Univerzální principy mechaniky

Přednáška 1

21. února 2020

Přehled látky

Vektorový prostor, jeho prvky (vektory), základní operace a jejich vlastnosti, nulový vektor. Lineární kombinace vektorů, lineární nezávislost, báze, dimenze. Složky vektoru vzhledem k bázi, reprezentace vektoru pomocí báze, souvislost mezi vektorem a sloupcovou maticí jeho složek (vzhledem ke zvolené bázi).

Definice lineární formy jako zvláštního případu lineárního zobrazení, operace s lineárními formami (sčítání lineárních forem a násobení lineární formy skalárem), prostor všech lineárních forem (duální prostor). Složky lineární formy vzhledem k bázi, výpočet hodnoty lineární formy pro daný vektor s využitím složek vzhledem ke zvolené bázi (tj. hodnot, kterých nabývá pro bázové vektory), reprezentace této operace pomocí maticového násobení. Názorná představa o lineární formě, např. síla jako lineární forma, která každému vektoru představujícímu posun přiřazuje skalár představující práci vykonanou touto silou na daném posunu.

Definice bilineární formy. Složky bilineární formy vzhledem k dané bázi, tj. hodnoty, kterých nabývá pro dvojice bázových vektorů. Označení složek indexy, výpočet hodnoty bilineární formy pro dané vektory s využitím složek vzhledem ke zvolené bázi, reprezentace této operace pomocí maticového násobení. Zjednodušení zápisu pomocí sčítací (Einsteinovy) konvence.

Skalární součin jako symetrická pozitivně definitní bilineární forma, norma vektoru, úhel mezi dvěma vektory, ortogonalita vektorů. Ortonormální báze, Kroneckerovo delta. Výpočet složek vektoru vzhledem k dané ortonormální bázi pomocí průmětu (skalárního součinu s bázovými vektory).

Lineární forma odpovídající danému vektoru (v prostoru vybaveném skalárním součinem). Tvrzení, že každou lineární formu lze zapsat jako skalární součin s jistým vektorem. Ztotožnění lineárních forem s vektory, a tedy ztotožnění duálního prostoru (prostoru všech lineárních forem) s výchozím vektorovým prostorem. Souvislost mezi vektory, lineárními formami a sloupcovými maticemi jejich složek.

Doplňující poznámky

Ortonormální báze (již probráno)

Na přednášce byla definována báze vektorového prostoru V jako skupina lineárně nezávislých vektorů $\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_n$, pomocí kterých lze vyjádřit libovolný vektor $\mathbf{u} \in V$ jako jejich lineární kombinaci

$$\mathbf{u} = u_i \mathbf{b}_i \tag{1}$$

Připomeňte si, že v tomto vztahu je použita sčítací konvence a výraz na pravé straně představuje součet $u_1\mathbf{b}_1 + u_2\mathbf{b}_2 + \ldots + u_n\mathbf{b}_n$. Čísla $u_1, u_2, \ldots u_n$ jsou složky vektoru $\mathbf{u}$ vzhledem k dané bázi.

Otázka je, jak složky vektoru určit, jestliže nejsou zadány a k dispozici máme pouze vektor $\mathbf{u}$ a bázové vektory. Při řešení této úlohy lze využít skalární součin a přenásobit obě strany vztahu (1) postupně jednotlivými bázovými vektory. Po skalárním vynásobení vektorem $\mathbf{b}_j$ dostaneme[1]

$$\mathbf{b}_j \cdot \mathbf{u} = \mathbf{b}_j \cdot \mathbf{b}_i u_i \tag{2}$$

V tomto vztahu je i sčítací index, zatímco j může postupně nabývat hodnot 1, 2, $\ldots n$, přičemž každé z těchto hodnot odpovídá jedna rovnice. Po přehození pravé a levé strany a rozepsání jednotlivých rovnic bychom dostali

$$(\mathbf{b}_1 \cdot \mathbf{b}_1)u_1 + (\mathbf{b}_1 \cdot \mathbf{b}_2)u_2 + \ldots + (\mathbf{b}_1 \cdot \mathbf{b}_n)u_n = \mathbf{b}_1 \cdot \mathbf{u} \tag{3}$$

$$(\mathbf{b}_2 \cdot \mathbf{b}_1)u_1 + (\mathbf{b}_2 \cdot \mathbf{b}_2)u_2 + \ldots + (\mathbf{b}_2 \cdot \mathbf{b}_n)u_n = \mathbf{b}_2 \cdot \mathbf{u} \tag{4}$$

$$\ldots$$

$$(\mathbf{b}_n \cdot \mathbf{b}_1)u_1 + (\mathbf{b}_n \cdot \mathbf{b}_2)u_2 + \ldots + (\mathbf{b}_n \cdot \mathbf{b}_n)u_n = \mathbf{b}_n \cdot \mathbf{u} \tag{5}$$

