji skupni delitelj $d(X) \in L[X]$ polinomov $p(X)$ in $q(X)$. Ker imata skupno ničlo a, d ni konstanten, $d(a) = 0$ in d deli p. Ker je p nerazcepen, mora biti $d(X) = p(X)$, torej p res deli q. Vsaka ničla b polinoma p je torej tudi ničla q. Ker je E normalna, je $b \in E$. □

Izrek (Fundamentalni izrek Galoisove teorije)**.** *Naj bo E/F končna Galoisova razširitev.*

- *Predpisa $L \mapsto \mathrm{Gal}(E/L)$ in $G \mapsto E^G$ sta paroma inverzni preslikavi med vmesnimi polji razširitve in podgrupami Galoisove grupe.*
- *Za poljubni vmesni polji $F \subseteq L \subseteq M \subseteq E$ velja $[M : L] = |\mathrm{Gal}(E/F) : Gal(E/M)|$.*
- *Za vmesno polje $F \subseteq L \subseteq E$ je L/F normalna natanko tedaj, ko je $\mathrm{Gal}(E/L) \lhd \mathrm{Gal}(E/F)$. Velja še $\mathrm{Gal}(L/F) \cong \mathrm{Gal}(E/F)/\mathrm{Gal}(E/L)$.*

Dokaz. Za prvo točko: Dokažimo, da za vsak L velja

$$E^{\mathrm{Gal}(E/L)} = L.$$

Očitno je $L \subseteq E^{\mathrm{Gal}(E/L)}$. Trdimo, da za vsak $a \in E \backslash L$ obstaja $\sigma \in \mathrm{Gal}(E/L)$, ki ga ne fiksira. Naj bo $q(X) \in L[X]$ minimalni polinom a. Ker $a \notin L$, je $\operatorname{st} q > 1$, torej ima q od a različno ničlo $b \in E$. Polji $L(a)$ in $L(b)$ sta izomorfni, izomorfizem lahko razširimo do izomorfizma razpadnih polj polinoma q; ta izomorfizem fiksira elemente L in slika a v b.

Sedaj pokažimo $\operatorname{Gal}(E/E^G) = G$. Očitno je $G \subseteq \operatorname{Gal}(E/E^G)$, za drugo inkluzijo pa je dovolj dokazati (ker so vse grupe končne), da je $[E : E^G] \leq |G|$. Zaradi separabilnosti je $E = E^G(a)$. Naj bo $q(X) \in E^G[X]$ minimalni polinom a. Velja $\operatorname{st} q = [E : E^G]$. Definiramo

$$p(X) = \prod_{\tau \in G} (X - \tau(a)) \in E[X].$$

Trdimo, da je $p(X) \in E^G[X]$, za kar moramo pokazati, da vsak $\sigma \in G$ fiksira koeficiente polinoma. Označimo polinom slik koeficientov z

$$\sigma(p(X)) = \prod_{\tau \in G} (X - \sigma(\tau(a))).$$

Ko τ preteče cel G, tudi $\sigma\tau$ preteče cel G, torej $\sigma(p(X)) = p(X) \in E^G[X]$. Velja $p(a) = 0$, torej q deli p in zato $[E : E^G] = \operatorname{st} q \leq \operatorname{st} p = |G|$.

Druga točka: Preprost račun

$$[M : L] = \frac{[E : L]}{[E : M]} = \frac{|\operatorname{Gal}(E/L)|}{|\operatorname{Gal}(E/M)|} = |\operatorname{Gal}(E/L) : \operatorname{Gal}(E/M)| \,.$$

Tretja točka: Dokaz v več korakih. Prvo pokažimo, da je L/F normalna natanko tedaj, ko za vsak $\sigma \in \operatorname{Gal}(E/F)$ velja $\sigma(L) = L$. V desno naj bo $\sigma \in \operatorname{Gal}(E/F)$. Velja $L = F(a_1, \ldots, a_k)$, označimo s p_i minimalni polinom a_i. Ker je razširitev normalna, so vse ničle p_i v L, torej se a_i s σ slika v neko ničlo p_i. Torej je $\sigma(a_i) \in L$ za vse i, drugo inkluzijo pa dobimo z inverzom σ^{-1}. Podobno v drugo smer.

Sedaj pokažimo, da za $F \subseteq L \subseteq E$ in poljuben $\sigma \in \operatorname{Gal}(E/F)$ velja $\operatorname{Gal}(E/\sigma(L)) = \sigma \operatorname{Gal}(E/F)\sigma^{-1}$. To velja zato, ker je $\tau \in \operatorname{Gal}(E/\sigma(L))$ natanko tedaj, ko je $\tau|_{\sigma(L)} = \mathrm{id}$, oziroma $\tau(\sigma(x)) = \sigma(x)$ za vsak $x \in L$. Če množimo z leve s σ^{-1}, dobimo $\sigma^{-1}\tau\sigma(x) = x$ za $x \in L$, kar je ekvivalentno zahtevi $\sigma^{-1}\tau\sigma \in \operatorname{Gal}(E/L)$.

V tretjem koraku opazimo, da je desni del ekvivalence v prvem koraku enaka zahtevi, da za vsak $\sigma \in \operatorname{Gal}(E/F)$ velja $\operatorname{Gal}(E/F) = \operatorname{Gal}(E/\sigma(L))$, oziroma $\sigma \operatorname{Gal}(E/L)\sigma^{-1} = \operatorname{Gal}(E/L)$, kar pa je po definiciji enako $\operatorname{Gal}(E/L) \lhd \operatorname{Gal}(E/F)$.

V zadnjem koraku recimo, da je L/F normalna, in definirajmo $\phi : \operatorname{Gal}(E/F) \to \operatorname{Gal}(L/F)$ kot

$$\phi(\sigma) = \sigma|_L \,.$$

Po prvem koraku dokaza prejšnje točke zožena preslikava res slika L v L, iz izreka o izomorfizmu torej dobimo $\operatorname{Gal}(E/F)/\operatorname{Gal}(E/L) \cong \operatorname{Gal}(L/F)$. □

Vprašanje 6. Povej in dokaži fundamentalni izrek Galoisove teorije.

5.1.1 Rešljive grupe

Definicija. Grupa G je REŠLJIVA, če obstaja končno zaporedje podgrup

$$\{e\} = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_k = G,$$

da je za vsak i $G_i \triangleleft G_{i+1}$ ter G_{i+1}/G_i Abelova grupa.

Vprašanje 7. Kaj je rešljiva grupa? Povej nekaj pozitivnih in negativnih primerov.

Odgovor: Abelove grupe so rešljive, enostavne nekomutativne grupe pa niso. Za $H = \langle(123)\rangle$ je $\{id\} \triangleleft H \triangleleft S_3$, za $K = \langle(12)(34), (14)(32)\rangle$ pa $\{id\} \triangleleft K \triangleleft A_4 \triangleleft S_4$, preverimo lahko, da sta tako S_3 kot S_4 rešljivi. ⊠

Trditev. *Vsaka podgrupa rešljive grupe je rešljiva.*

Dokaz. Naj bo G rešljiva, $\{e\} = G_0 \triangleleft \cdots \triangleleft G_k = G$, in naj bo $H \leq G$. Potem je $G_i \cap H \triangleleft G_{i+1} \cap H$, za kvocient pa velja

$$\frac{G_{i+1} \cap H}{G_i \cap H} = \frac{G_{i+1} \cap H}{G_i \cap G_{i+1} \cap H} \cong \frac{(G_{i+1} \cap H) \cdot G_i}{G_i} \leq \frac{G_{i+1}}{G_i},$$

kjer smo uporabili drugi izrek o izomorfizmu. Grupa G_{i+1}/G_i je Abelova, torej je

$$\{e\} = G_0 \cap H \triangleleft \cdots \triangleleft G_k \cap H = H.$$

□

Vprašanje 8. Pokaži, da je podgrupa rešljive grupe rešljiva.

Trditev. *Kvocient rešljive grupe je rešljiv.*

Dokaz. Naj bo G rešljiva in $N \triangleleft G$. Preverimo lahko, da velja

$$\{e\} \triangleleft \frac{G_0 N}{N} \triangleleft \cdots \triangleleft \frac{G_k N}{N} = \frac{G}{N}.$$

Z uporabo tretjega in drugega izreka o izomorfizmu izračunamo

$$\frac{G_{i+1}N/N}{G_i N/N} \cong \frac{G_{i+1}/N}{G_i/N} = \frac{G_{i+1}G_i/N}{G_i/N} \cong \frac{G_{i+1}}{G_{i+1} \cap G_i N} \cong \frac{G_{i+1}/G_i}{(G_{i+1} \cap G_i N)/G_i}.$$

Kvocient Abelove grupe je Abelova grupa. □

Vprašanje 9. Pokaži, da je kvocient rešljive grupe rešljiv.

Trditev. *Naj bo $N \triangleleft G$ in naj bosta N in G/N rešljivi. Potem je G rešljiva grupa.*

Dokaz. Naj bosta $N_0 \triangleleft \cdots \triangleleft N_k$ in $G_0/N \triangleleft \cdots G_l/N$ zaporedji edink. Iz $G_i/N \triangleleft G_{i+1}/N$ sledi $G_i \triangleleft G_{i+1}$. Po tretjem izreku o izomorfizmu je G_{i+1}/G_i Abelova (ker je $\frac{G_{i+1}/N}{G_i/}$ Abelova). Sedaj imamo zaporedje

$$N_0 \triangleleft N_i \triangleleft \cdots \triangleleft N_k = N = G_0 \triangleleft G_1 \triangleleft \cdots \triangleleft G_l = G.$$

□

Vprašanje 10. Kako sestaviš rešljivo grupo iz rešljive edinke in kvocienta? Dokaži.

5.1.2 Rešljivost polinomskih enačb z radikali

Dan je polinom $p(X) \in F[X]$. Rešljivost polinomske enačbe z $p(X) = 0$ z radikali je ekvivalentna obstoju razširitve E/F, ki jo dobimo z zaporedjem radikalskih razširitev $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_k = E$.

