Универзитет у Нишу
Електронски факултет

Математички методи у вештачком учењу

Аутор: Јована Цунић

Ниш, 2023

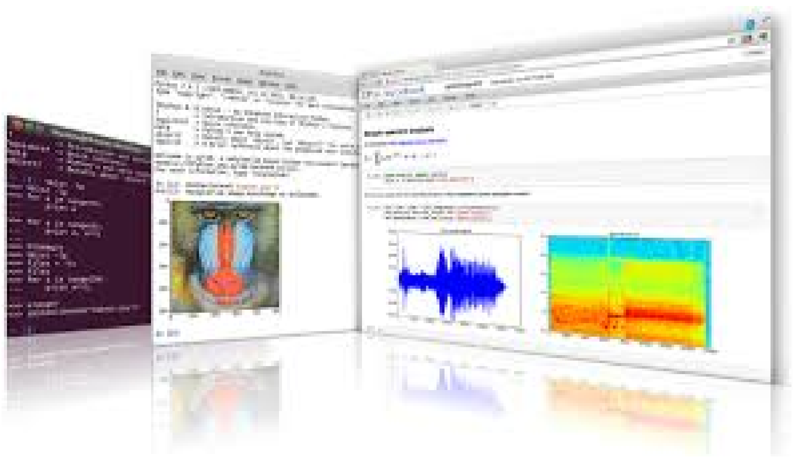
Садржај

Глава 1

Инсталација и основно о раду у **Python** програмском окружењу

Нумеричка математика и њени алгоритми развијају се са циљем имплементације теоријских резултата на некој рачунској машини. Од математичке теорије до имплементације је дуг пут који укључује дубоко разумевање особина и начина рада рачунске машине. Поред потпоре у теорији, примена нумеричких алгоритама захтева пуно експеримената и израчунавања са променама услова и вредности параметара, и најбитније - анализу добијених резултата. Са намером да се студенти подстакну на овакав приступ проучавању нумеричких алгоритама, у оквиру предмета Матрични методи представићемо основе интерактивног и нумеричког израчунавања у програмском окружењу Python. Користићемо програмско окружење за израчунавања преко уграђених функција програмског језика Python. Ово прво поглавље садржи информације потребне за инсталацију Python Jupyter и Python Idle окружења, и служи за почетно упознавање са њима. Наставак учења рада у Python-у одвијаће се кроз часове вежби и Python Jupyter радне свеске које нису саставни део уџбеника већ пратеће збирке задатака у е-формату.

Python је програмски језик високог нивоа и опште намене. Карактерише га синтакса која је једноставна и лака за читање. То га чини изузетно примамљивим инжењерима и научницима за креирање прототипова, дизајн симулација или тестирања, уз значајно мањи број кодних редова. Тиме се битно штеди на времену приликом дизајна програма или докумената. Лакоћа изражавања алгоритма кроз код чини Python језиком који не отежава већ подстиче креативни процес. Главна предност овог програмског окружења је лак пренос података и рутина кроз многе различите платформе и програмска окружења, као и постојање огромне базе поузданих open-source кодова. На тај начин програмирање у Python-у поспешује продуктивност у раду. Ипак, све ове предности долазе са оптерећењем времена извршења програма