Jedná se o soustavu n lineárních rovnic pro n neznámých složek $u_1, u_2, \ldots u_n$, kterou lze řešit např. Gaussovou eliminační metodou.

[1] Na přednášce jsme rovnici (1) přenásobili bázovými vektory zprava a v tomto textu zleva, ale vzhledem ke komutativitě skalárního součinu jsou oba postupy zcela ekvivalentní. Všimněte si, že matice koeficientů v rovnicích (3)–(5) je symetrická.

Situace se výrazně zjednoduší, pokud zvolíme bázové vektory tak, aby byly navzájem kolmé (neboli ortogonální). Protože skalární součin dvou ortogonálních vektorů je nulový, rozpadne se soustava (3)–(5) na n nezávislých rovnic a její řešení bude snadné. Navíc můžeme bázové vektory volit jako jednotkové (tj. jako vektory s jednotkovou normou). Pak budou jednotkové i koeficienty $\mathbf{b}_1 \cdot \mathbf{b}_1$, $\mathbf{b}_2 \cdot \mathbf{b}_2$, atd., a neznámé budou přímo rovny pravé straně. Báze tohoto speciálního typu se nazývá *ortonormální* a její bázové vektory budeme značit $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$.

Již zmíněné požadavky na ortonormální bázi můžeme přepsat ve tvaru

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} \tag{6}$$

kde δ_{ij} je tzv. *Kroneckerovo delta*, definované vztahy

$$\delta_{ij} = \begin{cases} 0 & \text{pro } i \neq j \\ 1 & \text{pro } i = j \end{cases} \tag{7}$$

Pokud pracujeme s ortonormální bází, můžeme místo (2) napsat

$$\mathbf{e}_j \cdot \mathbf{u} = \mathbf{e}_j \cdot \mathbf{e}_i u_i = \delta_{ji} u_i = u_j \tag{8}$$

Tím získáme pohodlnou metodu pro určení složek vektoru vůči ortonormální bázi. Složka u_j se spočítá jako skalární součin vektoru $\mathbf{u}$ s j-tým bázovým vektorem $\mathbf{e}_j$. Není třeba řešit žádnou soustavu rovnic a v tom spočívá velká výhoda ortonormální báze.

Dobře si rozmyslete poslední rovnost v rovnici (8). Index i je zde sčítací, zatímco index j je pro jednu rovnici tohoto typu pevně zvolený (i když může nabývat různých hodnot, ale každé z nich odpovídá samostatná rovnice). Pro určitost položme např. $j = 2$. Výraz $\delta_{2i} u_i$ představuje podle sčítací konvence součet $\delta_{21} u_1 + \delta_{22} u_2 + \ldots + \delta_{2n} u_n$. Koeficienty δ_{2i} jsou ale nulové pro všechna $i \neq 2$, takže nenulový je pouze koeficient δ_{22}, který je roven jedné a násobí neznámou u_2. Obecně jsou v j-té rovnici nulové všechny koeficienty s výjimkou toho, který násobí neznámou u_j a má jednotkovou hodnotu. Proto můžeme místo $\delta_{ji} u_i$ napsat jednoduše u_j. Tato úvaha se bude v úpravách výrazů obsahujících Kroneckerovo delta často opakovat a můžeme z ní vyvodit nedbale formulované pravidlo, že Kroneckerovo δ_{ij} "udělá" z indexu i index j (nebo z indexu j index i) a tím "zanikne". Například výraz $F_{kj} \delta_{ij}$ můžeme přepsat jako F_{ki}.