Lema. *Naj bo F polje, ki vsebuje primitivni n-ti koren enote. Naj bo $a \in F$. Potem je Galoisova grupa polinoma $X^n - a$ ciklična.*

Dokaz. Naj bo ε primitivni n-ti koren enote. Če je b neka ničla $X^n - a$, so vse ničle oblike $b, \varepsilon b, \ldots, \varepsilon^{n-1}b$. Potem je $F(X^n - a) = F(b, \varepsilon b, \ldots, \varepsilon^{n-1}b) = F(b)$, ker je $\varepsilon \in F$. Preslikava $\sigma \in \mathrm{Gal}(X^n - a)$ je določena s sliko b;

$$b \mapsto \varepsilon^i b.$$

S tem je določena injektivna preslikava $Gal(X^n - a) \to \mathbb{Z}_n$, ki je hkrati homomorfizem grup. □

Izrek. *Enačba $p(X) = 0$ je rešljiva z radikali natanko tedaj, ko je Galoisova grupa polinoma p rešljiva grupa.*

Dokaz. Dokažemo le implikacijo v desno; v drugo smer je podobno. Enačba je rešljiva, torej obstajajo razširitve $F = E_0 \subseteq E_1 \subseteq \cdots \subseteq E_k = E$, da je $E_{i+1} = E_i(a_i)$ in $a_i^{n_i} \in F$. Definiramo $n = n_0 \ldots n_{k-1}$. Naj bo ε primitivni n-ti koren enote. Za ta dokaz se ne bomo ukvarjali s separabilnostjo; privzamemo npr. da je karakteristika polja 0.

Po izreku o primitivnem elementu obstaja a, da je $E = F(a)$. Naj bo $g(X) \in F[X]$ minimalni polinom a in $\Omega = F(g(X)(X^n - 1))$. To je normalna razširitev F, ki vsebuje E in ε. Vzemimo še $\tilde{E}$ kot normalno zaprtje polja $E(\varepsilon)$ v Ω tj. najmanjša normalna razširitev $E(\varepsilon)$, ki je vsebovana v Ω.

Polje $\tilde{E}$ vsebuje vse $\sigma(E(\varepsilon))$ za $\sigma \in \mathrm{Gal}(\Omega/F)$, torej je enako

$$\tilde{E} = F(\varepsilon, a_0, a_1, \ldots, a_{k-1}, \sigma_1(a_0), \ldots, \sigma_2(a_0), \ldots).$$

Razširitev $\tilde{E}/F$ lahko naredimo po korakih, po vrstnem redu kot zgoraj. V vsaki vmesni razširitvi dodamo ničlo polinoma $X^{n_i} - b_i$ za nek b_i, torej so razširitve radikalske. Preverimo lahko, da so tudi normalne.

Po Galoisovi korespondenci dobimo verigo podgrup

$$\mathrm{Gal}(\tilde{E}/F) \supseteq \mathrm{Gal}(\tilde{E}/F(\varepsilon)) \supseteq \cdots \supseteq \mathrm{Gal}(\tilde{E}/\tilde{E}) = \{e\}.$$

Zaradi normalnosti je to zaporedje edink. Kvocienti so po lemi ciklične grupe, torej je $\mathrm{Gal}(\tilde{E}/F)$ rešljiva. Grupa

$$\frac{\mathrm{Gal}(\tilde{E}/F)}{\mathrm{Gal}(\tilde{E}/F(p))} \cong \mathrm{Gal}(F(p)/F) = \mathrm{Gal}(p)$$

je kvocient rešljive grupe, torej je rešljiva. □

Vprašanje 11. Dokaži, da je rešljivost polinomske enačbe $p(X) = 0$ z radikali implicira rešljivost grupe $\mathrm{Gal}(p)$.

5.2 Moduli

Definicija. Naj bo R kolobar in M neprazna množica. LEVI R-MODUL je $(M, +, \cdot)$, kjer sta $+ : M \times M \to M$ in $\cdot : R \times M \to M$. Pri tem zahtevamo

- $(M, +)$ je Abelova grupa
- $r(m_1 + m_2) = rm_1 + rm_2$
- $(r_1 + r_2)m = r_1 m + r_2 m$
- $(r_1 r_2)m = r_1(r_2 m)$
- $1 \cdot m = m$

Analogno lahko definiramo desne module. Vsak levi R-modul je pravzaprav tudi desni R^{opp} modul, kjer je R^{opp} kolobar z enakim seštevanjem kot v R in množenjem v obratnem vrstnem redu.

Vprašanje 12. Definiraj modul.

Primer. Vektorski prostor nad poljem F je levi (in desni) F-modul.

Primer. Abelova grupa (tu pisana aditivno) je $\mathbb{Z}$-modul za množenje

$$n \cdot g = \underbrace{g + g + \ldots + g}_{n}.$$

Primer. Naj bo I levi ideal kolobarja R. Potem je I levi R-modul.

Primer. Če je V vektorski prostor nad poljem F in R množica endomorfizmov $V \to V$, je V levi R-modul za množenje

$$A \cdot v = A(v).$$

Vprašanje 13. Naštej nekaj enostavnih primerov modulov.

Definicija. Naj bo M R-modul. Množica $N \subseteq M$ je PODMODUL v M, če je N za podedovano seštevanje in množenje s skalarjem tudi modul. Oznaka $N \leq M$.

Definicija. Naj bo M R-modul in $X \subseteq M$ neprazna množica. Najmanjšemu podmodulu, ki vsebuje X, pravimo PODMODUL, GENERIRAN Z MNOŽICO X. Oznaka $\langle X \rangle$. Modul je KONČNO GENERIRAN, če obstaja končna X, da je $M = \langle X \rangle$. Če je modul generiran z enim samim elementom, je CIKLIČNI.

Definicija. Naj bosta N, M R-modula. Preslikava $\varphi : M \to N$ je HOMOMORFIZEM R-modulov, če velja $\varphi(m_1 + m_2) = \varphi(m_1) + \varphi(m_2)$ in $\varphi(rm) = r\varphi(m)$. JEDRO homomorfizma je množica

$$\ker \varphi = \{m \in M \mid \varphi(m) = 0\},$$

SLIKA pa

$$\operatorname{im} \varphi = \{\varphi(m) \mid m \in M\}.$$

Definicija. Naj bo M R-modul in $N \leq M$. Za relacijo

$$m_1 \sim m_2 \Leftrightarrow m_1 - m_2 \in N$$

definiramo KVOCIENTNI MODUL $M/_{\sim} = M/_N$.

Izrek. *Za kvocientne module veljajo naslednje lastnosti:*

- *če je $\varphi : M \to N$ homomorfizem, potem je $M/_{\ker \varphi} \cong \operatorname{im} \varphi$,*
- *če sta N, K podmodula v M, potem je $(N + K)/_K \cong N/_{N \cap K}$,*
- *velja*

$$\frac{M/_N}{L/_N} \cong M/_L.$$

Definicija. Naj bodo $N_1, \ldots, N_k$ R-moduli. ZUNANJI DIREKTNI PRODUKT je R-modul $N_1 \times N_2 \times \cdots \times N_k$ s seštevanjem in množenjem s skalarjem po komponentah. Oznaka $N_1 \oplus N_2 \oplus \cdots \oplus N_k$.

Opomba. Če imamo neskončno družino modulov, lahko naredimo podobno konstrukcijo, ki ji pravimo KARTEZIČNI PRODUKT.

Definicija. Naj bodo $N_1, \ldots, N_k \leq M$. Vsoti $N_1 + \cdots + N_k$ pravimo NOTRANJA DIREKTNA VSOTA, če za vsak i velja $N_i \cap (N_1 + \cdots + N_{i-1} + N_{i+1} + \cdots + N_k) = \{0\}$.

Opomba. Zunanja in notranja definicija sta ekvivalentni.

Definicija. Naj bo M R-modul in $X \subseteq M$. Pravimo, da je X BAZA modula M, če je $M = \langle X \rangle$ in če za vsak $k \in \mathbb{N}$ in vse $x_1, \ldots, x_k \in X$ iz enakosti $r_1 x_1 + \cdots + r_n x_n = 0$ sledi $r_i = 0$.

Primer. Obstajajo moduli, ki nimajo baz. Modul $M = \mathbb{Z}_n = \mathbb{Z}/_{n\mathbb{Z}}$ je $\mathbb{Z}$-modul. Recimo, da je nek $x + n\mathbb{Z}$ v bazi. Ampak $n(x + n\mathbb{Z}) = 0$, n pa v $\mathbb{Z}$ ni enak 0.

Definicija. Modul je PROST, če ima bazo.

Vprašanje 14. Kaj je baza modula? Podaj primer modula, ki ni prost.

Izrek. *Naj bo M R-modul. Naslednje trditve so ekvivalentne:*

- *M je prost R-modul.*
- *Obstaja indeksna množica Λ, da je M kot R-modul izomorfen zunanji direktni vsoti kopij R-modula R:*
$$M \cong \bigoplus_{\lambda \in \Lambda} R.$$
- *Obstaja indeksna množica Λ in podmoduli $M_\lambda \leq M$, $M_\lambda \cong R$, da je*
$$M = \bigoplus_{\lambda \in \Lambda} M_\lambda.$$

Dokaz. Druga in tretja točka sta očitno ekvivalentni. Naj bo X baza modula M. Trdimo $M = \bigoplus_{x \in X} \langle x \rangle$. Ker je $M = \langle X \rangle$, je M vsota podmodulov $\langle x \rangle$. Vzemimo $z \in \langle x \rangle \cap \sum_{y \neq x} \langle y \rangle$. Velja $z = rx = r_1 y_1 + \cdots + r_k y_k$, iz definicije baze pa sledi $r = r_i = 0$. Preveriti moramo še, da je $\langle x \rangle \cong R$. Ustrezen izomorfizem $R \to \langle x \rangle$ bo $\varphi(r) = rx$.

Sedaj dokažimo, da iz druge točke sledi prva. Imamo izomorfizem
$$\varphi : \bigoplus_{\lambda \in \Lambda} R \to M.$$

Izberimo e_λ kot Λ-terico, ki ima na mestu λ enico, drugje pa 0, in definirajmo $x_\lambda = \varphi(e_\lambda)$. Preverimo lahko, da je $X = \{x_\lambda \mid \lambda \in \Lambda\}$ baza. $\square$

Vprašanje 15. Karakteriziraj prostost modula in dokaži karakterizacijo.

6 Analiza 4

6.1 Osnovni tipi PDE

Uporabljamo notacijo s predavanj:

- $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ je multiindeks,
- $|\alpha| = \alpha_1 + \dots + \alpha_n$,
- odvod funkcije u z multiindeksom α je

$$D^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}},$$

- za $k \in \mathbb{N}$ označimo skupek odvodov k-tega reda kot $D^k u = \{D^\alpha u \mid |\alpha| = k\}$. To včasih obravnavamo kot množico, včasih pa kot vektor ali matriko.

Definicija. Parcialna diferencialna enačba je enačba, ki vsebuje neznano funkcijo u vsaj dveh spremenljivk ter nekatere njene parcialne odvode.