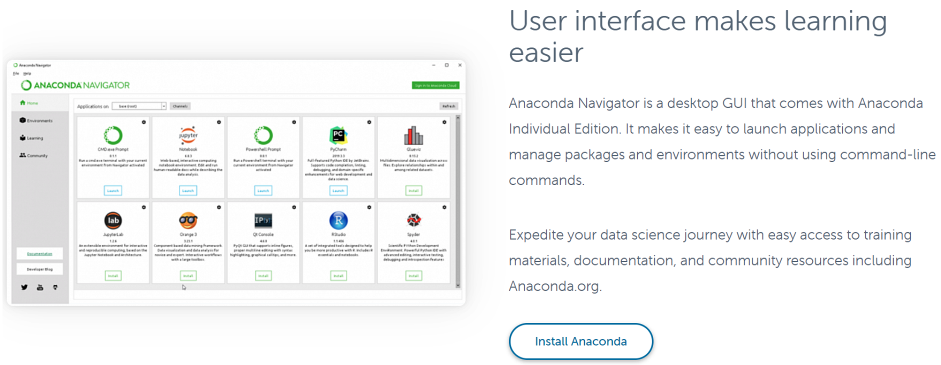
Слика 1.1: Изглед Jupyter Notebook

писаних у Python-у. Свеједно, овај програмски језик је одлично окружење за прве кораке у нумеричкој математици. То потврђују многи курсеви на основним студијама различитих усмерења и универзитета широм света у оквиру којих се обрађује Python за неку специјализовану примену.

За израчунавања у програмском језику Python неопходан је активан Python интерпретатор. У овом поглављу представићемо инсталацију два радна окружења Anaconda и IDLE, мада постоје и многа друга. Anaconda navigator се уобичајено користи online и представља системски захтевније окружење, али удобније за рад. За теме које се обрађују током курса и IDLE окружење пружа сву неопходну функционалност. IDLE конзола има далеко једноставнији кориснички интерфејс, уобичајен за програмске језике треће или четврте генерације (4GL). Ипак, препорука је да се, ако је могуће, за потребе предмета инсталира Anaconda Navigator. Jupyter радне свеске (Notebooks) као радно окружење интегришу много лепих корисничких додатака у једном. Због ових додатака резултати рада у Jupyter-у могу имати изглед веома професионалних пројектних извештаја, са интегрисаним и функционалним свим компонентама програмског кода за генерисање нумеричких и графичких резултата на лицу места.

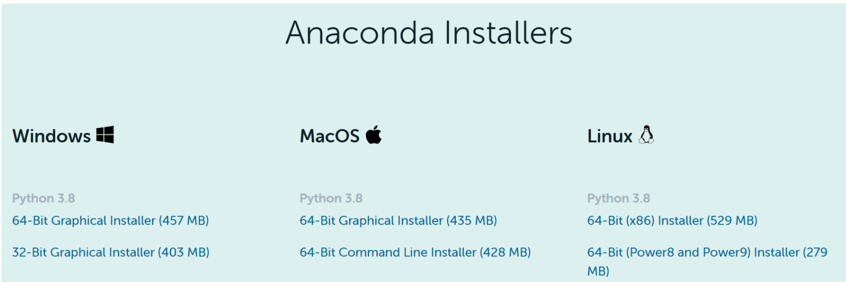
1.1 Инсталација **Anaconda Navigator**-а

Anaconda се може бесплатно преузети са званичне web стране. Кликом на дугме Install Anaconda (видети слику 1.2) отвара се опција за избор верзије. Одаберите инсталацију верзије која одговара конфигурацији вашег рачунара и оперативном систему.



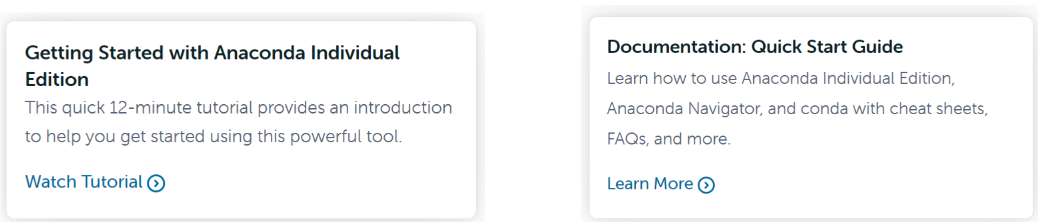
Слика 1.2: Почетак инсталације

По покретању инсталације пратите упутства Setup-а, тј. само прелазите на понуђен следећи корак. За активацију и рад Anaconda Navigator-а неопходно је одабрати адекватан оперативни систем рачунара, слика 1.3.