Přemýšlivého studenta může napadnout otázka, zda ortonormální báze vůbec existuje, a pokud ano, jak ji lze získat. Pokud by pro některý vektorový prostor žádná ortonormální báze neexistovala, byly by předchozí úvahy o elegantním výpočtu složek vektoru nepoužitelné. Naštěstí je k dispozici konstruktivní postup, který umožňuje z jakékoliv báze sérií jasně popsaných kroků vytvořit bázi ortonormální, čímž je zároveň dokázána její existence. Jde o tzv. Gramovu-Schmidtovu ortogonalizační metodu, jejíž popis lze najít pod příslušným heslem na internetu. Pro naše účely postačí informace o tom, že taková metoda existuje a můžeme tedy vždy předpokládat, že ortonormální báze je k dispozici.

Korespondence mezi vektory a lineárními formami (již probráno)

Na přednášce jsme se zabývali souvislostí mezi vektory a lineárními formami. Pro pevně zvolenou bázi lze každý vektor charakterizovat jeho složkami vzhledem k této bázi, uspořádanými do sloupcové matice. Počet složek n odpovídá dimenzi vektorového prostoru. Podobně i lineární formu lze charakterizovat pomocí n složek a uspořádat je do sloupcové matice.

Pro daný vektor $\mathbf{f}$ můžeme definovat lineární formu f předpisem

$$f(\mathbf{u}) = \mathbf{f} \cdot \mathbf{u} \tag{9}$$

Jinými slovy, forma f přiřazuje každému vektoru $\mathbf{u}$ jeho skalární součin s pevně zvoleným vektorem $\mathbf{f}$. Volbou vektoru $\mathbf{f}$ je odpovídající lineární forma jednoznačně určena. Úvahu lze ale i obrátit. Pokud je dána lineární forma f, můžeme nalézt vektor $\mathbf{f}$ tak, aby platilo (9). K tomu využijeme složek dané lineární formy vzhledem k libovolně (ale pevně) zvolené ortonormální bázi. Složky f_i formy f jsou definovány předpisem

$$f_i = f(\mathbf{e}_i) \tag{10}$$

Jestliže sestrojíme vektor

$$\mathbf{f} = f_i \mathbf{e}_i \tag{11}$$

pak pro libovolný vektor $\mathbf{u}$ bude platit (promyslete si podrobně jednotlivé úpravy)

$$\mathbf{f} \cdot \mathbf{u} = f_i \mathbf{e}_i \cdot \mathbf{u} = f_i u_i = f(\mathbf{e}_i) u_i = f(u_i \mathbf{e}_i) = f(\mathbf{u}) \tag{12}$$

Je tedy vidět, že vektory a lineární formy si vzájemně jednoznačně odpovídají. *Každý vektor můžeme zároveň chápat jako lineární formu a každou lineární formu jako vektor.*

Definice tenzoru

Obecně jsou jako tenzory chápány určité typy lineárních zobrazení. Pro naše účely postačí, když za tenzory 1. řádu označíme lineární formy a za tenzory 2. řádu označíme bilineární formy. Protože už víme, že lineární formy lze ztotožnit s vektory, jsou tenzory 1. řádu vlastně vektory. Uvedenou definici tenzorů 1. a 2. řádu lze snadno zobecnit na vyšší řády. Například tenzory 3. řádu jsou trilineární formy, tedy funkce, které každé uspořádané trojici vektorů přiřazují reálné číslo a zároveň se chovají jako lineární formy vůči každému z těchto vektorů zvlášť, pokud zbývající dva zafixujeme. Podrobně je tento požadavek popsán podmínkami

$$F(\mathbf{u}_1+\mathbf{u}_2,\mathbf{v},\mathbf{w}) = F(\mathbf{u}_1,\mathbf{v},\mathbf{w})+F(\mathbf{u}_2,\mathbf{v},\mathbf{w}) \tag{13}$$

$$F(\alpha\mathbf{u},\mathbf{v},\mathbf{w}) = \alpha F(\mathbf{u},\mathbf{v},\mathbf{w}) \tag{14}$$

$$F(\mathbf{u},\mathbf{v}_1+\mathbf{v}_2,\mathbf{w}) = F(\mathbf{u},\mathbf{v}_1,\mathbf{w})+F(\mathbf{u},\mathbf{v}_2,\mathbf{w}) \tag{15}$$