Definicija. Parcialna enačba k-tega reda je enačba oblike

$$F(x, u(x), Du(x), \dots, D^k u(x)) = 0,$$

kjer je $x \in U^{\text{odp}} \subseteq \mathbb{R}^n$ in $F : U \times R \times \mathbb{R}^n \times \dots \times R^{n^k} \to \mathbb{R}$ dana.

Rešitev PDE je vsaka funkcija, ki enačbi zadošča. Lahko je klasična ali posplošena rešitev; za klasično obstajajo vsi odvodi $Du, \dots, D^k u$, ki jih vstavimo v F, posplošena rešitev pa je rešitev v smislu integracije po delih.

Primer. Dana je enačba $\Delta u = f$. Če je $u \in \mathcal{C}^2$ in označimo $\Delta u = g$, za vsak $\phi \in \mathcal{C}^\infty(U)$ s kompaktnim nosilcem velja

$$\langle \Delta u, \phi \rangle = \langle g, \phi \rangle = \int_U g(x)\phi(x) dx.$$

Iz integracije po delih sledi $\langle \Delta u, \phi \rangle = \langle u, \Delta \phi \rangle$ oziroma $\langle u, \Delta \phi \rangle = \langle g, \phi \rangle$. Tu zahtevamo samo, da je u dovolj regularen, da lahko izračunamo integral v skalarnem produktu. Šibka rešitev je vsaka $u \in \mathcal{C}(U)$, ki zadošča $\langle u, \Delta \phi \rangle = \langle f, \phi \rangle$ za vsak ϕ.

Poznamo štiri osnovne tipe parcialnih diferencialnih enačb.

- Linearna PDE je linearna kombinacija parcialnih odvodov

$$\sum_{|\alpha| \le k} a_\alpha(x) D^\alpha u(x) = f(x).$$

- Semilinearna PDE je linearna samo v odvodu najvišjega reda

$$\sum_{|\alpha| = k} a_\alpha(x) D^\alpha u + b(x, u, Du, D^{k-1} u) = 0.$$

- Kvazilinearna PDE je enačba oblike

$$\sum_{|\alpha|=k} a_\alpha(x, u, Du, \ldots, D^{k-1}u) \cdot D^\alpha u + b(x, u, \ldots, D^{k-1}u) = 0.$$

- Povsem nelinearna PDE je enačba, ki je nelinearno odvisna od odvodov najvišjega reda.

Vprašanje 1. Definiraj linearne, semilinearne in kvazilinearne PDE.

Definicija. Pravimo, da je problem PDE z začetnim ali robnim pogojem DOBRO POSTAVLJEN, če ima naslednje lastnosti:

- rešitev obstaja,
- rešitev je enolična,
- rešitev je zvezno odvisna od začetnih pogojev.

6.2 Kvazilinearne enačbe prvega reda v dveh spremenljivkah

Obravnavamo enačbe oblike

$$a(x, y, u)u_x + b(x, y, u)u_y = c(x, y, u). \tag{6.1}$$

Reševanja se bomo lotili z metodo karakteristik. Enačbo prvo prepišemo v obliko

$$\langle (a, b, c), (u_x, u_y, -1)\rangle = 0.$$

Ploskvi $z = u(x, y)$, ki reši kvazilinearno enačbo, pravimo INTEGRALNA PLOSKEV enačbe (**??**). Če je $u : D \subseteq \mathbb{R}^2 \to \mathbb{R}$, lahko njen graf parametriziramo z $\vec{r}(x, y) = (x, y, u(x, y))$. Ploskev opišemo z normalo $\vec{r}_x \times \vec{r}_y = -(u_x, u_y, -1)$. Če u že imamo, tedaj vektor (a, b, c) leži v tangentni ravnini na graf.

Definicija. Vektor (a, b, c) se imenuje KARAKTERISTIČNA SMER za PDE.

Pišimo $V = (a, b, c)$, čemur pravimo pripadajoče (karakteristično) vektorsko polje. Iz eksistenčnega izreka sledi, da za vsak $p_0 = (x_0, y_0, z_0)$ obstaja natanko ena tokovnica polja V, ki poteka skozi točko p_0. Če označimo $\varphi_t(p) = \gamma_p(t)$, je φ tok, in velja $\varphi_{t+s} = \varphi_t \circ \varphi_s$.

Trditev. *Naj bodo $a, b, c \in \mathcal{C}^1(\mathbb{R}^3)$, $S = \{z = u(x, y)\}$ poljubna integralna ploskev za* (6.1) *in $p_0 \in S$. Tedaj vsaka karakteristična krivulja za* (6.1)*, ki gre skozi točko p_0, cela leži v S.*

Dokaz. Naj bo $\gamma(t) = (x, y, z)$ karakteristična krivulja z $\gamma(0) = p_0$. Dokazati želimo, da velja $z = u(x, y)$, za kar je dovolj pokazati $\dot{z} = \partial_t u(x, y)$. Označimo $w(t) = z(t) -$

$u(x(t), y(t))$ in računamo

$$\begin{aligned}\dot{w} &= \dot{z} - u_x\dot{x} - u_y\dot{y}\\ &= c(\gamma(t)) - u_x(x,y)a(\gamma(t)) - u_y(x,y)b(\gamma(t))\\ &= c(x,y,w+u(x,y)) - u_x(x,y)a(x,y,w+u(x,y)) - u_y(x,y)b(x,y,w+u(x,y)).\end{aligned}$$

Velja $w(0) = 0$. Dobili smo Cauchyjevo nalogo oblike $\dot{w}(t) = f(t, w(t)), w(0) = 0$. Preverimo lahko, da za f lokalno velja Lipschitzov pogoj, torej po eksistenčnem izreku obstaja enolična rešitev. Ker je u po predpostavki rešitev (6.1), je $w = 0$ rešitev Cauchyjeve naloge, ki je enolična. □

Vprašanje 2. Pokaži, da tokovnice ne zapustijo integralne ploskve.

Posledica. *Ob predpostavkah trditve je vsaka integralna ploskev unija karakterističnih krivulj.*

Posledica. *Ob predpostavkah trditve:*

- *Če se dve integralski ploskvi S_1, S_2 sekata v točki $p \in S_1 \cap S_2$, cela tokovnica skozi p leži v $S_1 \cap S_2$.*
- *Če je $C = S_1 \cap C_2$ krivulja, ki je presek dveh integralskih ploskev S_1, S_2, ki se sekata ne-tangentno, je karakteristična krivulja.*

Dokaz. Tokovnica ne zapusti ne S_1 ne S_2, torej mora biti vsebovana v preseku. Za drug del naj bo $p \in S_1 \cap S_2$. Tangentni ravnini T_pS_1 in T_pS_2 se sekata ne-tangentno, torej sta različni, in je njun presek enodimenzionalen. Ker sta S_1 in S_2 integralski ploskvi, njuni tangentni ravnini vsebujeta smer karakteristike in je zato

$$T_pS_1 \cap T_pS_2 = \{\lambda(a(p), b(p), c(p)) \mid \lambda \in \mathbb{R}\} = T_p(S_1 \cap S_2) = T_pC.$$

Torej je C v vsaki točki tangentna na smer karakteristike, torej je karakteristična krivulja. □

Vprašanje 3. Dokaži, da je presek integralskih ploskev karakteristična krivulja.

Vsaka integralska ploskev $z = u(x, y)$ za (6.1) je unija karakterističnih krivulj polja $V = (a, b, c)$. Naj bo Γ gladka krivulja, parametrizirana z $\gamma(s) = (f(s), g(s), h(s))$. Skozi $\gamma(s)$ napeljemo karakteristično krivuljo, ki vsebuje točko $\gamma(s)$. Privzamemo lahko, da gre ta krivulja skozi točko $\gamma(s)$ ob času $t = 0$. Iščemo vektorsko funkcijo $R(s, t) = (x(s, t), y(s, t), z(s, t))$, ki zadošča

- za vsak s funkcija $R(s, \cdot)$ reši karakteristični sistem $\dot{x} = a(x, y, z), \dot{y} = b(x, y, z), \dot{z} = c(x, y, z)$,
- začetni pogoji: $x(s, 0) = f(s), y(s, 0) = g(s), z(s, 0) = h(s)$.

Naj bo J kompakten interval, $\gamma : J \to \mathbb{R}^3$, $s \in J$ in $\vec{F}_s = (f(s), g(s), h(s))$. Rešujemo sistem

$$\frac{d}{dt}\vec{x}_s = V(\vec{x}_s), \quad \vec{x}_s(0) = \vec{F}_s.$$

Rešitev bo podana z $R(s,t) = \vec{x}_s(t)$. Po eksistenčnem izreku za sisteme NDE obstaja $\varepsilon > 0$, da na intervalu $(-\varepsilon, \varepsilon)$ obstaja natanko ena rešitev $\vec{x}_s : (-\varepsilon, \varepsilon) \to \mathbb{R}^3$. Želeli bi, da je R parametrizacija integralske ploskve. Problem je, če je Γ karakteristična krivulja; tedaj je slika R enodimenzionalna, torej potrebujemo dodatne pogoje.

Poskusimo sistem enačb $x = x(s,t), y = y(s,t)$ prevesti na sistem $s = s(x,y), t = t(x,y)$ in vstaviti v tretjo komponento R. To nam da $z = z(s(x,y), t(x,y)) = u(x,y)$. Kdaj pa to smemo narediti? Spomnimo se: če za neki $s_0 \in J$ velja

$$\frac{\partial(x,y)}{\partial(s,t)}(s_0, 0) = \begin{vmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{vmatrix} \neq 0,$$

inverz lahko poiščemo lokalno na neki okolici $(s_0, 0)$. Po privzetkih na $\vec{F}_s$ je determinanta enaka

$$\begin{vmatrix} f'(s_0) & a(p_0) \\ f'(s_0) & b(p_0) \end{vmatrix} = b(p_0)f'(s_0) - a(p_0)f'(s_0)$$

za $p_0 = (f(s_0), g(s_0), h(s_0))$. Geometrijsko želimo, da je projekcija na prvi dve komponenti tangente $\dot{\gamma}(s_0)$ linearno neodvisna od projekcije vektorja $V(p_0)$ na prvi dve komponenti.