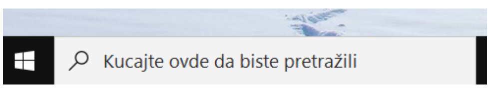
Слика 1.3: Опције инсталације

По завршетку инсталације појавиће се прозор о успеху уколико сте изабрали верзију са адекватним системским захтевима. На истом прозору постоје и линкови за 12-минутно видео упутство за снала жење у новом окружењу и линк за документациони центар, слика 1.4.



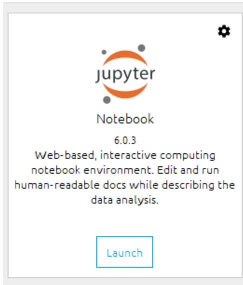
Слика 1.4: Линкови ка корисним ресурсима

Уколико користите Windows оперативни систем Anaconda Navigator-у, или директно Jupyter апликацији, можете приступити кроз search box, слика 1.5.



Слика 1.5: Windows search box

1.2 Jupyter и IPython



Слика 1.6: Jupiter покретање

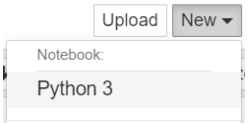
У оквиру Anaconda Navigator-а од интереса за овај курс су Jupyter и IPython (Interactive Python). Разлог лежи у посебном формату који је од 2011. године доступан у IPython-у – interactive Notebook, тј. Jupiter Notebook. У оквиру курса зваћемо га радна свеска. Jupyter се покреће кликом на одговарајућу апликацију унутар Anaconda Navigator-а, видети слику 1.6, или позивом саме апликације.

Формат рада радне свеске пружа битне предности у односу на стандардна програмска радна окружења. Jupyter радне свеске прављене су по угледу на популарна окружења за израчунавања попут MATLAB (live script, .mlx) и Wolfram Mathematica (Notebook, .nb). Jupyter радна свеска се покреће унутар Jupyter претраживача и омогућава обједињен web интерфејс где програмски код, текст, математичке једначине, графици и интерактивне контроле могу да се комбинују унутар јединственог документа. Унутар једне радне свеске могуће је покретање делова кода на извршење, бележење излаза програмске секвенце и документовање података или резултата уз пратеће текстуално објашњење. То је нарочито повољно у пројектима са сложеном анализом података или у приликама када решење једног проблема обухвата доста израчунавања и повезаности међурезултата. Више о овом окружењу може се наћи на званичној web страни. У оквиру курса Матрични методи инструкције за рад у Python-у, задаци на вежбама и пројектни задаци креирани су управо у овом формату. Полазници курса Матрични методи охрабрују се да своје домаће задатке предају такође у овом формату.

Радне свеске имају .ipynb екстензију. Могу се креирати као нов или већ постојећи документ у оквиру Jupyter web интерфејса, слика 1.7. Отпремање докумената у Jupyter окружење омогућено је опцијом Upload.

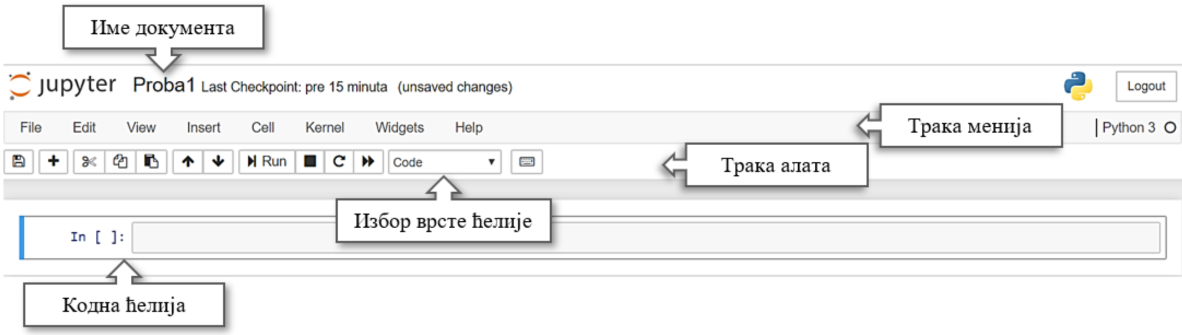
Слика 1.7: Управљање Jupyter радном свеском

Отварање новог документа у коме је могуће извршавање Python програмског кода подразумева избор Python3 опције у оквиру New падајућег енија, слика 1.8.