$$F(\mathbf{u},\alpha\mathbf{v},\mathbf{w}) = \alpha F(\mathbf{u},\mathbf{v},\mathbf{w}) \tag{16}$$

$$F(\mathbf{u},\mathbf{v},\mathbf{w}_1+\mathbf{w}_2) = F(\mathbf{u},\mathbf{v},\mathbf{w}_1)+F(\mathbf{u},\mathbf{v},\mathbf{w}_2) \tag{17}$$

$$F(\mathbf{u},\mathbf{v},\alpha\mathbf{w}) = \alpha F(\mathbf{u},\mathbf{v},\mathbf{w}) \tag{18}$$

Pokud pracuje forma se čtveřicí vektorů, mluvíme o kvadrilineární formě, která představuje tenzor 4. řádu, atd. Lineární, bilineární, trilineární, kvadrilineární a další podobné formy se obecně nazývají multilineární formy.

Můžeme tedy jednoduše říci, že tenzory chápeme jako multilineární formy. Pro úplnost mezi ně zahrnujeme i samotné skaláry, které lze považovat za tenzory nultého řádu.

Transformace složek vektoru při změně báze

Pomocí tenzorového zápisu lze snadno odvodit elegantní vzorce pro transformaci složek tenzoru při změně báze (tedy například pro výpočet složek napětí vůči pootočené soustavě souřadnic, jestliže jsou známy složky napětí vůči původní soustavě souřadnic).

Začneme vektorem, tedy tenzorem 1. řádu, který vyjádříme vzhledem ke dvěma různým ortonormálním bázím:

$$\mathbf{f} = f_i\mathbf{e}_i = f'_i\mathbf{e}'_i \tag{19}$$

Přitom $\mathbf{e}_1,\mathbf{e}_2,\ldots\mathbf{e}_n$ a $\mathbf{e}'_1,\mathbf{e}'_2,\ldots\mathbf{e}'_n$ jsou dvě ortonormální báze stejného vektorového prostoru (např. báze odpovídající dvěma kartézským soustavám souřadnic v trojrozměrném prostoru, pokud $n=3$). Jestliže známe složky f_i a chceme z nich vypočítat složky f'_i, stačí rovnici (19) skalárně přenásobit jedním z vektorů "čárkované" báze:

$$f_i\mathbf{e}_i\cdot\mathbf{e}'_k = f'_i\mathbf{e}'_i\cdot\mathbf{e}'_k \tag{20}$$

Jelikož je čárkovaná báze ortonormální, platí $\mathbf{e}'_i\cdot\mathbf{e}'_k=\delta_{ik}$ a výraz na pravé straně je roven $f'_i\delta_{ik}=f'_k$. Dostáváme tedy

$$f'_k = t_{ki}f_i \tag{21}$$

kde

$$t_{ki} = \mathbf{e}_i\cdot\mathbf{e}'_k = \mathbf{e}'_k\cdot\mathbf{e}_i \tag{22}$$

jsou transformační koeficienty, které představují kosiny úhlů mezi vektory původní a pootočené báze.[2] V maticovém zápisu bychom rovnici (21) mohli přepsat jako součin transformační matice 3×3 obsahující směrové kosiny a sloupcové matice obsahující složky vektoru vzhledem k původní bázi.

Uvědomte si, že koeficienty t_{ki} definované vztahem (22) lze interpretovat nejen jako kosiny úhlů, ale také jako složky vektorů jedné báze vzhledem k druhé. Konkrétně je koeficient t_{ki} i-tou složkou vektoru $\mathbf{e}'_k$ vzhledem k "nečárkované" bázi a zároveň také k-tou složkou vektoru $\mathbf{e}_i$ vzhledem k "čárkované" bázi. Zobecnění na tenzory vyššího řádu je předmětem domácího úkolu.

Transformační vzorec odvozený pomocí multilineárních forem

Pomocí interpretace vektoru jako lineární formy lze provést o něco jednodušší, ale abstraktnější odvození transformačních vzorců.