Trditev. *Naj bo $\gamma = (f, g, h)$ pot v $\mathbb{R}^3$ razreda $\mathcal{C}^1$, ki v dani točki $p_0 = \gamma(s_0)$ zadošča pogoju transverzalnosti*

$$\begin{vmatrix} f'(s_0) & a(p_0) \\ g'(s_0) & b(p_0) \end{vmatrix} \neq 0.$$

Tedaj v neki okolici točke $(x_0, y_0) = (f(s_0), g(s_0))$ obstaja natanko ena rešitev $u = u(x,y)$ kvazilinearne PDE (6.1)*, ki zadošča pogoju $h(s) = u(f(s), g(s))$ za vsak s blizu s_0.*

Dokaz. Vse razen enoličnosti smo že pokazali. Denimo, da neka integralna ploskev vsebuje točko p_0. Tedaj po že dokazanem vsebuje vse karakteristične krivulje skozi točko p_0. Posledično (vsaj lokalno) celo ploskev parametriziramo z $(x(s,t), y(s,t), z(s,t))$. Če privzamemo, da v neki okolici p_0 obstaja še ena rešitev $w = w(x,y)$, za katero je $h = w(f, g)$ na okolici s_0, tedaj je $W = \{z = z(x,y)\}$ unija tokovnic. Če gre neka tokovnica iz W skozi točko p_0, je cela vsebovana v U. Torej $W \subseteq U$. □

Vprašanje 4. Kako poiščeš rešitev kvazilinearne PDE prvega reda dveh spremenljivk? Katere predpostavke potrebuješ? Pokaži, da je rešitev enolična.

6.2.1 Linearna PDE

To je enačba oblike

$$a(x,y)u_x + b(x,y)u_y = c(x,y)u + d(x,y).$$

V pripadajočem polju $V = (a, b, cu + d)$ funkcije a, b, c, d niso odvisne od u. Enačbe karakteristik imajo obliko

$$\begin{aligned}\dot{x} &= a(x,y),\\ \dot{y} &= b(x,y),\\ \dot{z} &= c(x,y)z + d(x,y).\end{aligned}$$

Vprašanje 5. Kakšne so enačbe karakteristik za linearno PDE?

6.2.2 Ovojnica družine ravnin

Naj bo

$$\Pi_\lambda = \{(x,y,z) \in \mathbb{R}^3 \mid z = \psi(x,y,\lambda) = a(\lambda)x + b(\lambda)y\}$$

družina ravnin, kjer za interval $I \subseteq R$ funkciji $a, b \in \mathcal{C}^2(I)$ taki, da je $\vec{n}(\lambda) = (a(\lambda), b(\lambda), -1)$ injektivna in

$$W(a', b') = \begin{vmatrix} a' & b' \\ a'' & b'' \end{vmatrix} \neq 0$$

za vsak λ. Ovojnica družine Π_λ je ploskev

$$\mathcal{O} = \{(x,y,z) \in \mathbb{R}^3 \mid z = \varphi(x,y)\},$$

za katero velja:

- $\mathcal{O} \subseteq \bigcup_\lambda \Pi_\lambda$,
- če je $p \in \mathcal{O} \cap \Pi_\lambda$, potem je normala na $\mathcal{O}$ v točki p vzporedna $\vec{n}(\lambda)$.

Trivialna izbira je $\mathcal{O} = \Pi_\lambda$ za nek λ, te pa v nadaljnje ne bomo upoštevali.

Vprašanje 6. Definiraj ovojnico družine ravnin v $\mathbb{R}^3$.

Vzemimo projekcijo $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ na prvi dve komponenti. Ker zahtevamo, da je $\vec{n}(\lambda)$ injektivna, za vsak par $(x,y) \neq (0,0)$ obstaja natanko določena točka $p \in \mathcal{O}$, da je $(x,y) = \pi(p)$. To nam definira preslikavo $\Lambda : \pi(\mathcal{O}) \to \mathbb{R}$, ki slika par (x,y) v pripadajoči λ.

Naš cilj je poiskati funkcijo φ, ki definira $\mathcal{O}$. Ker je $\mathcal{O} \subseteq \bigcup_\lambda \Pi_\lambda$, velja

$$\varphi(x,y) = a(\Lambda(x,y))x + b(\Lambda(x,y))y.$$

Zahtevali smo, da je normala vzporedna z $\vec{n}(\lambda)$, torej lahko parametriziramo $\mathcal{O}$ z

$$u(x,y) = (x, y, \varphi(x,y)).$$

Upoštevaje definicijo Π_λ lahko zapišemo normalo kot

$$u_x \times u_y = (1, 0, \varphi_x) \times (0, 1, \varphi_y) = (-\psi_x - \psi_\lambda \Lambda_x, -\phi_y - \psi_\lambda \Lambda_y, 1)$$

oziroma nasprotno vrednot kot

$$-u_x \times u_y = (\psi_x, \psi_y, -1) + \psi_\lambda(\Lambda_x, \Lambda_y, 0).$$

Velja $\psi_x = a$ in $\psi_y = b$, torej $-u_x \times u_y = \vec{n}(\Lambda(x, y)) + \psi_\lambda(\Lambda_x, \Lambda_y, 0)$. Zahtevamo, da je to vzporedno z $\vec{n}(\Lambda(x, y))$, torej mora biti drugi člen ničeln. Če je $\Lambda_x = \Lambda_y = 0$, je Λ konstantna preslikava in je $\mathcal{O} = \Pi_\Lambda$; to je izključen trivialni primer. Velja torej $\psi_\lambda = 0$, kar nam da dodatno enačbo za ovojnico.

Vprašanje 7. Kako poiščeš ovojnico družine ploskev v $\mathbb{R}^3$? Izpelji potrebni pogoj.

Ta enačba je ekvivalentna pogoju

$$D(\lambda) := \frac{a'(\lambda)}{b'(\lambda)} = \frac{-y}{x}.$$

Po predpostavki o determinanti Wronskega je odvod D neničeln, torej jo lahko na množici $\{x \neq 0\}$ obrnemo in dobimo $\Lambda_1(x, y) = D^{-1}(-y/x)$. Podobno lahko naredimo za $y \neq 0$, iz česar izpeljemo $\Lambda_2(x, y) = E^{-1}(-x/y)$ za $E(\lambda) = b'(\lambda)/a'(\lambda)$. Če pokažemo, da je $\Lambda_1 = \Lambda_2$, kjer sta obe definirani, bomo lahko zapisali

$$\Lambda(x, y) = \begin{cases} \Lambda_1(x, y), & x \neq 0 \\ \Lambda_2(x, y). & y \neq 0 \end{cases}$$

Vemo, da za $(x, y) \in \pi(\mathcal{O})$ obstaja natanko določen λ, pri katerem je (x, y) slika neke točke na $\mathcal{O}$ s π. Ker je tretja koordinata $z = a(\Lambda_1)x + b(\Lambda_1)y = a(\Lambda_2)x + b(\Lambda_2)y$ pri tem natančno določena, iz pogoja injektivnosti $\vec{n}$ lahko izpeljemo $\Lambda_1 = \Lambda_2$.

Iz zgornje izpeljave hitro vidimo, da velja $\Lambda(\sigma x, \sigma y) = \Lambda(x, y)$ za $\sigma \neq 0$, iz česar sledi naslednja trditev.

Trditev. *Če je $p \in \mathcal{O}$, potem je $\sigma p \in \mathcal{O}$ za poljuben $\sigma \neq 0$.*

6.3 Nelinearne enačbe prvega reda

Rešujemo enačbo oblike

$$F(x, y, u, u_x, u_y) = 0. \tag{6.2}$$

Tu je $F \in \mathcal{C}^1$ funkcija petih realnih spremenljivk in $u = u(x, y)$ iskana funkcija. Vpeljemo oznake $p = u_x$, $q = u_y$. Normala na ploskev $\{z = u(x, y)\}$ je $(u_x, u_y, -1) = (p, q, -1)$. Enačba (6.2) je zveza med točko (x, y, z) na Γ_u in normalo. Družino potencialnih normal parametriziramo z $\vec{n}(\lambda) = (p(\lambda), q(\lambda), -1)$. Za vsak izbrani λ dobimo potencialno tangentno ravnino na ploskev. Enačba te ravnine, ki jo označimo s $\Pi_{\vec{r}_0,\lambda} = \Pi_\lambda$, je

$$\langle \vec{r} - \vec{r}_0, \vec{n}(\lambda) \rangle = 0.$$

Definicija. Ogrinjača ravnin Π_λ se imenuje MONGEOV STOŽEC v točki (x_0, y_0, z_0).

Ogrinjačo dobimo tako, da enačbo ravnin odvajamo po λ in enačimo z 0. Imamo torej $\langle \vec{r} - \vec{r}_0, \vec{n}(\lambda) \rangle = 0$ in $\langle \vec{r} - \vec{r}_0, \vec{n}'(\lambda) \rangle = 0$. Vektor $\vec{r} - \vec{r}_0$ je torej vzporeden z vektorskih produktom $\vec{n}(\lambda) \times \vec{n}'(\lambda)$. Če je $F(x_0, y_0, z_0, p(\lambda), q(\lambda)) = 0$ in odvajamo po λ, dobimo

$$F_p p' + F_q q' = 0.$$

Sledi $(q', -p') \parallel (F_p, F_q)$, torej (v kombinaciji s prejšnjo enačbo) $\vec{r} - \vec{r}_0 \parallel (F_p, F_q, pF_p + qF_q)$. Dobili smo parametrizacijo Mongeovega stožca

$$\begin{aligned}
\vec{r} &= \vec{r}_0 + \mu(P(\lambda), Q(\lambda), R(\lambda)),\\
P(\lambda) &= F_p(x_0, y_0, z_0, p(\lambda), q(\lambda)),\\
Q(\lambda) &= F_q(x_0, y_0, z_0, p(\lambda), q(\lambda)),\\
R(\lambda) &= p(\lambda)P(\lambda) + q(\lambda)Q(\lambda).
\end{aligned}$$

Privzemimo, da imamo integralno ploskev $S = \{z = u(x, y)\}$. Naj bo $\gamma(t) = (x(t), y(t), z(t))$ neka krivulja na tej ploskvi. Tedaj $\dot{\gamma}(t)$ leži na tangentni ravnini, zato $\vec{r}_0 + \dot{\gamma}(t)$ leži v Mongeovem stožcu. Posledično mora veljati (oz. lahko zahtevamo)

$$\begin{aligned}
\dot{x} &= F_p,\\
\dot{y} &= F_q,\\
\dot{z} &= pF_p + qF_q,
\end{aligned}$$

kjer je $p = p(\lambda)$ in $q = q(\lambda)$ za neki (ne vemo kateri) λ. Ker p in q v splošnem ne poznamo, ju obravnavamo kot novi neznanki. Vsaki točki na γ bomo dodali še en košček ravnine, določene z normalo $(p, q, -1)$, s čimer dobimo KARAKTERISTIČNI TRAK.