Слика 1.8: Отварање нове IPython радне свеске

Било да је нов или постојећи документ Python Notebook поседује исте компоненте. То су заглавље и ћелије. Заглавље носи важне информације о самом документу као и уређивачке механизме. Ћелије су основни елементи корисничког уноса унутар радне свеске.



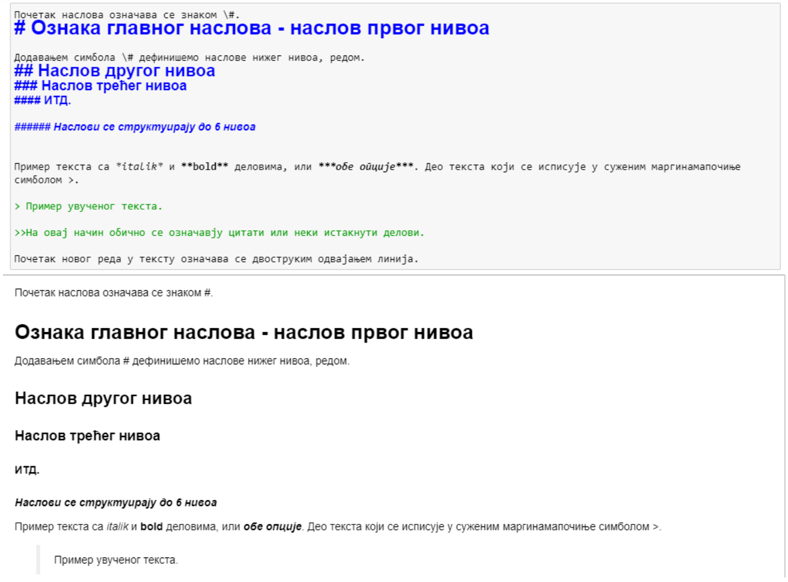
Слика 1.9: Заглавље

Представићемо најпре главне компоненте стандардног заглавља једне радне свеске означене на слици 1.9. Компоненте су описане према редоследу појављивања на слици 1.9, редом одозго на доле:

- Име документа екстензије ipynb може се дефинисати или мењати кликом на њега. На слици 1.9 је то име Proba1. У наставку, поред имена може се пратити статус опције Autosave.
- Кроз траку менија омогућен је приступ многим активностима које се тичу карактеристика радних свезака било као текстуалних докумената, било њених програмских делова. Трака менија на својој десној старни садржи статуса језгра тј. интерпретатора.
- Трака алата садржи иконе (Shortcuts) за уобичајене корисничке активности: чување документа, додавање ћелија, copy-paste актовности, покретање и паузирање програмског кода, промена редоследа ћелија, итд. Посебно је важан, падајући мени Code који омогућава измену типа ћелије уноса.
- Главне компоненте корисничког интерфејса радне свеске јесу ћелије. Радна свеска састоји се од линеарне листе ћелија. У зависности од избора врсте ћелије, њихов садржај ће се другачије третирати у интерпретатору. Садржај, тип и број ћелија може се мењати кроз траке менија и алата, и директном променом садржаја. На тај начин ћелије се могу додавати, брисати, примењивати copy-paste опције, вршити вертикално померање, итд.

Структура једне ћелије у радној свесци може бити Markdown или Code. Ови типови ћелија одговарају режиму текстуалних белешки и програмског кода, редом.

Пример једне Markdown ћелије са њеним излазом дат је на слици 1.10, а кодне на слици 1.11. Markdown ћелија служи за текстуалне описе у документу. Уз класич-



Слика 1.10: Улаз и излаз Markdown ћелије

не опције форматирања текста омогућава додавање линкова, слика и других HTML елемената, као и LaTeX математичких једначина итд.

