

Metoda končnih elementov, ki minimizira kvadrat ostanka aproksimacije (LSFEM)

Seminarska naloga pri Naprednih numeričnih metodah

Numerično reševanje parcialnih diferencialnih enačb (PDE) je zaradi pomanjkanja vsestranskega algoritma še zmeraj bolj umetnost kot ustaljena znanost [1]. Pri zapletenih problemih hitro prispemo do vznožja gore matematične teorije, ki je ni moč zaobiti. Zaradi množice različnih pristopov reševanja ter raztresene in neprijazno napisane literature, lahko le ugibamo, kako visoko se bomo na poti do prelaza morali povzpeli. Zapletenim problemom prostorske dinamike v:

- dinamiki tekočin,
- termodinamiki,
- elektrodinamiki,
- kvantni teoriji,
- splošni teoriji relativnosti,

kjer naletimo na PDE, se tako tudi v višjem izobraževanju najraje izognemo. Metoda končnih elementov (FEM), ki minimizira kvadrat ostanka aproksimacije (LSFEM: Least Squares FEM), obeta razvoj vsestranskega algoritma za reševanje PDE in s tem približanje omenjenih problemov širšemu krogu raziskovalcev.

1 Podlaga za temelje LSFEM

Kadar obravnavamo prostorsko dinamiko (npr. tok tekočine), lahko fizični prostor modeliramo kot 1, 2 ali 3-mnogoterost. Temelje LSFEM bomo polagali na splošnem primeru d -mnogoterosti, za ponazoritev pa na njih sproti gradili konkretni 2D primer.

Naj bo torej prizorišče dogajanja d -mnogoterost Ω (slika 1), opremljena s krajevnim vektorjem:

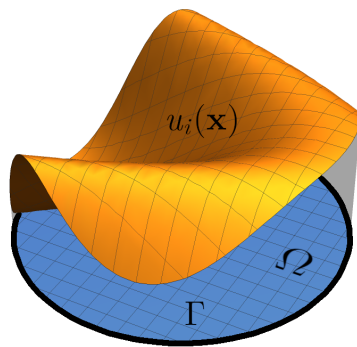
$$\mathbf{x} = \{x_1, \dots, x_d\}.$$

Pri reševanju sistema m PDE iščemo nabor funkcij:

$$\mathbf{u}(\mathbf{x}) = \{u_1(\mathbf{x}), \dots, u_m(\mathbf{x})\},$$

ki v vsaki točki domene Ω zadosti sistemu PDE, na meji Γ pa robnim pogojem. Konkretni primer bomo gradili na 2D primeru s štirimi spremenljivkami, pri katerem bosta krajevni vektor in vektor odvisnih spremenljivk enaka:

$$\mathbf{x} = \{x, y\} \quad \text{in} \quad \mathbf{u} = \{u, v, p, \omega\}.$$



Slika 1: Domena Ω , meja domene Γ in komponenta rešitve $u_i(\mathbf{x})$.

Dinamiko naj opiše **sistem Stokesovih enačb** za nestisljive tekočine v obliki hitrost-tlak-vrtinčnost, ki jim v prid nazornosti primera umetno dodamo koeficiente $\alpha(\mathbf{x})$, $\beta(\mathbf{x})$, $\gamma(\mathbf{x})$ in $\delta(\mathbf{x})$:

$$\alpha \frac{\partial p}{\partial x} + \beta \frac{\partial \omega}{\partial y} = f_x, \quad (1) \quad \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad (3)$$

$$\gamma \frac{\partial p}{\partial y} - \delta \frac{\partial \omega}{\partial x} = f_y, \quad (2) \quad \omega + \frac{\partial u}{\partial y} - \frac{\partial v}{\partial x} = 0. \quad (4)$$

Stokesove enačbe ustrezajo stacionarnim Navier-Stokesovim enačbam brez nelinearnih konvektivnih členov, ki jih moramo pri numeričnem reševanju linearizirati. Ker ta korak za ponazoritev LSFEM ni ključen, se mu na tak način izognemo. Stokesove enačbe opisujejo plazeče se tokove, pri katerih je konvekcija gibalne količine (zaradi gibanja) majhna v primerjavi z njeno difuzijo (zaradi viskoznosti).