Trditev. *Če krivulja $\gamma(t) = (x, y, z)$ zadošča pogoju $\dot{z} = p\dot{x} + q\dot{y}$, leži v karakterističnem traku, določenem s $p = u_x$ in $q = u_y$.*

Dokaz. Računamo $\langle (\dot{x}, \dot{y}, \dot{z}), (p, q, -1) \rangle = \langle (\dot{x}, \dot{y}, p\dot{x} + q\dot{y}), (p, q, -1) \rangle = 0$. □

Če odvajamo zgornjo enačbo, dobimo

$$\begin{aligned}
F_x + F_u u_x + F_p p_x + F_q q_x &= 0,\\
F_y + F_u u_y + F_p p_y + F_q q_y &= 0.
\end{aligned}$$

Za $x = x(t)$ in $y = y(t)$ imamo torej

$$\begin{aligned}
\dot{p} = p_x \dot{x} + p_y \dot{y} &= p_x F_p + p_y F_q,\\
\dot{q} &= q_x F_p + q_y F_q,
\end{aligned}$$

iz česar sledi

$$\dot{p} = -F_x - F_u u_x - F_q q_x + p_y F_q = F_q(p_y - q_x) - F_x - F_u u_x$$

Pričakujemo $p = u_x$ in $q = u_y$, torej $p_y - q_x = u_{xy} - u_{yx} = 0$. Posledično je smiselno zahtevati $\dot{p} = -F_x - F_u p$ in $\dot{q} = -F_y - F_u q$. S tem smo dobili karakteristični sistem za x, y, z, q, p:

$$\begin{aligned}
\dot{x} &= F_p \\
\dot{y} &= F_q \\
\dot{z} &= pF_p + qF_q \\
\dot{p} &= -F_x - F_u p \\
\dot{q} &= -F_y - F_u q
\end{aligned}$$

Rešitve imenujemo KARAKTERISTIKE ENAČBE (6.2). Funkcija F je konstantna vzdolž rešitev tega sistema.

Vprašanje 8. Izpelji karakteristični sistem za nelinearno PDE prvega reda.

6.3.1 Cauchyjeva naloga za PDE prvega reda

Vzemimo karakteristiko $\lambda = \lambda(t) = (x, y, z, p, q) = (\gamma, p, q)$. Imamo še začetni pogoj $(x_0, y_0, z_0, p_0, q_0) \in \mathbb{R}^5$.

Definicija. Krivulja $\lambda = \lambda(s)$, ki zadošča $F(\lambda) = 0$, se imenuje INTEGRALNA KRIVULJA, če velja $z_s = px_s + qy_s$.

Rešitev karakterističnega sistema je integralna krivulja. Cauchyjeva naloga je reševanje (6.2), pri čemer graf rešitve vsebuje neko vnaprej dano krivuljo $\gamma(s) = (f(s), g(s), h(s))$. Želimo poiskati smiselna začetna pogoja za p in q, torej iščemo funkciji φ in ψ, da bo $(f, g, h, \varphi, \psi)(s)$ začetni pogoj za sistem. Rešitev sistema bo tedaj funkcija $R = R(s, t)$, za katero bo veljalo $R(s, 0) = (f, g, h, \varphi, \psi)(s)$, in da $(x(s, t), y(s, t), z(s, t))$ parametrizira $\{z = u(x, y)\}$, kjer je u rešitev enačbe (6.2).

Da bo R parametrizirala rešitev enačbe, mora veljati $F(R(s, t)) = 0$. Ker je F vzdolž tokovnic konstantna, zadošča zahtevati, da je $F(R(s, t)) = 0$, torej $F(f, g, h, \varphi, \psi) = 0$. Veljati mora

$$\begin{aligned}
\dot{z} &= u_x\dot{x} + u_y\dot{y}, \\
z_s &= u_x x_s + u_y y_s.
\end{aligned}$$

Ker želimo $p = u_x$ in $q = u_y$, pa dodatno

$$\begin{aligned}
\dot{z} &= p\dot{x} + q\dot{y}, \\
z_s &= px_s + qy_s.
\end{aligned}$$

Iz pogojev pri $t = 0$ torej dobimo

$$h'(s) = \varphi(s)f'(s) + \psi(s)g'(s).$$

7 Izbrane teme iz analize podatkov

Pri strojnem učenju obravnavamo MNOŽICO OPAZOVANIH OBJEKTOV O, in opazujemo lastnosti teh objektov. LASTNOST modeliramo kot preslikavo $V : O \to D_V$, kjer je zaloga vrednosti D_V lahko bodisi poljubna končna množica (v primeru diskretne spremenljivke), ali pa podmnožica $\mathbb{R}$, v primeru numerične spremenljivke. Za diskretne spremenljivke dodatno zahtevamo, da so vrednosti D_V neurejene. Lastnosti ustrezajo SPREMENLJIVKAM, ki jih ločimo na dva tipa. Prve so NAPOVEDNE SPREMENLJIVKE $\mathbf{X} = (X_1, \ldots, X_p)$, ki predstavljajo podatke, druge pa so CILJNE SPREMENLJIVKE, ki jih želimo napovedati. Vsako ciljno spremenljivko lahko obravnavamo posebej, torej si mislimo, da imamo le eno, Y. Pri nenadzorovanem učenju ciljnih spremenljivk nimamo, tam nas namesto napovedi zanimajo drugačna vprašanja.

Vprašanje 1. Opiši podatkovno množico pri strojnem učenju.

Napovedni model predstavimo s funkcijo $m : D_{X_1} \times D_{X_2} \times \cdots \times D_{X_p} \to D_Y$. Funkcija za podane vrednosti $\mathbf{x}$ izračuna napoved ciljne spremenljivke $m(\mathbf{x}) = \hat{y}$. Napovedne modele ločimo glede na to, kakšno spremenljivko napovedujejo. REGRESIJSKI MODEL napoveduje numerično spremenljivko, KLASIFIKACIJSKI pa diskretno.

Za mero, kateri napovedni model je boljši od drugega, definiramo FUNKCIJO IZGUBE $L : D_Y \times D_y \to [0, \infty)$, ki izračuna napako ene napovedi $\hat{y}$ glede na točno vrednost y. Pri regresijskih modelih običajno uporabljamo kvadratno napako

$$L_{\mathrm{SE}}(y, \hat{y}) = (y - \hat{y})^2,$$

pri klasifikacijskih pa

$$L_{01}(y, \hat{y}) = \begin{cases} 0 & y = \hat{y}, \\ 1 & \text{sicer.} \end{cases}$$

Napovedna napaka modela je potem povprečna napaka na vseh primerih iz S,

$$\mathrm{Error}(m, S) = \frac{1}{|S|} \sum_{(\mathbf{x}, y) \in S} L(y, m(\mathbf{x})).$$

Pri nadzorovanem učenju razdelimo množico S na učno in testno množico, S_{train} in S_{test}, ki sta disjunktni in pokrijeta vse primere. Pravimo, da je model TOČEN, če ima majhno napako na učni množici, in SPLOŠEN, če ima majhno napako na testni množici.

Vprašanje 2. Kako definiramo napako napovednega modela? Kaj sta točnost in splošnost?

Optimalen in nepristranski napovedni model bo takšen, ki bo minimiziral kvadratno napako $E((Y - m(\mathbf{X}))^2)$. V idealnem svetu bo tak model imel obliko $m^*(\mathbf{x}) = E(Y \mid \mathbf{X} = \mathbf{x})$. Če predpostavimo, da so vsa opažanja v S enako verjetna, je približek tega modela funkcija

$$m^*(\mathbf{x}_0) = \frac{1}{|S_0|} \sum_{(\mathbf{x}, y) \in S_0} y,$$

kjer je $S_0 = \{(\mathbf{x}, y) \in S_{\text{train}} \,|\, \mathbf{x} = \mathbf{x}_0\}$ množica primerov, kjer je $\mathbf{x} = \mathbf{x}_0$. Problem je v tem, da je v praksi ta množica najverjetneje prazna, ali pa vsebuje zelo malo primerov. Namesto tega lahko za S_0 vzamemo k najbližjih primerov iz S_{train}, čemur pravimo SOSEŠČINA točke $\mathbf{x}_0$.

Vprašanje 3. Opiši delovanje metode najbližjih sosedov.

7.1 Linearna regresija

Pri numeričnih napovednih spremenljivkah in numerični ciljni spremenljivki lahko uporabimo linearno regresijo,

$$\hat{y} = \beta_0 + \sum_{i=1}^{p} \beta_i X_i,$$

kjer so β_i neznani koeficienti. Če jih zložimo v vektor $\boldsymbol{\beta} = [\beta_0, \ldots, \beta_p]^T$, lahko model prepišemo v $\hat{y} = \mathbf{X}\boldsymbol{\beta}$ za $\mathbf{X} = [1, X_1, \ldots, X_p]$. Optimalna izbira za $\boldsymbol{\beta}$ bo tedaj

$$\arg\min_{\boldsymbol{\beta}} (Y - \mathbf{X}\boldsymbol{\beta})(Y - \mathbf{X}\boldsymbol{\beta})^T$$

za stolpec ciljnih vrednosti Y. Funkciji, ki jo zgoraj minimiziramo, pravimo tudi RSS. Optimalen $\boldsymbol{\beta}$ dobimo z reševanjem predoločenega sistema.

Vprašanje 4. Pojasni linearno regresijo.

Napako linearnega modela lahko merimo na več načinov. Ena možnost je RESIDUALNA STANDARDNA NAPAKA

$$\text{RSE} = \sqrt{\frac{\text{RSS}}{|S| - p - 1}},$$

kjer je podobno kot prej

$$\text{RSS} = \sum_{(\mathbf{x},y) \in S} \left(y - \mathbf{x}^T \boldsymbol{\beta}\right)^2.$$

Dober model imamo, ko je RSE blizu 0.

Drug način merjenja napake je DELEŽ POJASNJENE VARIANCE

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}}$$

za

$$\text{TSS} = \sum_{(\mathbf{x},y) \in S} (y - \bar{y})^2,$$

kjer z $\bar{y}$ označimo povprečje. Mera R^2 ima vrednosti v intervalu $[0, 1]$, dober model pa dobimo za R^2 blizu 1.

Nazadnje imamo še POPOLNO NAPAKO MODELA oziroma RMSE

$$\mathrm{RMSE} = \sqrt{\frac{1}{|S|} \sum_{(\mathbf{x},y)\in S} (y - \mathbf{x}^T\boldsymbol{\beta})^T}.$$

Tudi tu ima dober model vrednost blizu 0.

Vprašanje 5. Kako meriš napako linearnega modela?