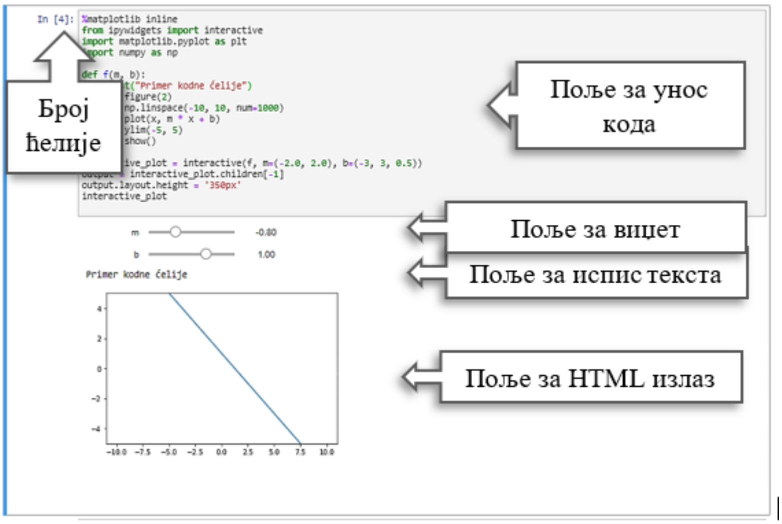
Све ћелије користе два режима рада, режим уређивања и командни режим. На горњем делу слике 1.10 приказан је режим уноса Markdown ћелије. Доњи део слике је у командом режиму и даје приказ одговарајуће Markdown ћелије након њене активације. Једним кликом на ћелију корисник је доводи у фокус. У фокус се могу ставити једна или више узастопних ћелија. У командном режиму рада, диркама стрелица горе-доле премешта се фокус по ћелијама. Стрелицама из траке алата вршимо вертикално померање позиције ћелија - замена места суседних ћелија. Ћелије у фокусу могу да се копирају, дуплирају или исецају (бришу).

Режим уређивања садржаја ћелије покреће се двоструким кликом на неку од ћелија или притиском тастера Enter док је одговарајућа ћелија у фокусу. У овом режиму садржај било које ћелије се може мењати. Излазак из режима уређивања постиже се притиском поља Run у траци алата, или притиском дирке ESC, или коришћењем неке комбинације тастера са тастатуре (Shift-Enter, Ctrl+Enter, Alt+Enter) за извршење ћелије.

Кодне ћелије, као што им само име сугерише, намењене су за програмски код који извршава језгро - Kernel. Програмски језик који одговара језгру може се мењати, али је овде од интереса искључиво Python интерпретатор. Садржај једне кодне ћелије се извршава на захтев корисника. .

Кодне ћелије састоје се од неколико делова, видети слику 1.11. Опис најважнијих од њих дат је у наставку:

- Кодна ћелија на својој левој страни почиње бројем ћелије. То је њен идентификациони број. Овај број се повећава са сваким извршењем кодне ћелије и на његову вредност се не може другачије утицати. Редослед извршавања кодних ћелија не мора бити линеаран, тј. не мора да одговара њиховом физичком распореду у оквиру радне свеске. Редослед извршења ових ћелија може да се



Слика 1.11: Кодна ћелија

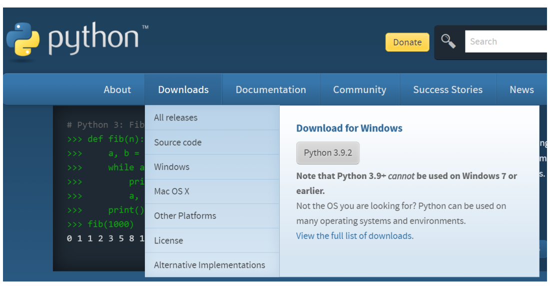
прати управо кроз број ћелије. Кодна ћелија се поставља у фокус кликом курсора на простор око броја ћелије.