V enačbah ni časovnih odvisnosti (razen preko časovno odvisnih robnih pogojev), zato so takšni tokovi časovno obrnljivi: časovno obrnjena rešitev enačb je prav tako rešitev (slika 2).



Slika 2: Zabaven eksperiment, pri katerem se v ozkem prostoru med dvema koncentričnima valjema nahaja viskozna tekočina, ki jo na dveh mestih označimo z liso barvila. Valja pet minut vrtimo v nasprotnih smereh (Stokesov tok, ki tako nastane, imenujemo Taylor-Couettov tok), da se lisi pomešata, nato smeri vrtenja obrnemo in po petih minutah se lisi ponovno sestavita. Pridobljeno iz [2].

Sistem PDE, ki ga obravnavamo, zapišemo bolj jedrnato v matrični obliki. To je enostavno, če je sistem linearen. Uvedemo diferencialni operator \mathbf{A} :

$$\mathbf{A}(\mathbf{x}) = \mathbf{A}_0(\mathbf{x}) + \mathbf{A}_1(\mathbf{x})\frac{\partial}{\partial x_1} + \mathbf{A}_2(\mathbf{x})\frac{\partial}{\partial x_2}, \quad (5)$$

s katerim lahko sistem enačb zapišemo kot:

$$\left(\mathbf{A}_0(\mathbf{x}) + \mathbf{A}_1(\mathbf{x})\frac{\partial}{\partial x_1} + \mathbf{A}_2(\mathbf{x})\frac{\partial}{\partial x_2} \right) \cdot \mathbf{u}(\mathbf{x}) = \mathbf{f}(\mathbf{x}), \quad (6)$$

oziroma na kratko:

$$\boxed{\mathbf{A}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) = \mathbf{f}(\mathbf{x})} \quad \text{sistem PDE.} \quad (7)$$

V matriko \mathbf{A}_0 spravimo vse koeficiente pred členi z odvisnimi spremenljivkami, v matriko \mathbf{A}_1 vse koeficiente pred členi z odvodi odvisnih spremenljivk po koordinati x_1 in v \mathbf{A}_2 vse koeficiente pred členi z odvodi odvisnih spremenljivk po koordinati x_2 . Ostale člene zložimo v vektor \mathbf{f} . Enačbe (1) - (4) lahko v duhu enačbe (6) zapišemo kot:

$$\left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & -\delta \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial x} + \begin{pmatrix} 0 & 0 & 0 & \beta \\ 0 & 0 & \gamma & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \frac{\partial}{\partial y} \right] \cdot \begin{pmatrix} u \\ v \\ p \\ \omega \end{pmatrix} = \begin{pmatrix} f_x(\mathbf{x}) \\ f_y(\mathbf{x}) \\ 0 \\ 0 \end{pmatrix}.$$

2 Temelji LSFEM

Vse različice FEM vsaj okvirno temeljijo na variacijskem pristopu, kjer ne operiramo neposredno na PDE, ampak jih najprej pretvorimo v enakovreden variacijski problem: omislimo si **poskusno funkcijo** $\mathbf{w}(\mathbf{x})$, ki jo napnemo nad domeno Ω , in izberemo funkcional $I[\mathbf{w}(\mathbf{x})]$, ki za vsako $\mathbf{w}(\mathbf{x})$ vrne neko realno število. Za uspešnost variacijskega pristopa moramo izbrati funkcional, ki vrne najmanjšo vrednost, ko je $\mathbf{w}(\mathbf{x})$ enaka rešitvi. Kadar obstaja s sistemom PDE povezan energijski potencial, je le-ta fizikalno najintuitivnejša izbira za konstrukcijo funkcionala. Zato ni presenetljivo, da je bila **Rayleigh-Ritzeva različica FEM (RRFEM)**, ki jo na tak način dobimo, razvita prva [3]. Konstrukcija funkcionala in njegova minimizacija sta tipična koraka variacijskega pristopa in nista specifična za RRFEM: vzamemo neko funkcijo poskusne funkcije $F(\mathbf{w})$ in jo integriramo po domeni Ω :

$$\boxed{I[\mathbf{w}(\mathbf{x})] = \int_{\Omega} F(\mathbf{w}(\mathbf{x})) \, d\Omega} \quad \text{funkcional poskusne funkcije} . \quad (8)$$