Víme, že složky lineární formy se získají jejím vyhodnocením pro jednotlivé bázové vektory. Pokud je f lineární forma odpovídající jistému vektoru a $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ je ortonormální báze, můžeme hodnotu formy f pro libovolný vektor $\mathbf{u}$ o složkách u_1, u_2, u_3 vypočítat jako

$$f(\mathbf{u}) = f(u_i\mathbf{e}_i) = u_if(\mathbf{e}_i) = u_if_i \tag{23}$$

kde

$$f_i = f(\mathbf{e}_i) \tag{24}$$

[2]Uvědomte si, že všechny bázové vektory jsou normované, takže pro úhel ϕ_{ki} mezi k-tým vektorem pootočené báze a i-tým vektorem původní báze platí $\cos\phi_{ki}=\mathbf{e}'_k\cdot\mathbf{e}_i$.

jsou právě složky formy f a zároveň i vektoru $\mathbf{f}$. Pracujeme-li s jinou bází $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$, určí se složky formy f vzhledem k této bázi jako

$$f'_k = f(\mathbf{e}'_k) \tag{25}$$

Abychom nalezli vyjádření složek f'_k pomocí složek f_i, stačí vyjádřit bázové vektory $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ pomocí bázových vektorů $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, dosadit do (25) a využít vlastností lineární formy. Jestliže tedy

$$\mathbf{e}'_k = t_{ki}\mathbf{e}_i \tag{26}$$

pak

$$f'_k = f(\mathbf{e}'_k) = f(t_{ki}\mathbf{e}_i) = t_{ki}f(\mathbf{e}_i) = t_{ki}f_i \tag{27}$$

To přesně odpovídá vzorci (21). Výhodou je snadné zobecnění tohoto alternativního postupu na tenzory vyšších řádů.

Vlastnosti transformačních koeficientů

Názorný význam transformačních koeficientů už byl vysvětlen. Podle (26) koeficient t_{ki} představuje i-tou souřadnici (vůči “nečárkované” bázi) k-tého bázového vektoru “čárkované” báze. Explicitní výraz pro transformační koeficienty se získá přenásobením vztahu (26) bázovým vektorem $\mathbf{e}_j$:

$$\mathbf{e}'_k \cdot \mathbf{e}_j = t_{ki}\mathbf{e}_i \cdot \mathbf{e}_j = t_{ki}\delta_{ij} = t_{kj} \tag{28}$$

Výsledek podle očekávání souhlasí se vztahem (22) (až na formální rozdíl, spočívající v tom, že index i je nyní nazván j). Je tedy $t_{kj} = \mathbf{e}'_k \cdot \mathbf{e}_j$ = skalární součin k-tého bázového vektoru čárkované báze a j-tého bázového vektoru nečárkované báze = kosinus úhlu mezi k-tou čárkovanou a j-tou nečárkovanou souřadnicovou osou. Pokud provádíme transformaci “opačným směrem”, tedy z čárkované báze do nečárkované, použijí se transformační koeficienty, které pracovně označíme $\bar{t}_{ik} = \mathbf{e}_i \cdot \mathbf{e}'_k = t_{ki}$. Po transformaci “tam a zase zpátky” musíme vždy dospět k původním složkám, jinými slovy, tyto transformace jsou navzájem inverzní. Přitom složená transformace vede k bázovým vektorům

$$\mathbf{e}''_i = \bar{t}_{ik}\mathbf{e}'_k = t_{ki}\mathbf{e}'_k = t_{ki}t_{kj}\mathbf{e}_j \tag{29}$$

které odpovídají původním bázovým vektorům $\mathbf{e}_i$ pouze, pokud je $t_{ki}t_{kj} = \delta_{ij}$ pro každé i a j. Tato podmínka znamená, že pokud uspořádáme transformační koeficienty do čtvercové matice $\mathbf{T}$, která má v k-tém řádku a j-tém sloupci koeficient t_{kj}, pak platí $\mathbf{T}^T\mathbf{T} = \mathbf{I}$ = jednotková matice, neboli $\mathbf{T}^T = \mathbf{T}^{-1}$. Matice splňující tento vztah se nazývají ortogonální, což úzce souvisí se skutečností, že báze, mezi nimiž jsme transformovali, jsou obě ortonormální. Všimněte si, že pro ortogonální matice také platí $\mathbf{T}\mathbf{T}^T = \mathbf{I}$ a $(\mathbf{T}^T)^{-1} = (\mathbf{T}^{-1})^T = \mathbf{T}$.