- Поље за унос програмског кода је вишелинијски едитор текста који омогућава унос једне или више кодних линија. Садржај кодне ћелије је спреман за измене постављањем курсора на ово поље. Унос програмског кода подржава синтаксно истицање које је карактеристика Python језика. Другим речима, увлачењем реда описују се унутрашњи делови програмских блокова као што су гране и петље. Резервисане речи програмског језика наглашене су другом бојом у односу на имена уведених променљивих и др.

- Поље излаза може садржати више делова као што су део за виџет, стандардни текстуални испис, испис грешке, или HTML излаз (табеле и графици).

1.3 Инсталација **IDLE** едитора

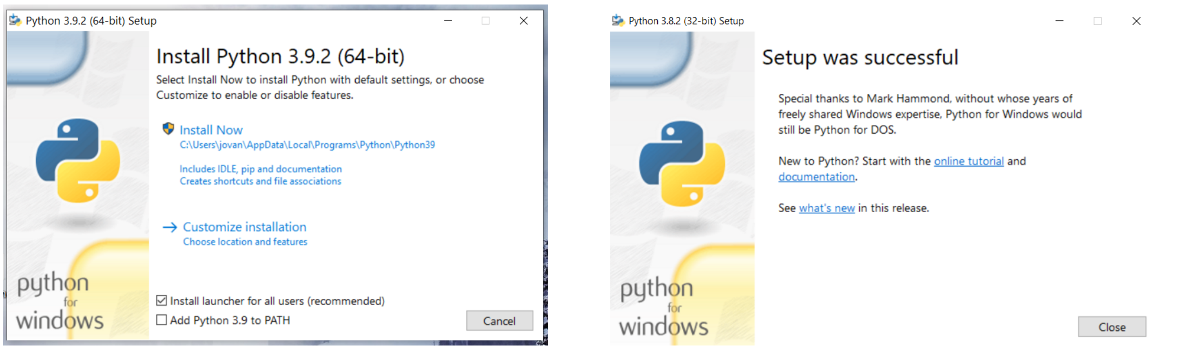
Уколико из било ког разлога вам није доступно Jupiter радно окружење, представљамо једну могућу резервну опцију. IDLE едитор доступан је за бесплатну ин-



Слика 1.12: Python Language Website и картица Downloads

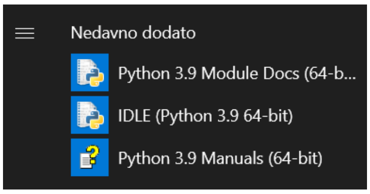
сталацију на сајту Python Language Website, слика 1.12 за верзију из 2020. године. У оквиру картице Downloads налазе се опције оперативног система на који се софтвер инсталира као и информације о актуелној верзији

Покретањем инсталације аутоматски се инсталирају и Python и едитор IDLE (Integrated DeveLopment Environment) који омогућава извршење кодова директно из овог едитора. Током инсталације потребно је да се прате упутства Setup-а.



Слика 1.13: Покретање и завршетак инсталацијом

По завршетку инсталације појавиће се порука о успешној инсталацији, слика 1.13. Из овог прозора можете одмах приступити онлајн приручнику или документационом центру. Сам IDLE едитор може се покренути на исти начин као Anaconda Navigator из Windows поља за претрагу, или из поља недавно додатих апликација у Windows менију, слика 1.14.



Слика 1.14: Покретање IDLE едитора

За математичке функције, симболичка и нумеричка израчунавања, рад са матрицама и графички приказ биће у употреби Python пакети (модули и библиотеке) Math, Numpy, Scipy, Matplotlib и Sympy. Потребне додатне пакете инсталираћете отварањем прозора PowerShell или Command prompt. PowerShell прозор отвара се притиском на SHIFT и десно дугме миша и избором Open PowerShell window here опције, слика 1.15.

