1 skyrius

Matematiniai modeliai ir jų korektiškumas

1.1. Matematinių uždavinių klasifikacija

Matematinis modeliavimas yra svarbus naujas žinių gavimo būdas, kuris vis dažniau naudojamas sprendžiant technologinius uždavinius, tiriant gamtos ir socialinius procesus. Taikydami šį metodą naudojame naujausius matematikos pasiekimus, kita vertus matematinis modeliavimas formulavo ir formuluoja naujus matematikos uždavinius, įtakoja pačios matematikos vystymąsi.

Matematinių modelių bendroji schema. Pirmiausia aptarsime labai bendrą matematinių uždavinių sprendimo schemą. Paaiškinsime, kokius reikalavimus turėtų tenkinti bet koks matematinis modelis, kad jį galėtume naudoti modeliuodami realius procesus, technologinius įrenginius ir socialinius reiškinius.

Tarkime, kad turime *pradinius* duomenis u, tada naudodami pasirinktąjį algoritmą *apibrėžiame* uždavinio sprendinį

$$v = R(u).$$

Šiame apibrėžime kol kas nepatiksliname, kaip sudaromas matematinis modelis ir jį realizuojantis algoritmas (šį apibrėžimą tenkina netgi tokie egzotiški algoritmai, kaip astrologinės prognozė ar atsitiktinis spėjimas).

1.1 pavyzdys. Tiesinių lygčių sistema. Spręsime tiesinių lygčių sistemą

$$AX = F,$$

čia pradinius duomenis sudaro $n \times n$ dydžio matrica A ir vektorius F, o reikia rasti vektorių X. Naudodami bendrąją schemą, pažymėsime

$$u = (A, F), \quad v = X.$$

Uždavinio sprendinį apibrėžiame tokiu būdu

$$X = A^{-1}F,$$

kur A^{-1} yra atvirkštinė matrica, tenkinanti sąlygą

$$AA^{-1} = I.$$

Jokio konkretaus algoritmo, kaip skaičiuoti atvirkštinę matricą kol kas nenurodome.

Sakykime, kad pradiniai duomenys u priklauso metrinei erdvei U, o sprendinys v – metrinei erdvei V. Atstumus tarp dviejų elementų šiose erdvėse atitinkamai žymėsime

$$\rho_U(u_1, u_2), \quad \rho_V(v_1, v_2).$$

1.1.1. Korektiškai suformuluotas uždavinys

Apibrėžiame du pirmuosius reikalavimus, kuriuos turi tenkinti *korektiškas* matematinis uždavinys.

1. Kiekvienam elementui $u \in U$ egzistuoja sprendinys $v = R(u) \in V$.
2. Toks sprendinys yra vienintelis.

Pavyzdžiui, jei matricos A determinantas nelygus nuliui, tai tiesinių lygčių sistema turi vienintelį sprendinį bet kokiems vektoriams F. Tačiau jei $\det A = 0$, tai tokia sistema turi sprendinį (netgi ne vienintelį) tik tam tikriems vektoriams F.

Trečiasis reikalavimas yra tiesiogiai susijęs su matematinio modeliavimo poreikiais. Daugelis reiškinių pasižymi savybe, kad, esant panašioms pradinėms sąlygoms, stebime labai artimus rezultatus. Tokią savybę dažnai vadiname stabilumu. Todėl natūralu reikalauti, kad ir atitinkami matematiniai modeliai būtų *stabilūs*. Stabilumas tiesiogiai siejasi su sprendinio tolydžia priklausomybe nuo pradinių duomenų.

Apibrėžimas. Sprendinio $v = R(u) \in V,\ u \in U$ radimo uždavinys vadinamas *stabiliu* metrinėse erdvėse (V, U), jei bet kokiam $\varepsilon > 0$ egzistuoja toks

$\delta(\varepsilon) > 0$, kad imdami $u_1, u_2 \in U$, tenkinančius nelygybę $\rho_U(u_1, u_2) \le \delta(\varepsilon)$, gauname įvertį $\rho_V(v_1, v_2) \le \varepsilon$, čia

$$v_1 = R(u_1), \quad v_2 = R(u_2), \quad v_1, v_2 \in V.$$

Tada formuluojame trečiąjį korektiškų uždavinių reikalavimą.