Če vemo (ali predpostavljamo), da ima neka spremenljivka nelinearen vpliv, lahko poskusimo dodati nove spremenljivke, npr. X_1^2, X_1X_2, ipd.

7.2 Logistična regresija

Če imamo diskretno spremenljivko V z domeno $\{v_1, \ldots, v_k\}$, jo spremenimo v k numeričnih spremenljivk, ki so indikatorji dogodka $V = v_i$. Dovolj je torej znati napovedati vrednost binarne diskretne spremenljivke. V nadaljevanju napovedujemo Y z vrednostmi $D_Y = \{\oplus, \ominus\}$. Namesto vrednosti $\oplus$ in $\ominus$ lahko poskusimo napovedati verjetnost $P(Y = \oplus \,|\, X = x)$, problem pa je, da nam linearna regresija hitro zbeži izven intervala $[0,1]$. Rešitev je, da napovedujemo logaritem obetov $\log \frac{p}{1-p}$, ki leži na intervalu $(-\infty, \infty)$. Odločitvena meja pri $p = 0.5$ sovpada z vrednostjo $z = 0$.

Težava je v tem, da verjetnosti $P(Y = \oplus \,|\, X = x)$ ne poznamo, torej ne moremo formulirati najmanjših kvadratov. Deluje pa metoda največjega verjetja

$$L(\mathbf{X}, Y, \boldsymbol{\beta}) = \prod_{(\mathbf{x},y)\in S} P(Y = y \,|\, \mathbf{X} = \mathbf{x}, \boldsymbol{\beta}) = \prod_{y=1} \frac{e^{\mathbf{x}^T\boldsymbol{\beta}}}{1 - e^{\mathbf{x}^T\boldsymbol{\beta}}} \prod_{y=0} \frac{1}{1 - e^{\mathbf{x}^T\boldsymbol{\beta}}}.$$

Najlažje maksimiziramo logaritem verjetja, torej

$$\log L = \sum_{y=1} (\mathbf{x}^T\boldsymbol{\beta} - \log(1 + e^{\mathbf{x}^T\boldsymbol{\beta}})) - \sum_{y=0} \log(1 + e^{\mathbf{x}^T\boldsymbol{\beta}}).$$

Če prvo vsoto pomnožimo z y, drugo pa z $1 - y$, nič ne spremenimo, torej je to enako

$$\log L = \sum_{(\mathbf{x},y)\in S} (\mathbf{x}^T\boldsymbol{\beta}y - \log(1 + e^{\mathbf{x}^T\boldsymbol{\beta}})).$$

Maksimum funkcije poiščemo z odvodom.

Vprašanje 6. Razloži logistično regresijo.

Pri logistični regresiji lahko merimo KLASIFIKACIJSKO NAPAKO

$$\mathrm{CE} = \frac{1}{|S|} \sum_{(\mathbf{x},y)\in S} \mathbb{1}(y \neq \mathbf{x}^T\boldsymbol{\beta}),$$

kjer imamo dober model, če je ta napaka blizu 0.

Vprašanje 7. Kaj je klasifikacijska napaka?

7.3 Najbližji sosedi

V metodi najbližjih sosedov definiramo SOSEŠČINO točke $\mathbf{x}_0$ kot množico S_0 najbližjih k meritev tej točki. Razdalja je običajno psevdometrika, kjer ne zahtevamo, da iz $d(a, b) = 0$ sledi $a = b$. Pri regresijskem modelu vzamemo (uteženo) povprečje množice S_0, v klasifikaciji pa večinsko glasovanje, tj. najpogostejšo vrednost v soseščini.

Vprašanje 8. Razloži metodo najbližjih sosedov.

8 Numerična linearna algebra

8.1 Singularni razcep

Izrek. *Za vsako matriko $A \in \mathbb{R}^{m\times n}$, kjer je $m \geq n$, obstaja razcep $A = U\Sigma V^T$, kjer je $U \in \mathbb{R}^{m\times m}$ ortogonalna, $V \in \mathbb{R}^{n\times n}$ ortogonalna in $\Sigma \in \mathbb{R}^{m\times n}$ oblike*

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{bmatrix}$$

za singularne vrednosti $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$ matrike A.

Dokaz. Če je $A = U\Sigma V^T$, potem je $A^T A = V\Sigma^T\Sigma V^T$. Ta matrika je simetrična nenegativno definitna, torej so njene lastne vrednosti realne in nenegativne ter jih lahko uredimo padajoče. Lastne vektorje lahko izberemo ortonormirane in jih zložimo v stolpce V.

Iz $AV = U\Sigma$ sledi $Av_i = \sigma_i u_i$ za stolpce v_i in u_i. Če je $\sigma_i \neq 0$, lahko tak u_i poiščemo. Če je $\sigma_r > \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$, tako določimo prvih r stolpcev U. Ker je v_j lastni vektor za $A^T A$, velja

$$u_i^T u_j = \frac{1}{\sigma_i\sigma_j} v_i^T A^T A v_j = \frac{1}{\sigma_i\sigma_j}\sigma_j^2 v_i^T v_j = \delta_{ij}.$$

Sedaj imamo

$$\begin{aligned} V &= \begin{bmatrix} V_1 & V_2 \end{bmatrix}, \\ \Sigma &= \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}, \\ U &= \begin{bmatrix} U_1 & U_2 \end{bmatrix}, \end{aligned}$$

kjer smo U_2 določili tako, da dopolnimo stolpce U_1 do ortonormirane baze prostora.

Preverimo, da te matrike res tvorijo singularni razcep:

$$U^T AV = \begin{bmatrix} U_1^T AV_1 & U_1^T AV_2 \\ U_2^T AV_1 & U_2^T AV_2 \end{bmatrix}.$$

Velja $(AV_2)^T AV_2 = V_2^T A^T AV_2 = V_2^T \cdot 0 = 0$, ker so stolpci V_2 lastni vektorji za lastno vrednost 0. Iz definicije u_i pa vidimo, da je $U_1^T AV_1 = S$. Velja $AV_1 = [\sigma_1 u_1, \ldots, \sigma_r u_r]$, ti stolpci so pravokotni vrednostim u_{r+j} po definiciji. □

Če je $m < n$, singularni razcep dobimo tako, da transponiramo razcep za A^T. V dokazu je r rang matrike A. Numerično je singularni razcep najboljše orodje za računanje ranga; dobimo tudi bazi za jedro in sliko matrike A (stolpci V_2 in U_1).

Vprašanje 1. Pokaži, da za vsako matriko obstaja singularni razcep.

Če je A ranga $r < n$, potem za vektor x, ki reši predoločen sistem $Ax = b$, izberemo tistega, ki minimizira $\|Ax - b\|_2$ (ta ni več enolično določen), in ki ima med takimi minimalno normo $\|x\|_2$.

Izrek. *Naj bo $A \in \mathbb{R}^{m\times n}$, $A = U\Sigma V^T$ ranga r. Potem je rešitev predoločenega sistema $Ax = b$ enaka*

$$x = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i.$$

Dokaz. Zapišimo $U = [U_1 U_2]$, $V = [V_1 V_2]$ in $\Sigma = \operatorname{diag}(S, 0)$, kjer so prve komponente širine r, druge pa širine $n - r$. Potem je

$$\|Ax - b\|_2 = \left\|U\Sigma V^T x - b\right\|_2 = \left\|\Sigma V^T x - U^T b\right\|_2 = \left\|\begin{bmatrix} Sy_1 - U_1^T b \\ -U_2^T b \end{bmatrix}\right\|$$

za $V^T x = y = (y_1, y_2)$. Minimum bo dosežen, če je $Sy_1 = U_1^T b$, pri čemer je y_2 lahko poljuben. Velja $\|x\|_2 = \|y\|_2$, torej izberemo $y_2 = 0$. □

Vprašanje 2. Kakšna je rešitev problema najmanjših kvadratov $Ax = b$ za matriko $A \in \mathbb{R}^{m\times n}$? Dokaži.

Definicija. Za matriko $A \in \mathbb{R}^{m\times n}$ je PSEVDOINVERZ matrika $X \in \mathbb{R}^{n\times m}$, ki zadošča naslednjim točkam:

- $AXA = A$
- $XAX = X$
- $(AX)^T = AX$
- $(XA)^T = XA$

Označimo $X = A^+$.

Če je rang matrike A enak $r < n$, potem je

$$A^+ = (A^T A)^{-1} A^T.$$

Vprašanje 3. Definiraj psevdoinverz. Čemu je enak?

Izrek. *Naj bo $A = U\Sigma V^T \in \mathbb{R}^{m\times n}$ in $\operatorname{rang} A = r$. Potem je psevdoinverz enak*

$$A^+ = \sum_{i=1}^{r} \frac{1}{\sigma_i} v_i u_i^T$$

oziroma $A^+ = V\Sigma^+ U^T$ za $\Sigma^+ = \operatorname{diag}(S^{-1}, 0)$.

Dokaz. Iščemo $X \in \mathbb{R}^{m\times n}$, ki zadošča Moore-Penroseovim pogojem. Velja $X = VYU^T$ za nek Y. Veljati mora $(AX)^T = AX$, iz česar z zapisom po blokih dobimo, da je zgornji desni blok Y enak 0; podobno iz drugih pogojev izpeljemo enakosti v ostalih blokih. □

Vprašanje 4. Kako izračunaš psevdoinverz s pomočjo singularnega razcepa?