Ko smo prepričani, da ima funkcional (8) minimum pri rešitvi $\mathbf{u}(\mathbf{x})$, sledimo znanemu Euler-Lagrangevemu postopku. Ta nas pripelje do variacijske izjave, ki velja le, kadar je poskusna funkcija $\mathbf{w}(\mathbf{x})$ enaka rešitvi $\mathbf{u}(\mathbf{x})$. Poskusno funkcijo razvijemo okoli rešitve:

$$\tilde{\mathbf{w}}(\mathbf{x}, \varepsilon) = \mathbf{u}(\mathbf{x}) + \varepsilon \mathbf{v}(\mathbf{x}) , \quad (9)$$

kjer je $\mathbf{v}(\mathbf{x})$ poljubna odmična funkcija, ε pa realno število. Kadar gre ε proti nič, gre $\tilde{\mathbf{w}}(\mathbf{x}, \varepsilon)$ proti rešitvi problema $\mathbf{u}(\mathbf{x})$, hkrati pa vemo, da ima funkcional I pri $\mathbf{u}(\mathbf{x})$ minimum. Minimum funkcionala poiščemo tako, da razvoj (9) vstavimo v funkcional (8) namesto $\mathbf{w}(\mathbf{x})$ in izraz odvajamo po ε :

$$\frac{dI}{d\varepsilon} = \int_{\Omega} \frac{d}{d\varepsilon} F(\tilde{\mathbf{w}}) \, d\Omega = \int_{\Omega} \left(\frac{dF}{d\tilde{\mathbf{w}}} \right)^{\top} \cdot \frac{d\tilde{\mathbf{w}}}{d\varepsilon} \, d\Omega = \int_{\Omega} \left(\frac{dF}{d\tilde{\mathbf{w}}} \right)^{\top} \cdot \mathbf{v} \, d\Omega ,$$

nato pa ε v (2) pošljemo proti nič in celoten izraz enačimo z nič:

$$\lim_{\varepsilon \rightarrow 0} \frac{dI}{d\varepsilon} = \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \left(\frac{dF(\tilde{\mathbf{w}}(\mathbf{x}, \varepsilon))}{d\tilde{\mathbf{w}}} \right)^{\top} \cdot \mathbf{v}(\mathbf{x}) \, d\Omega = 0 .$$

Limita deluje le na prvi člen v integrandu, zato se variacijska izjava glasi:

$$\boxed{\int_{\Omega} \lim_{\varepsilon \rightarrow 0} \left(\frac{dF(\tilde{\mathbf{w}})}{d\tilde{\mathbf{w}}} \right)^{\top} \cdot \mathbf{v} \, d\Omega = 0 , \quad \forall \mathbf{v}(\mathbf{x})} \quad \text{variacijska izjava} . \quad (10)$$

V jeziku funkcionalne analize, kjer na funkcije gledamo kot na vektorje, izjava pove naslednje: projekcija izraza v oklepaju na katerokoli odmično funkcijo $\mathbf{v}(\mathbf{x})$ mora biti enaka nič, ali krajše: izraz mora biti ortogonalen na katerokoli $\mathbf{v}(\mathbf{x})$.

$F(\mathbf{w}(\mathbf{x}))$ v funkcionalu poskusne funkcije je pri RRFEM energijski potencial, funkcional pa torej skupna potencialna energija sistema, ki jo rešitev $\mathbf{u}(\mathbf{x})$ minimizira. Zaradi tega ima RRFEM lastnost najboljšega približka, diskretizacija pa vodi do simetričnega in pozitivno-definitnega sistema algebraskih enačb, ki je zelo prikladen za reševanje s hitrimi iteracijskimi metodami. Različica metode se je izkazala v gradbenem inženirstvu, kjer je s problemom vedno povezan energijski potencial. Večina računalniških programov s tega področja zato še danes temelji na RRFEM.