3. Uždavinys yra stabilus metrinėse erdvėse (V, U).

Tokia savybė labai svarbi, kai uždavinio sprendinio ieškome skaitiniais metodais. Tada tenka aproksimuoti pradinius duomenis ir uždavinio operatorių, tačiau šią paklaidą galime padaryti kiek norima maža $\delta(\varepsilon)$. Jei uždavinys yra stabilus, tai gauname tikslaus sprendinio artinį, kuris skiriasi nuo tikslaus sprendinio tik dydžiu ε.

Uždavinio korektiškumas priklauso nuo metrinių erdvių poros (V, U) parinkimo. Vienose erdvėse jis yra korektiškas, kitose – ne. Tai susiję su informacijos apie uždavinio pradinius duomenis ir sprendinį kiekiu. Tokius pavyzdžius pateiksime kitame skirsnyje.

Korektiškų uždavinių apibrėžimą pirmasis suformulavo prancūzų matematikas Žakas Adamaras (*Hadamard*).

1.1.2. Nekorektiškų uždavinių pavyzdžiai

Egzistuoja daug matematinių modelių, aprašančių mus dominančius reiškinius, bet netenkinančių pateiktojo korektiškų uždavinių apibrėžimo. Tai nereiškia, kad mes nespręsime tokių uždavinių, tik būtina patikslinti uždavinio apibrėžimą, nes klasikiniai sprendimo metodai jau nepakankami. Svarbu mokėti *pažinti* nekorektiškus uždvinius ir sukurti specialius jų sprendimo metodus.

Pirmiausia pateiksime *bendrąją schemą*, tinkamą daugeliui uždavinių, sprendžiamų matematinio modeliavimo metodu. Mus domina uždavinio

$$Av_T = u_T, \quad u_T \in U, \ \ v_T \in V$$

sprendinys. Tarkime, kad tikslus sprendinys $v_T \in V$ egzistuoja ir yra vienintelis. Dažniausiai vietoj tikslių pradinių duomenų u_T turime tik jų artinį $u \in U$, pavyzdžiui, šiuos duomenis gauname atlikdami eksperimentus ar matavimus. Todėl neišvengiamai atsiranda pradinių duomenų paklaidos, jas įvertiname metrinės erdvės U, kurioje apibrėžti pradiniai duomenys, metrikoje:

$$\rho_U(u, u_T) \le \delta, \quad \delta > 0.$$

Todėl realiai sprendžiame modifikuotą uždavinį

$$Av = u.$$

Tada gali susidaryti kelios nepalankios situacijos:

- Tokiems pradiniams duomenims $u \in U$ uždavinio sprendinys

$$v = A^{-1}u \in V$$

  neegzistuoja, t.y. $u \notin AV$. Reikia sugalvoti, kaip apibrėžti *apibendrintąjį* arba *kvazi-* sprendinį $v = R(u)$, kuris egzistuotų visiems pradiniams duomenims $u \in U$ ir būtų tam tikra prasme artimas tiksliam sprendiniui v_T.

- Jei sprendinys (gal būt tik apibendrintasis) egzistuoja, jis gali nekonverguoti į tikslų sprendinį v_T, kai $\delta \to 0$, t.y. mažus pradinių duomenų pokyčius atitinka kiek norima dideli sprendinio pokyčiai.

Dabar pateiksime pavyzdžius uždavinių, kuriuos jau nagrinėjome matematinės analizės kurse, ir įrodysime, kad jie nekorektiški.

1. Išvestinės skaičiavimas. Reikia rasti diferencijuojamos funkcijos $u_T(t) \in C^1(D)$, $D = [0, 2\pi]$ išvestinę, kai šios funkcijos reikšmes žinome tik su paklaida (atsirandančia, pavyzdžiui, dėl matavimo prietaiso skiriamosios gebos). Pažymėkime $v_T(t) = u'_T(t)$ funkcijos $u_T(t)$ išvestinę. Tarkime, kad vietoj funkcijos $u_T(t)$ turime tik jos artinį

$$u(t) = u_T(t) + \varepsilon \sin(\omega t), \quad \varepsilon, \omega > 0.$$

Nagrinėkime tolydžių funkcijų metrinę erdvę $C(D)$, kai atstumas apibrėžiamas taip:

$$\rho_C(u_1, u_2) = \max_{t \in D} |u_1(t) - u_2(t)|.$$

Šioje metrinėje erdvėje atstumas tarp funkcijų $u(t)$ ir $u_T(t)$ yra

$$\rho_C(u, u_T) \le \varepsilon.$$

Taigi mūsų turimi pradiniai duomenys $u(t)$ mažai skiriasi nuo tikslių pradinių duomenų $u(t)$, kai parametras ε yra mažas skaičius.