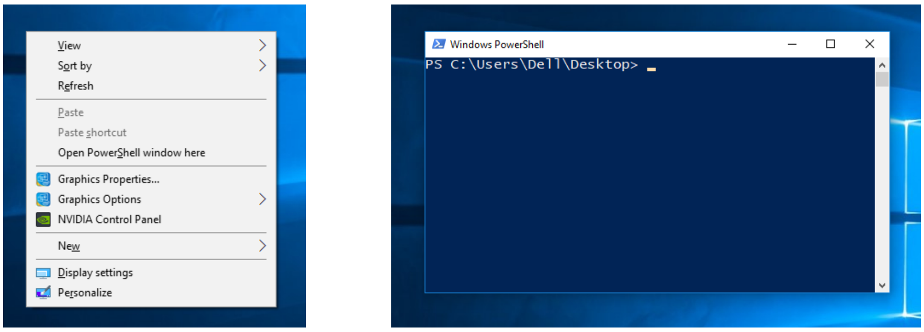
Када се појави PowerShell командни промпт као на слици 1.15, унесите наредбу за инсталацију пакета:

```
python -m pip install --user numpy scipy matplotlib sympy
```

По завршетку инсталација добићете поруку о успешности. Ажурирање pip верзије врши се следећом командом.

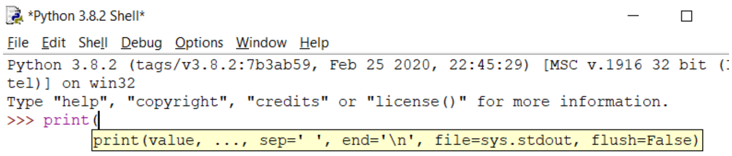
```
python -m pip install --upgrade pip
```

Окружење IDLE долази уз стандардну инсталацију Python система. И њега карактерише интерактивност у раду. Корисник може да унесе наредбу и одмах види



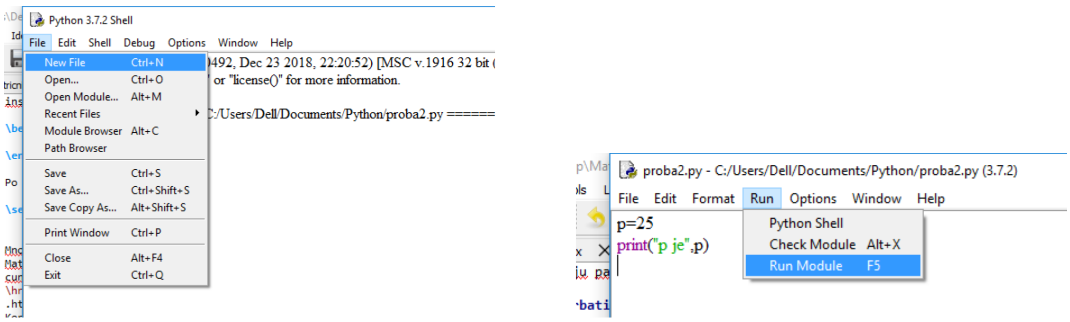
Слика 1.15: PowerShell

њене ефекте. Овакав начин рада омогућава да се грешке лако уочавају и исправљају. Интерактивно окружење је веома погодно за анализу појединачних корака у процесу решавања проблема израчунавања као и у процесу креирања и извођења експеримената. Окружење IDLE садржи интерактивну платформу, едитор, debugger, као и прегледач класа и модула програмског језика Python. Једна од згодних опција IDLE окружења је помоћ коју пружа за стандардне функције Python језика. Приликом уноса имена уграђене Python функције појави се мали прозор са навођењем низа аргумената те функције и линијом за кратак опис дејства саме функције, слика 1.16.



Слика 1.16: Idle синтаксна помоћ

Осим интерактивног рада IDLE пружа могућност и 'batch mode' уноса. Отварањем новог фајла у IDLE улази се у batch mode, слика 1.17.



Слика 1.17: Idle мени

Након уноса низа наредби потребно је запамтити документ са екстензијом .py и покренути Python интерпретатор позивом Run module из Run менија на врху прозора. Овакав начин рада одговара програмској структури уноса и погодује ситуацијама када је исти низ кодних секвенци потребно применити на различите скупове улазних података. Наредбе је потребно распоредити у логичке линије преласком у нови ред. Више о раду у овом окружењу можете наћи на Getting started with Python Idle.