Žal energijski potencial povezan s sistemom PDE ne obstaja vedno, kar je značilno za sisteme PDE v dinamiki tekočin. To je motiviralo razvoj Galerkinove različice FEM (GFEM), ki je zastavljena kot posplošitev RRFEM, vendar na precej neroden način. Ideja GFEM je, da lahko za vsak sistem PDE (7) in za poljubno poskusno funkcijo $\mathbf{w}(\mathbf{x})$ definiramo vektor ostanka $\mathbf{R}(\mathbf{w}(\mathbf{x}))$. Vse člene v jedrnatem zapisu sistema PDE (7) damo na eno stran in namesto $\mathbf{u}(\mathbf{x})$ pišemo $\mathbf{w}(\mathbf{x})$:

$$\boxed{\mathbf{R}(\mathbf{w}(\mathbf{x})) = \mathbf{A}(\mathbf{x}) \cdot \mathbf{w}(\mathbf{x}) - \mathbf{f}(\mathbf{x})} \quad \text{vektor ostanka} . \quad (11)$$

Funkcija $\mathbf{w}(\mathbf{x})$, za katero je ostanek $\mathbf{R}(\mathbf{w}(\mathbf{x}))$ enak nič, je rešitev problema $\mathbf{u}(\mathbf{x})$. Idejo za izničenje ostanka vzamemo iz variacijske izjave (10): naj bo ostanek $\mathbf{R}(\mathbf{w}(\mathbf{x}))$ ortogonalen na katerokoli odmično funkcijo $\mathbf{v}(\mathbf{x})$:

$$\int_{\Omega} \mathbf{R}(\mathbf{w}(\mathbf{x}))^{\top} \cdot \mathbf{v}(\mathbf{x}) \, d\Omega = 0 \quad \text{načelo metode uteženih ostankov} . \quad (12)$$

Pristop se imenuje **metoda uteženih ostankov** in nas za sebi-adjungirane ter pozitivno-definitne $\mathbf{A}(\mathbf{x})$ pripelje do istega sistema algebraskih enačb kot RRFEM. Uporabimo pa ga lahko tudi za sisteme PDE, ki ne posedujejo teh lastnosti, zato daje vtis posplošitve RRFEM. Akademiki so pričakovali, da bo GFEM v dinamiki tekočin enako uspešna, kot je bila RRFEM v gradbenem inženirstvu, a se to ni zgodilo [1]. Ko \mathbf{A} ni sebi-adjungiran, namreč načelo (12) ni nujno zvesto osnovnemu problemu PDE.

Metoda v tem primeru ne poseduje lastnosti najboljšega približka in v rešitvi se pogostokrat pojavijo lažne oscilacije (*wiggles*). Teh se je mogoče znebiti le s hudimi izboljšavami mreže, kar očitno okrnjuje praktičnost metode. Metoda končnih diferenc in metoda končnih volumnov sta zato v dinamiki tekočin še vedno v modi.

Podobno kot pri GFEM se tudi pri LSFEM opremo na vektor ostanka (11), a reševanja variacijskega problema se lotimo na legitimen način. Ne zanašamo se na ad hoc načela, kot je zahteva (12), ampak začnemo od začetka - s konstrukcijo funkcionala (8). Sestavimo ga s kvadratom ostanka:

$$F(\mathbf{w}(\mathbf{x})) = \mathbf{R}(\mathbf{w}(\mathbf{x}))^\top \cdot \mathbf{R}(\mathbf{w}(\mathbf{x})) \quad \text{kvadrat vektorja ostanka ,} \quad (13)$$

in tako se funkcional, ki ga minimiziramo, glasi:

$$\boxed{I[\mathbf{w}(\mathbf{x})] = \int_{\Omega} \mathbf{R}(\mathbf{w}(\mathbf{x}))^\top \cdot \mathbf{R}(\mathbf{w}(\mathbf{x})) \, d\Omega} \quad \text{funkcional LSFEM ,} \quad (14)$$

od koder dobi metoda svoje ime. Želimo dobiti variacijsko izjavo za ta specifični funkcional, zato v splošno izjavo (10) vstavimo kvadrat vektorja ostanka (13) in postopek izpeljemo do konca. Najprej torej izračunamo odvod integranda:

$$\left(\frac{dF(\tilde{\mathbf{w}})}{d\tilde{\mathbf{w}}} \right)^\top = \left(\frac{d(\mathbf{R}(\tilde{\mathbf{w}})^\top \cdot \mathbf{R}(\tilde{\mathbf{w}}))}{d\tilde{\mathbf{w}}} \right)^\top = 2\mathbf{R}(\tilde{\mathbf{w}})^\top \cdot \frac{d\mathbf{R}(\tilde{\mathbf{w}})}{d\tilde{\mathbf{w}}} = 2(\mathbf{A} \cdot \tilde{\mathbf{w}} - \mathbf{f})^\top \cdot \mathbf{A}$$

ter nato limto, ko gre ε proti nič:

$$\lim_{\varepsilon \rightarrow 0} \left(\frac{dF(\tilde{\mathbf{w}})}{d\tilde{\mathbf{w}}} \right)^\top = \lim_{\varepsilon \rightarrow 0} 2(\mathbf{A} \cdot (\mathbf{u} + \varepsilon \mathbf{v}) - \mathbf{f})^\top \cdot \mathbf{A} = 2(\mathbf{A} \cdot \mathbf{u} - \mathbf{f})^\top \cdot \mathbf{A} .$$

Variacijska izjava za minimizacijo funkcionala LSFEM se zato glasi:

$$\int_{\Omega} 2(\mathbf{A} \cdot \mathbf{u} - \mathbf{f})^\top \cdot (\mathbf{A} \cdot \mathbf{v}) \, d\Omega = 0 , \quad \forall \mathbf{v}(\mathbf{x}) .$$

Celoten izraz transponiramo (zamenjamo vrstni red členov v zunanjem skalarnem produktu, s čemer rezultata ne spremenimo) ter delimo z dve:

$$\boxed{\int_{\Omega} (\mathbf{A} \cdot \mathbf{v})^\top \cdot (\mathbf{A} \cdot \mathbf{u} - \mathbf{f}) \, d\Omega = 0 , \quad \forall \mathbf{v}} \quad \text{variacijska izjava LSFEM .} \quad (15)$$

Izjava pravzaprav ustreza Galerkinovi formulaciji (12), kjer namesto odmičnih funkcij samih (\mathbf{v}) uporabimo njihove odvode ($\mathbf{A} \cdot \mathbf{v}$).

3 Diskretizacija problema

Za reševanje problema z računalnikom moramo abstraktno zastavitev (15) z neskončno prostostnimi stopnjami narediti oprijemljivejšo. Domena Ω namreč kljub omejenosti vsebuje neskončno število točk, v katerih funkcijam dopuščamo neodvisne vrednosti. Število prostostnih stopenj omejimo tako, da poljubno funkcijo $v(\mathbf{x})$ razvijemo v vrsto N baznih funkcij, ki jih izberemo sami:

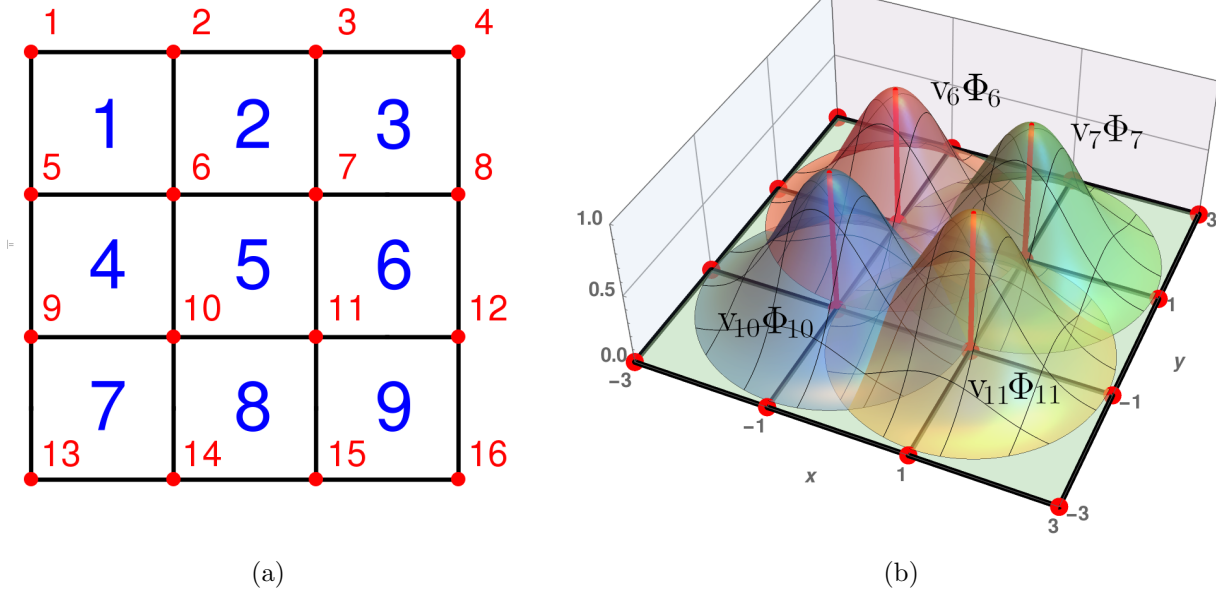
$$\Phi_i(\chi), \quad i = 1, \dots, N \quad \text{vozliščne funkcije ,}$$

Vzemimo skrajno preprosto kvadratno domeno $[-3, 3] \times [-3, 3]$ s krajevnim vektorjem $\chi = \{\xi, \eta\}$. Nanjo postavimo pravokotno mrežo z n ($= 9$) **elementi** in N ($= 16$) **vozlišči** (slika 3a). Nad vozlišči napnimo prav toliko vozliščnih funkcij:

z nastavljlivo višino (**vozliščno vrednostjo**) v_i (slika 3b).

Poljubno skalarno funkcijo $v(\mathbf{x})$ na domeni nato aproksimiramo s sestavljanjo N vozliščnih funkcij:

$$v(\mathbf{x}) = \sum_{i=1}^N v_i \Phi_i(\mathbf{x}) . \quad (16)$$



Slika 3: (a) Pravokotna domena z devetimi elementi (modre številke) in šestnajstimi vozlišči (rdeče številke) ter (b) nad vozlišči napete vozliščne funkcije. V prid nazornosti rišemo le osrednje vozliščne funkcije.

Nobena vozliščna funkcija Φ_i se ne sme raztezati nad elementi, ki niso v stiku z njenim vozliščem. S tem dosežemo, da je $v(\mathbf{x})$ nad nekim elementom sestavljena le iz funkcij v neposredni bližini tega elementa. Tako je $\mathbf{v}(\mathbf{x})$ na sliki 4a nad osrednjim elementom popolnoma določena z vrednostmi v_6, v_7, v_{10} in v_{11} .

Ozrmo se na variacijsko izjavo (15) ter si predstavljajmo funkcije $\mathbf{A}(\mathbf{x})$, $\mathbf{u}(\mathbf{x})$, $\mathbf{v}(\mathbf{x})$ in $\mathbf{f}(\mathbf{x})$ zapisane v smislu razvoja po vozliščnih funkcijah (16). Zaslutimo, da bomo računali prekrivne integrale vozliščnih funkcij:

$$\int \Phi_i(\mathbf{x}) \Phi_j(\mathbf{x}) d\Omega . \quad (17)$$

To je enostavno dokler so vsi elementi iste oblike in velikosti, kot na sliki 3. Takrat je dovolj, da izračunamo prekrivne integrale za vozlišča enega elementa. Kaj pa, če želimo uporabljati elemente poljubne oblike? Kako naj čim učinkoviteje, če so elementi poljubne oblike