Pirmiausia pastebėsime, kad išvestinę $v(t) = u'(t)$ galime apskaičiuoti bet kokioms parametrų ε, ω reikšmėms, t.y. nereikia ieškoti apibendrintojo sprendinio (šiuo atveju apibendrintos išvestinės) apibrėžimo.

Tarkime, kad metrinėje erdvėje V naudojame tą pačią metriką $C(D)$. Tada apskaičiuotoji išvestinės reikšmė skiriasi nuo tikslios išvestinės dydžiu

$$\rho_C(v, v_T) = \max_{t \in D} |\varepsilon\omega \cos(\omega t)| = \varepsilon\omega.$$

Parinkus pakankamai didelį parametrą ω, šis skirtumas gali būti kiek norima didelis, net kai ε yra mažas skaičius. Taigi išvestinės skaičiavimas tokioje metrinių erdvių poroje (U, V) yra nekorektiškas uždavinys.

Pasirinkę kitą metrinę erdvę U, galime gauti ir korektišką uždavinį. Pavyzdžiui, tarkime kad $u(t) \in C^1(D)$ yra tolydžiai diferencijuojamos funkcijos ir atstumas tarp dviejų funkcijų apibrėžiamas taip:

$$\rho_{C^1}(u_1, u_2) = \max_{t \in D} \big(|u_1(t) - u_2(t)| + |u_1'(t) - u_2'(t)|\big).$$

Kadangi teisingas įvertis

$$\rho_C(v, v_T) \le \rho_{C^1}(u, u_T),$$

tai sprendinys $v(t) = u'(t)$ tolydžiai priklauso nuo pradinių duomenų. Įrodėme, kad išvestinės skaičiavimas yra korektiškas uždavinys, kai turime metrinių erdvių porą $(U, V) = (C^1(D), C(D)$.

Tačiau patikrinti, kad eksperimentiniai duomenys $u(t)$ priklauso tokiai metrinei erdvei $U = C^1(D)$ yra sudėtinga, taigi ši metrinė erdvė netinkama, kai modeliuojame realius fizinius arba technologinius procesus.

2. Furjė eilutės sumavimas. Skaičiuokime eilutės

$$f_T(t) := \sum_{n=0}^{\infty} a_n \cos(nt), \quad 0 \le t \le \pi$$

sumą. Šio uždavinio pradinius duomenis sudaro funkcijos $f_T(t)$ Furjė koeficientai $u_T = \{a_n, n = 0, 1, \ldots\} \in U$. Erdvėje U apibrėžiame l_2 atstumo funkciją:

$$\rho_{l_2}(u_1, u_2) = \Big(\sum_{n=0}^{\infty} (a_{1n} - a_{2n})^2\Big)^{1/2}.$$

Tarkime, kad vietoj u_T turime koeficientus

$$c_n = a_n + \frac{\varepsilon}{n}, \quad n \ge 1, \quad c_0 = a_0,$$

tada sumuojame eilutę

$$f(t) := \sum_{n=0}^{\infty} c_n \cos(nt).$$

Metrinėje erdvėje l_2 abiejų eilučių koeficientai skiriasi dydžiu:

$$\varepsilon_1 := \rho_{l_2}(u_T, u) = \Big(\sum_{n=0}^{\infty}(a_n - c_n)^2\Big)^{1/2} = \varepsilon\Big(\sum_{n=1}^{\infty}\frac{1}{n^2}\Big)^{1/2} = \varepsilon\sqrt{\frac{\pi^2}{6}}.$$