9 Statistika

9.1 Centralni limitni izrek

Naj bo $S_n \sim \mathrm{Bin}(n, \frac{1}{2})$, da velja $P(S_n = k) = \binom{n}{k} 2^{-n}$. Poglejmo si razmerje

$$\frac{P(S_n = k+1)}{P(S_n = k)} = \frac{\binom{n}{k+1}}{\binom{n}{k}} = \frac{n-k}{k+1}.$$

Za sedaj se omejimo na sode $n = 2m$. Potem je $P(S_n = k)$ največja za $k = m$, in za $d > 0$ dobimo

$$\begin{aligned}
\frac{P(S_{2m} = m+d)}{P(S_{2m} = m)} &= \frac{P(S_{2m} = m+d)}{P(S_{2m} = m+d-1)} \frac{P(S_{2m} = m+d-1)}{P(S_{2m} = m+d-2)} \cdots \frac{P(S_{2m} = m+1)}{P(S_{2m} = m)} \\
&= \frac{2m-m-d+1}{m+d} \cdots \frac{2m-m}{m+1} \\
&= \frac{m-d+1}{m-d} \cdots \frac{m}{m+1} \\
&= \frac{1+\frac{1-d}{m}}{1+\frac{d}{m}} \cdots \frac{1}{1+\frac{1}{m}} \\
&= \frac{1-\frac{d-1}{m}}{1+\frac{1}{m}} \cdots \frac{1}{1+\frac{d}{m}}.
\end{aligned}$$

Če je d dovolj majhen, je to približno enako

$$\begin{aligned}
\frac{P(S_{2m} = m+d)}{P(S_{2m} = m)} &\approx \frac{e^{-\frac{d-1}{m}} e^{-\frac{d-2}{m}} \cdots e^0}{e^{\frac{1}{m}} e^{\frac{2}{m}} \cdots e^{\frac{d}{m}}} \\
&= \exp\left(-\frac{1}{m} - \frac{2}{m} - \cdots - \frac{d}{m} - \frac{d-1}{m} - \cdots - \frac{1}{m}\right) \\
&= \exp\left(-\frac{d^2}{m}\right).
\end{aligned}$$

Pri $d < 0$ je porazdelitev simetrična, torej dobimo enak rezultat. Izkaže se

$$\lim_{m\to\infty} \sum_{d=-\infty}^{\infty} \left| \frac{P(S_{2m} = m+d)}{P(S_{2m} = m)} - e^{-d^2/m} \right| = 0.$$

Vrsta

$$\sum_{d=-\infty}^{\infty} \frac{P(S_{2m} = m+d)}{P(S_{2m} = m)} = \frac{1}{P(S_{2m} = m)}$$

je približno enaka

$$\sum_{d=-\infty}^{\infty} e^{-d^2/m} \approx \int_{-\infty}^{\infty} e^{-x^2/m} dx = \sqrt{m\pi},$$

torej je verjetnost $P(S_{2m} = m) \approx \frac{1}{\sqrt{m\pi}}$. Upoštevaje zgornjo formulo potem dobimo

$$P(S_{2m} = m+d) \approx \frac{1}{\sqrt{m\pi}} e^{-\frac{d^2}{m}}$$

oziroma

$$P(S_n = k) \approx \sqrt{\frac{2}{n\pi}} \exp\left(-\frac{2(k - {}^n\!/_2)^2}{n}\right).$$

Vse to je res tudi za lihe n. Rezultatu pravimo DE MOIVREOVA LOKALNA FORMULA.

Vprašanje 1. Izpelji de Moivreovo lokalno formulo.

Za $a, b \in \mathbb{Z}$ lahko izpeljemo

$$\begin{aligned}
P(a \leq S_n \leq b) &= \sum_{k=a}^{b} P(S_n = k) \\
&\approx \sqrt{\frac{2}{n\pi}} \sum_{k=a}^{b} \exp\left(-\frac{2(k - \frac{m}{2})^2}{n}\right) \\
&\approx \sqrt{\frac{2}{n\pi}} \sum_{k=a}^{b} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \exp\left(-\frac{2(x - \frac{m}{2})^2}{n}\right) dx \\
&= \sqrt{\frac{2}{n\pi}} \int_{a}^{b} \exp\left(-\frac{2(x - \frac{m}{2})^2}{n}\right) dx,
\end{aligned}$$

kar se da razširiti za poljubna $a < b$. Dobljeni približek za verjetnost $P(S_n = k)$ je gostota normalne porazdelitve $N(\frac{n}{2}, \frac{n}{4})$ v točki k. Velja $E(S_n) = \frac{n}{2}$ in $\mathrm{var}(S_n) = \frac{n}{4}$. Binomska porazdelitev, s katero smo začeli, je enaka vsoti n neodvisnih Bernoullijevih slučajnih spremenljivk, ki je torej porazdeljena približno normalno.

Izrek (centralni limitni izrek)**.** *Naj bodo $X_1, X_2, \ldots$ neodvisne in enako porazdeljene slučajne spremenljivke s končnim drugim momentom. Označimo $\mu_1 = E(X_i)$ in $\sigma_1^2 = \mathrm{var}(X_i)$. Za $S_n = X_1 + \cdots + X_n$ veljata naslednji točki:*

- *Če definiramo $W_n = \frac{S_n - n\mu_1}{\sigma_1\sqrt{n}}$, za vsak $w \in \mathbb{R}$ velja*

$$\lim_{n\to\infty} P(W_n \leq w) = \phi(w).$$

- *Limita*

$$\lim_{n\to\infty} \sup_{w\in\mathbb{R}} |P(W_n < w) - \phi(w)| = \lim_{n\to\infty} \sup_{x\in R} \left| P(S_n < x) - \phi\left(\frac{x - n\mu_1}{\sigma_1\sqrt{n}}\right) \right| = 0.$$

Vprašanje 2. Formuliraj centralni limitni izrek.

Dokažemo lahko tudi drugačno formulacijo: za vsako zvezno in omejeno funkcijo $h : \mathbb{R} \to \mathbb{R}$ z največ kvadratno rastjo (tj. obstaja M, da je $|h(x)| \leq M(1 + x)^2$) velja

$$\lim_{n\to\infty} E(h(W_n)) = E(h(Z)),$$

kjer je $Z \sim N(0, 1)$. Funkcijam h pravimo POIZKUSNE ali TESTNE funkcije.

Trditev (Jensenova neenakost). *Naj bo X slučajna spremenljivka z vrednostmi na intervalu $I \subseteq \mathbb{R}$ in $\varphi : I \to \mathbb{R}$ konveksna funkcija. Tedaj je $\varphi(E(X)) \leq E(\varphi(X))$.*

Ideja dokaza je, da poiščemo linearno funkcijo, ki se v iskani točki dotika grafa φ, in vrednost aproksimiramo s pomočjo nje. Trditev je pomembna zaradi posledice.

Posledica. *Če je $p \geq q$, je $E(|X|^q) \leq (E(|X|^p))^{q/p}$.*

Posledica. *Če obstaja p-ti moment in je $q \leq p$, obstaja tudi q-ti moment.*

Za nadaljevanje naj bodo $Y_1, \ldots, Y_n$ neodvisne z vsoto S in $Z_1, \ldots, Z_n$ neodvisne z vsoto T. Privzemimo še, da tretji momenti Y_k in Z_k obstajajo, da je $E(Y_k) = E(Z_k) = 0$ in da je $\mathrm{var}(Y_k) = \mathrm{var}(Z_k) = \sigma_k^2$. Tedaj je $E(T) = E(S) = 0$ in

$$\mathrm{var}(S) = \mathrm{var}(T) = \sum_k \sigma_k^2.$$

Za primerno $h : \mathbb{R} \to \mathbb{R}$ bomo ocenili razliko $E(h(S)) - E(h(T))$. Privzeli bomo, da je $h \in \mathcal{C}^3$ in $M_3 = \sup_{x\in\mathbb{R}} |h'''(X)| < \infty$. Definirajmo kombinirane vsote $V_k = Y_1 + \ldots + Y_k + Z_{k+1} + \ldots + Z_n$.

Dodatno privzemimo, da so Y_i in Z_i vsi medsebojno neodvisni. Velja

$$E(h(S)) - E(h(T)) = \sum_{k=1}^n E(h(V_k) - h(V_{k-1})) = \sum_{k=1}^n E(h(U_k + Y_k) - h(U_k + Z_k))$$

za $U_k = Y_1 + \ldots + Y_{k-1} + Z_{k+1} + \ldots + Z_n$. Velja

$$h(V_k) = h(U_k) + h'(U_k)Y_k + \frac{1}{2}h''(U_k)Y_k^2 + R_k,$$

kjer je R_k omejen z $\frac{1}{6}M_3 \left|Y_k\right|^3$. Poleg tega je

$$h(V_{k-1}) = h(U_k) + h'(U_k)Z_k + \frac{1}{2}h''(U_k)Z_k^2 + \tilde{R}_k,$$

kjer je $\tilde{R}_k$ podobno omejena. Zaradi neodvisnosti U_k, Y_k in Z_k potem dobimo

$$E(h(V_k) = E(h(U_k)) + E(h'(U_k))E(Y_k) + \frac{1}{2}E(h''(U_k))E(Y_k^2) + E(R_k),$$
$$E(h(V - k - 1)) = E(h(U_k)) + E(h'(U_k))E(Z_k) + \frac{1}{2}E(h''(U_k))E(Z_k^2) + E(\tilde{R}_k).$$

Torej

$$|E(h(V_k) - h(V_{k-1}))| \leq E(R_k) + E(\tilde{R}_k) \leq \frac{1}{6}M_3(E(|Y_k|^3) + E(|Z_k|^3)),$$

iz česar naposled dobimo

$$|E(h(S)) - E(h(T))| \leq \frac{1}{6}M_3 \sum_{k=1}^n \left(E(|Y_k|^3) + E(|Z_k|^3)\right).$$

Vprašanje 3. Izpelji oceno za razliko $|E(h(S)) - E(h(T))|$, kjer je $S = Y_1 + \ldots + Y_n$ in $T = Z_1 + \ldots + Z_n$, in so Y_i enako porazdeljeni neodvisni, ter Z_i enako porazdeljeni in neodvisni z $E(Y_i) = E(Z_i) = 0$, $\mathrm{var}(Y_i) = \mathrm{var}(Z_i) = \sigma_i^2$.

Če vzamemo $Z_k \sim N(0, \sigma_k^2)$, je $T \sim N(0, \sigma^2)$. Tedaj lahko z integralom izračunamo

$$E(|Z_k|^3) = \sigma_k^3 \frac{4}{\sqrt{2\pi}} = \frac{4}{\sqrt{2\pi}} \left(E(Y_k^2)\right)^{3/2} \leq \frac{4}{\sqrt{2\pi}} E(|Y_k|^3),$$

kjer smo v zadnjem koraku uporabili posledico Jensenove neenakosti. V tem primeru torej lahko ocenimo

$$|E(h(S)) - E(h(T))| \leq \left(\frac{1}{6} + \frac{2}{3\sqrt{2\pi}}\right) M_3 \sum_{k=1}^{n} E(|Y_k|^3).$$

Tu žal ne moramo vzeti $h(w) = \mathbb{1}(w \leq a)$. Velja pa naslednji izrek.

Izrek. *Obstaja taka univerzalna konstanta C, da za poljubne neodvisne slučajne spremenljivke $Y_1, \ldots, Y_n$, za katere je $E(Y_k) = 0$, velja*

$$\sup_{x\in\mathbb{R}} \left| P(S \leq x) - \phi\Big(\frac{x}{\sigma}\Big) \right| = \sup_{w\in\mathbb{R}} \left| P\left(\frac{S}{\sigma} < w\right) - \phi(w) \right| \leq \frac{C}{\sigma^3} \sum_{k=1}^{n} E(|Y_k|^3),$$

pri čemer je $S = Y_1 + \ldots + Y_n$ in $\sigma^2 = \mathrm{var}(S) = \sum_k \mathrm{var}(Y_k)$.