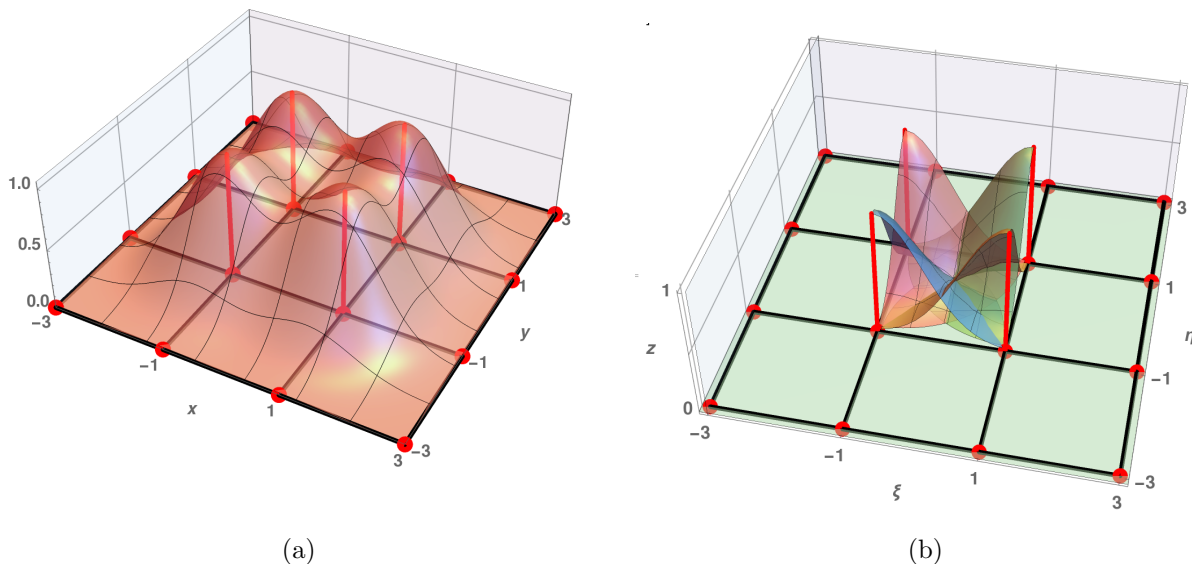
Segmente vozliščnih funkcij $\Phi_6, \Phi_7, \Phi_{10}$ in Φ_{11} , ki se nahajajo neposredno nad elementom 5, proglasimo za **elementarne funkcije** $\phi_{5j}(\mathbf{x})$ tega elementa (slika 4b). Tako lahko funkcijo \mathbf{v} na

še vedno Z natančno analitično izpeljavo se prebijemo do izjave (15), od tod dalje pa moramo iskanje funkcije $\mathbf{u}(\mathbf{x})$ z neskončno prostostnimi stopnjami poenostaviti v iskanje funkcije s končnim številom prostostnih stopenj N .

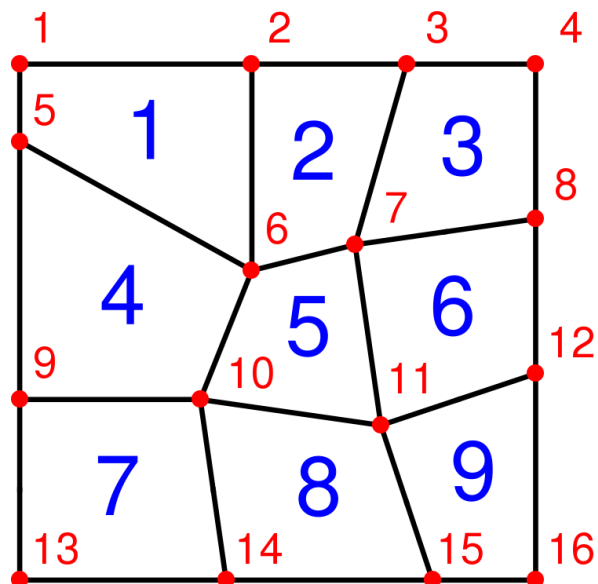
Skozi oči FI je $|\Phi_i\rangle$ eden izmed baznih vektorjev v razvoju vektorja $|v\rangle$, v_i pa pripadajoča komponenta. V jeziku funkcionalne analize (FI) pravimo, da smo omejili funkcijski prostor.