Pasirinkę mažą skaičių ε, šį skirtumą galime padaryti kiek norima mažu, t.y. pradiniai duomenys bus artimi vienas kitam. Tačiau eilučių sumos skirsis dydžiu

$$f(t) - f_T(t) = \varepsilon\sum_{n=1}^{\infty}\frac{1}{n}\cos(nt).$$

Taigi atstumas $\rho_{l_2}(f, f_T)$ negali būti padarytas kiek norima mažu parenkant mažą ε, nes dešinėje lygybės pusėje esanti eilutė diverguoja, jeigu $t = 0$:

$$\varepsilon\sum_{n=1}^{N}\frac{1}{n} \approx \varepsilon\ln N \to \infty, \quad N \to \infty.$$

Taigi, jei $V = C(D)$, tai Furjė eilutės sumavimas yra nekorektiškas uždavinys, nes neišpildytas sprendinio stabilumo reikalavimas.

Jeigu atstumą tarp funkcijų $f(t)$ ir $f_T(t)$ apibrėžiame naudodami $V = L_2(0, \pi)$ metriką:

$$\rho_{L_2}(f_1, f_2) = \Big(\int_0^{\pi}\big(f_1(t) - f_2(t)\big)^2 dt\Big)^{1/2},$$

tai tokioje metrinių erdvių poroje (U, V) Furjė eilutės sumavimo uždavinys jau korektiškas, nes, naudodami Parsevalio teoremą, gauname lygybę

$$\delta(\varepsilon_1) := \Big(\int_0^{\pi}\big(f(t) - f_T(t)\big)^2 dt\Big)^{1/2} = \Big(\sum_{n=0}^{\infty}\frac{\pi}{2}(a_n - c_n)^2\Big)^{1/2} = \varepsilon_1\sqrt{\frac{\pi}{2}}.$$

3. Integralinių lygčių sprendimas. Šiuolaikinė neardomosios diagnostikos, medicinos (pvz. tomografijos) aparatūra padeda gauti svarbią informaciją, kurios negalime tiesiogiai išmatuoti ar tirti, o naudojame tik nesunkiai matuojamų pagalbinių charakteristikų ar požymių rezultatus. Tokių metodų matematinis modelis aprašomas integraline lygtimi

$$Av := \int_a^b K(x, s)v(s)\, ds = u(x), \quad c \le x \le d,$$

čia $v(s) \in V$ yra ieškomoji funkcija, $u(x) \in U$ – eksperimentiniai matavimo duomenys, o funkcija $K(x,s)$ charakterizuoja matavimo aparatūrą (ji yra žinoma). Imdami skirtingus režimus $x \in [c,d]$ atliekame seriją matavimų $u(x), x \in [c,d]$ (pavyzdžiui rentgeno spinduliais peršviečiame tą patį objektą įvairiomis kryptimis), tada iš jų randame funkciją $v(s)$.

1 pastaba. Funkcijos $u(t)$ išvestinės skaičiavimo uždavinį irgi galime užrašyti kaip integralinę lygtį

$$\int_0^t v(s)\, ds = u(t) - u(0).$$

Kadangi vietoj tikslių pradinių duomenų $u_T(x)$ turime tik apytikslius eksperimentinius duomenis $u(x) \in U$, tai integralinės lygties

$$Av = u, \quad u(x) \in U$$

sprendinys $v \in V$ gali ir neegzistuoti, jei $u(x) \notin AV$. Taigi šiuo atveju gali būti neišpildyta ir pirmoji korektiškų uždavinių apibrėžimo sąlyga. Pavyzdžiui, taip bus tada, kai $K(x,s)$ yra tolydžiai diferencijuojama pagal x funkcija, o eksperimente gautoji funkcija $u(x)$ nėra diferencijuojama. Čia jau negalime taikyti klasikinio sprendinio apibrėžimo

$$v = A^{-1}u,$$

nes toks sprendinys egzistuoja tik tada, kai $u(x)$ išvestinė yra tolydi funkcija. Reikia nutarti, kaip apibrėšime sprendinį, kad jis egzistuotų ir neglodiems pradiniams duomenims. Kitoje paskaitoje pateiksime tokių apibrėžimų pavyzdžius.