Naj bodo $X_1, X_2, \ldots$ neodvisne in enako porazdeljene z $E(X_1) = \mu_1$ ter $\mathrm{var}(X_1) = \sigma_1^2$ in $E(|X_1 - \mu_1|^3) = \gamma^3$. Definiramo $S_n = X_1 + \ldots + X_n$ in

$$W_n = \frac{S_n - n\mu_1}{\sigma_1\sqrt{n}}.$$

Za fiksen n označimo

$$Y_k = \frac{X_k - \mu_1}{\sigma_1\sqrt{n}},$$

da je $E(Y_k) = 0$ in $W_n = \sum_k Y_k$. Velja $\sigma^2 = \mathrm{var}(W_n) = 1$ in

$$\sum_{k=1}^{n} E(|Y_k|^3) = nE(|Y_1|^3) = \frac{\gamma_1^3}{\sigma_1^3\sqrt{n}} \xrightarrow[n\to\infty]{} 0.$$

Posledica. *S temi oznakami*

$$\sup_{x\in\mathbb{R}} \left| P(S_n < x) - \phi\left(\frac{x - n\mu_1}{\sigma_1\sqrt{n}}\right) \right| = \sup_{w\in\mathbb{R}} |P(W_n < w) - \phi(w)| \leq \frac{C}{\sqrt{n}} \frac{\gamma_1^3}{\sigma_1^3},$$

za vsako funkcijo $h \in \mathcal{C}^3(\mathbb{R})$ pa velja še

$$|E(h(W_n)) - E(h(Z))| \leq \left(\frac{1}{6} + \frac{2}{3\sqrt{2\pi}}\right) \frac{\gamma_1^3}{\sigma_1^3} \frac{M_3}{\sqrt{n}}.$$

Vprašanje 4. Kaj lahko poveš o oceni razlike porazdelitve vsote od normalne porazdelitve?

9.2 Konvergenca porazdelitev

Definicija. Zaporedje realnih slučajnih spremenljivk $X_1, X_2, \ldots$ V PORAZDELITVI KONVERGIRA proti slučajni spremenljivki X, če za vsak $x \in \mathbb{R}$, za katerega je $P(X = x) = 0$, velja

$$\lim_{n\to\infty} P(X_n \leq x) = P(X \leq x).$$

Pišemo $X_n \xrightarrow[n\to\infty]{d} X$.

Opomba. Temu pravimo tudi ŠIBKA KONVERGENCA.

Definicija. Zaporedje slučajnih spremenljivk $X_1, X_2, \ldots$ z vrednostmi v topološkem prostoru S V PORAZDELITVI KONVERGIRA PROTI X, če za vsako zvezno in omejeno funkcijo $h : S \to \mathbb{R}$ velja

$$\lim_{n\to\infty} E(h(X_n)) = E(h(X)).$$

Funkcijam h rečemo tudi PREIZKUSNE ali TESTNE funkcije.

Izrek (Helly-Bray)**.** *Za porazdelitve na realni osi z običajno topologijo definiciji sovpadata.*

Izrek (Centralni limitni izrek)**.** *Če so $X_1, X_2, \ldots$ neodvisne in enako porazdeljene z $E(X_1) = \mu_1$ in $\operatorname{var}(X_1) = \sigma_1^2$, velja*

$$\frac{X_1 + \ldots + X_n - n\mu_1}{\sqrt{n}} \xrightarrow[n\to\infty]{d} N(0, \sigma_1^2).$$

Vprašanje 5. Definiraj konvergenco v porazdelitvi in reformuliraj centralni limitni izrek.

Trditev. *Naj bodo $X_1, X_2, \ldots, X$ slučajne spremenljivke z vrednostmi v topološkem prostoru S in T še en topološki prostor. Naj bo $g : S \to T$ zvezna. Če velja $X_n \xrightarrow[n\to\infty]{d} X$, velja tudi $g(X_n) \xrightarrow[n\to\infty]{d} g(X)$.*

Dokaz. Naj bo $h : T \to \mathbb{R}$ zvezna in omejena. Tedaj je tudi $h \circ g : S \to \mathbb{R}$ zvezna in omejena, zato je

$$E(h(g(X))) = \lim_{n\to\infty} E(h(g(X_n))) = \lim_{n\to\infty} E(h \circ g(X_n)) = E(h \circ g(X)).$$

□

Trditev. *Naj bo $S_0 \subseteq S$ in naj imajo $X_1, X_2, \ldots, X$ vrednosti v S_0. Če gre $X_n \xrightarrow[n\to\infty]{d} X$ v okviru S_0, to velja tudi v okviru prostora S.*

Dokaz. Naj bo $h : S \to \mathbb{R}$ zvezna in omejena, ter naj bo h_0 zožitev h na S_0. Tudi ta funkcija je zvezna in omejena, zato je $\lim_{n\to\infty} E(h_0(X_n)) = E(h_0(X))$. □

Trditev. *Naj bo S metrizabilen prostor, $S_0 \subseteq S$ pa zaprt podprostor. Če imajo $X_1, X_2, \ldots, X$ vrednosti v S_0 in $X_n \xrightarrow[n\to\infty]{d} X$ v okviru S, to velja tudi v okviru S_0.*

Dokaz. Vsaka zvezna omejena funkcija $h_0 : S_0 \to \mathbb{R}$ se da po Tietzejevem izreku razširiti do zvezne in omejene funkcije $h : S \to \mathbb{R}$. □

Vprašanje 6. V katerem primeru se konvergenca v porazdelitvi razširi iz večjega prostora v podprostor? Kaj je ideja dokaza?

Izrek. *Naj bo S metrizabilen in $S_0 \subseteq S$ odprt podprostor. Naj bodo $X_1, X_2, \ldots$ slučajne spremenljivke z vrednostmi v S in X z vrednostmi v S_0. Naj gre $X_n \xrightarrow[n\to\infty]{d} X$ v okviru prostora S. Nadalje naj bodo $X_1^*, X_2^*, \ldots$ slučajne spremenljivke z vrednostmi v S_0 ter naj bo $X_n^*(\omega) = X_n(\omega)$ brž ko je $X_n(\omega) \in S_0$. Tedaj gre $X_n \xrightarrow[n\to\infty]{d} X$ tudi v okviru S_0.*

Trditev. *Naj bodo $X_1, X_2, \ldots$ slučajne spremenljivke z vrednostmi v metričnem prostoru (S, d). Naj bo $c \in S$ konstantna. Tedaj gre $X_n \xrightarrow[n\to\infty]{d} c$ natanko tedaj, ko za vsak $r > 0$ velja*

$$\lim_{n\to\infty} P(d(X_n, c) \geq r) = 0.$$

Dokaz. V desno: Obstaja taka zvezna funkcija $h : S \to [0, 1]$, da je $h(c) = 0$ in $h(x) = 1$ za $d(x, c) \geq r$. Velja $\mathbb{1}(d(x, c) \geq r) \leq h(x)$, na tej neenakosti pa lahko uporabimo pričakovano vrednost in dobimo

$$P(d(X_n, c) \geq r) \leq E(h(X_n)) \xrightarrow[n\to\infty]{} E(h(c)) = 0.$$

V levo: Naj bo $h : S \to [0, 1]$ zvezna in $\varepsilon > 0$. Obstaja tak $r > 0$, da za vsak $x \in S$ z $d(x, c) \geq r$ velja $|h(x) - h(c)| < \varepsilon/2$. Sledi

$$\begin{aligned}
&|E(h(X_n)) - h(c)| \\
&\leq E(|h(X_n) - h(c)|) \\
&= E(|h(X_n) - h(c)|\,\mathbb{1}(d(X_n, c) < r)) + E(|h(X_n) - h(c)|\,\mathbb{1}(d(X_n, c) \geq r)) \\
&\leq \frac{\varepsilon}{2} + P(d(X_n, c) \geq r).
\end{aligned}$$

Za dovolj pozne n bo to manj od ε. □

Vprašanje 7. Kdaj zaporedje slučajnih spremenljivk konvergira k konstanti? Dokaži.

Trditev (Neenačba Markova)**.** *Za nenegativno slučajno spremenljivko W in poljuben $a > 0$ velja*

$$P(W \geq a) \leq \frac{E(W)}{a}.$$

Dokaz. Velja $\mathbb{1}(w \geq a) \leq w/a$. Na tem uporabimo pričakovano vrednost. □

Vprašanje 8. Povej in dokaži neenačbo Markova.

Trditev. *Če je* $p > 0$ *in* $\lim_{n\to\infty} E((d(X_n, c))^p) = 0$*, gre* $X_n \xrightarrow[n\to\infty]{d} c$*.*

Dokaz. Velja $P(d(X_n, c) \geq r) = P((d(X_n, c))^p \geq r^p)$. Na drugem delu uporabimo neenačbo Markova. □

Trditev (šibki zakon velikih števil)**.** *Če so* $X_1, X_2, \ldots$ *neodvisne in enako porazdeljene z* $E(X_1^2) < \infty$ *in* $E(X_1) = \mu$*, gre*

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{d} \mu.$$

Dokaz. Pričakovana vrednost izraza je očitno μ. Potem je

$$E\left(\left(\frac{X_1 + \cdots + X_n}{n} - \mu\right)^2\right) = \operatorname{var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\operatorname{var}(X_1)}{n},$$

to pa konvergira k 0 za $n \to \infty$. □

Vprašanje 9. Povej in dokaži šibki zakon velikih števil.

Opomba. Velja celo

$$P\left(\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mu\right) = 1.$$

Temu pravimo KREPKI ZAKON VELIKIH ŠTEVIL.

Izrek. *Naj bo* S *metrizabilen prostor, ki je števna unija svojih kompaktnih podprostorov. Naj bodo* $X_1, X_2, \ldots$ *slučajne spremenljivke z vrednostmi v* S*. Naj bo* T *še en metrizabilen prostor in naj imajo* $Y_1, Y_2, \ldots$ *vrednosti v* T*. Če gre* $X_n \xrightarrow[n\to\infty]{d} X$ *in* $Y_n \xrightarrow[n\to\infty]{d} c$*, kjer je* $c \in T$ *konstanta, gre tudi* $(X_n, Y_n) \xrightarrow[n\to\infty]{d} (X, c)$*.*

Vprašanje 10. Kaj velja za konvergenco parov slučajnih spremenljivk?