nadaljujemo z diskretizacijo problema, to je, pretvorbo na sistem N algebrskih enačb. Ta korak je enak pri vseh različicah FEM. Funkcije na domeni Ω imajo neskončno število prostostnih stopenj.

$$u_i(\mathbf{x}) = \sum_{a=1}^N \Phi^{a0} u_i^{a0} \quad (18)$$



Slika 4: (a) vsota vozliščnih funkcij s slike 3b in (b) elementarne funkcije, ki pripadajo elementu 5.



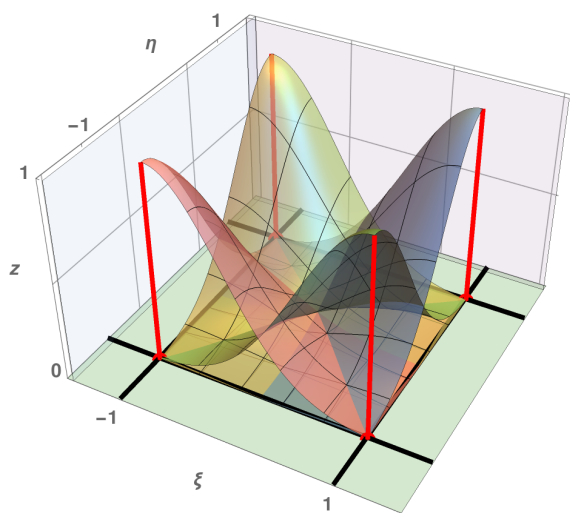
Mreža je v tem šolskem primeru strukturirana, kar pomeni, da je razporeditev elementov Kartezična. Mreža je lahko pri FEM tudi nestrukturirana, kar je ena izmed prednosti metode.

Potem omejimo Diskretizacija problema

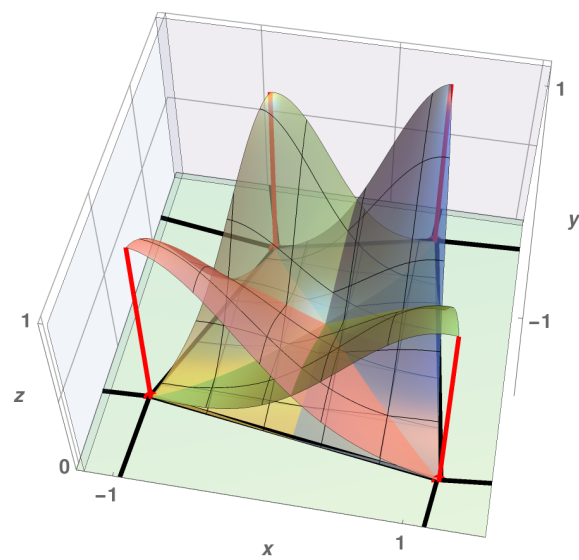
Galerkin, Najmanjših kvadratov [1] Basic lemma of variational principles: Temeljni lema variacijskih načel.

Literatura

- [1] B.-n. Jiang, *The Least-Squares Finite Element Method*. Springer-Verlag, 1998, Heidelberg.
- [2] Wikipedia. (2019). Stokes Flow, spletni naslov: https://en.wikipedia.org/wiki/Stokes_flow.
- [3] W. Ritz, „Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik“, *Journal für die Reine und Angewandte Mathematik*, let. 135, str. 1–61, 1909.



(a)



(b)

Slika 5