Tačiau net ir tada, kai apytiksliai pradiniai duomenys $u(x) \in AV$, sprendinys gali netenkinti *stabilumo* sąlygos. Tarkime, kad atstumas tarp dviejų funkcijų metrinėje erdvėje U apibrėžtas formule

$$\rho_U(u_1,u_2) = \Big(\int_c^d \big(u_1(x) - u_2(x)\big)^2 dx \Big)^{1/2},$$

t.y. turime $U = L_2[c,d]$ erdvę, o sprendinio pokyčius vertinsime $V = C[a,b]$ metrikoje:

$$\rho_C(v_1,v_2) = \max_{a \le x \le b} \big| v_1(x) - v_2(x) \big|.$$

Tarkime, kad apytiksliai pradiniai duomenys yra tokie:

$$u(x) = u_T(x) + M \int_a^b K(x,s) \sin(\omega s)\, ds, \quad M, \omega > 0.$$

Įvertinsime pradinių duomenų paklaidą

$$\rho_U(u,u_T) = M\Big[\int_c^d \Big(\int_a^b K(x,s)\sin(\omega s)\,ds\Big)^2 dx\Big]^{1/2}.$$

Nagrinėkime atvejį, kai funkcija $K(x,s)$ yra diferencijuojama pagal s ir išvestinė $\left|\frac{\partial K}{\partial s}\right| \le C$ yra aprėžta iš viršaus. Integruodami dalimis, gauname lygybę

$$\int_a^b K(x,s)\sin(\omega s)\,ds = \frac{1}{\omega}\Big(-K(x,s)\cos(\omega s)\Big|_a^b + \int_a^b \frac{\partial K}{\partial s}(x,s)\cos(\omega s)\,ds\Big).$$

Taigi pradinių duomenų pokyčius įvertiname nelygybe

$$\delta := \rho_U(u,u_T) \le \frac{CM}{\omega}.$$

Parinkę pakankamai didelį parametrą ω, šiuos pokyčius galime padaryti kiek norima mažais. Tačiau tokiems pradiniams duomenims atstumas tarp sprendinių erdvės $C[a,b]$ metrikoje

$$\rho_{C[a,b]}(v,v_T) := \max_{a\le s\le b}|v(s)-v_T(s)| = M\max_{a\le s\le b}|\sin(\omega s)| = M$$

yra kiek norima didelis, jei M didelis skaičius. Taigi neišpildyta sprendinio tolydžios priklausomybės nuo pradinių duomenų sąlyga.

Integralinės lygties sprendimo uždavinys nekorektiškas net ir tada, kai sprendinius apibrėžiame $L_2(a,b)$ metrinėje ervėje (šioje erdvėje pradinių duomenų pokyčius galime padaryti kiek norima mažais). Pasinaudojame paprastomis trigonometrinėmis formulėmis:

$$\sin^2 x = \frac{1-\cos(2x)}{2}, \quad \sin(2x)-\sin(2y) = 2\sin(x-y)\cos(x+y),$$

tada gauname lygybę:

$$\rho_{L_2}(v,v_T) = \Big(\int_a^b \big(v(s)-v_T(s)\big)^2 ds\Big)^{1/2} = M\Big(\int_a^b \sin^2(\omega s)\,ds\Big)^{1/2}$$
$$= M\Big(\frac{b-a}{2} - \frac{1}{2\omega}\sin\big(\omega(b-a)\big)\cos\big(\omega(b+a)\big)\Big)^{1/2} = CM.$$

Taigi $\rho_{L_2}(v,v_T)$ gali būti kiek norima didelis, o $\rho_U(u,u_T)$ kiek norima mažas, atitinkamai parinkus ω ir M. Vėl neišpildyta sprendinio tolydžios priklausomybės nuo pradinių duomenų sąlyga.

4. Adamaro pavyzdys. Šį uždavinį nagrinėjo Ž. Adamaras, formuluodamas korektiškų matematinių uždavinių reikalavimus. Srityje

$$D = \{(x, y): \ -\infty < x < \infty, \ \ 0 \le y < Y\}$$

spręsime dvimatę Laplaso lygtį

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0, \tag{1.1}$$

kai užduotos pradinės sąlygos

$$v(x, 0) = 0, \quad \left.\frac{\partial v}{\partial y}\right|_{y=0} = \varphi(x), \quad -\infty < x < \infty. \tag{1.2}$$

Pradinius duomenis ir sprendinį nagrinėsime tolydžiųjų funkcijų metrinėse erdvėse C ir $C(D)$, kurių metrikos yra apibrėžtos taip:

$$\rho_C(\varphi_1, \varphi_2) = \sup_{-\infty<x<\infty} |\varphi_1(x) - \varphi_2(x)|,$$

$$\rho_{C(D)}(v_1, v_2) = \sup_{(x,y)\in D} |v_1(x, y) - v_2(x, y)|.$$

Jeigu pradinė funkcija $\varphi_T(x) \equiv 0$, tai (1.1) – (1.2) uždavinio sprendinys

$$v_T(x, y) \equiv 0.$$

Pakeitę pradinę sąlygą į $\varphi(x) = \dfrac{1}{a} \sin ax$, gauname tokį uždavinio sprendinį

$$v(x, y) = \frac{1}{a^2} \sin(ax) \sinh(ay), \quad a > 0.$$

Parinkdami pakankamai didelį skaičių a, atstumą tarp pradinių duomenų

$$\rho_C(\varphi, \varphi_T) = \sup_{-\infty<x<\infty} \left|\frac{1}{a} \sin ax\right| = \frac{1}{a}$$

galime padaryti kiek norima mažu. Tačiau atstumas tarp sprendinių

$$\rho_{C(D)}(v, v_T) = \sup_{(x,y)\in D} \left|\frac{1}{a^2} \sin(ax) \sinh(ay)\right| = \frac{1}{a^2} \sinh(aY)$$

gali būti kiek norima dideliu, kai parenkame didelį skaičių a. Taigi tokio uždavinio sprendinys nėra stabilus pradinių duomenų atžvilgiu, todėl jis nekorektiškas.

2 pastaba. Laplaso lygčiai korektišką uždavinį gauname, kai vietoj antrosios pradinės sąlygos formuluojame kraštinę sąlygą taške $y = Y$, pavyzdžiui

$$v(x, Y) = \varphi(x).$$

5. Diferencialinės lygties koeficientų radimas. Labai dažnai tenka spręsti atvirkštinį koeficientų radimo uždavinį. Eksperimente matuojame tam tikrus sprendinio funkcionalus, o turime rasti ne tik patį sprendinį, bet ir diferencialinės lygties koeficientus. Nagrinėkime pradinį-kraštinį parabolinio tipo uždavinį, modeliuojantį temperatūros $u(x,t)$ pasiskirstymą vienmačiame strype. Turime diferencialinę lygtį

$$c\rho\frac{\partial u}{\partial t} = a\,\frac{\partial^2 u}{\partial x^2} + f(x,t), \quad (t,x) \in (0,T] \times (0,1),$$

pradinę sąlygą temperatūrai

$$u(x,0) = u_0(x), \quad x \in [0,1],$$

ir dvi kraštines sąlygas (matuojame temperatūrą strypo galuose):

$$u(0,t) = \mu_0(t), \;\; u(1,t) = \mu_1(t), \quad t \in (0,T].$$

Šilumos talpumo ir medžiagos tankio koeficientai c, ρ, bei išorinio šilumos šaltinio tankio funkcija $f(x,t)$ yra žinomos. Reikia rasti funkciją $u(x,t)$ ir difuzijos koeficientą a.

3 pastaba. Iš matematinės fizikos kurso žinome, kad parabolinis diferencialinis uždavinys yra korektiškas (vienintelis sprendinys egzistuoja ir tolydžiai priklauso nuo pradinių duomenų, kai parametras a irgi yra žinomas).

Norėdami rasti (u,a), papildomai matuojame dar vieną sprendinio funkcionalą, pavyzdžiui šilumos srautą

$$a\frac{\partial u}{\partial x}\Big|_{x=0} = \nu(t), \quad t \in (0,T].$$

Galima įrodyti, kad gautasis uždavinys nėra stabilus, taigi jis nekorektiškas.

Egzistuoja daug svarbių taikomųjų uždavinių, kai reikia rasti matematinio modelio koeficientus, o turime tik eksperimentinius matavimus. Tokie uždaviniai dažniausiai nekorektiški, todėl svarbu išmokti efektyviai spręsti šiuos uždavinius.