Глава 2

Векторски простори

У овом поглављу подсећамо се појмова из теорије векторских и унитарних простора. Ширимо скуп анализираних објеката који поседују својства вектора.

2.1 Увод

Дефиниција 1. Уређена четворка $(V, \mathbb{K}, +, \cdot)$, представља векторски простор над пољем $\mathbb{K}$ уколико су дефинисане операције са особинама:

1. $V \neq \emptyset$, и $(\mathbb{K}, +_{\mathbb{K}}, \cdot_{\mathbb{K}})$ је поље,
2. $+ : V \times V \to V, \quad (V, +)$ је Абелова група,

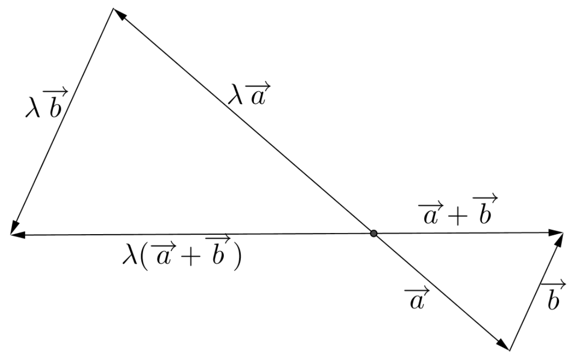
$$\begin{array}{ll} \text{а)} & (\forall v, u, w \in V)\ v + (u + w) = (v + u) + w, \\ \text{б)} & (\forall v, u \in V)\ v + u = u + v, \\ \text{в)} & (\exists \theta \in V)(\forall v \in V)\ v + \theta = \theta + v = v \\ \text{г)} & (\forall v \in V)(\exists -v \in V)\ v + (-v) = (-v) + v = \theta, \end{array}$$

3. $\cdot : \mathbb{K} \times V \to V, \quad \forall v, u \in V$ и $\forall \lambda, \mu \in \mathbb{K}$ важе следеће једнакости

$$\begin{array}{ll} \text{а)} & 1 \cdot v = v, \\ \text{б)} & \lambda \cdot (\mu \cdot v) = (\lambda \cdot_{\mathbb{K}} \mu) v, \\ \text{в)} & (\lambda +_{\mathbb{K}} \mu) \cdot v = \lambda \cdot v + \mu \cdot v \\ \text{г)} & \lambda \cdot (v + u) = \lambda \cdot v + \lambda \cdot u, \end{array}$$

Елементе скупа V називамо векторима, а елементе поља $\mathbb{K}$ скаларима. Операција $+$ зове се сабирање вектора, а операција $\cdot$ множење вектора скаларом. Операције сабирања и множења унутар поља $\mathbb{K}$ су у дефиницији 1 специјално наглашене $+_{\mathbb{K}}$ и $\cdot_{\mathbb{K}}$, мада их у наставку текста нећемо тако издвајати. Из контекста ће се унутар математичких формула јасно видети о којој од операција сабирања или множења је реч.

Бавићемо се углавном векторским просторима када је $\mathbb{K} = \mathbb{R}$, тј. реалним векторским просторима. Мада ће повремено бити од интереса да анализирамо и комплексни случај. Због тога су све дефиниције у наставку дате у случају реалних векторских простора. Тамо где комплесни случај доводи до разлика, то је специјално наглашено.



Слика 2.1: Сабирање и скалирање вектора за $\lambda < 0$

Приликом рада са векторима згодно је имати на уму геометријску интерпретацију основних операција, слика 2.1.

Кључни појам у векторским просторима је база. То је најмањи скуп вектора којим могу да се генеришу сви вектори векторског простора V. Начин генерисања јесте кроз линеарне комбинације.

Дефиниција 2. Скуп вектора $(b) \subset V$ векторског простора V представља базу ако има следеће две особине:

1. Било који коначан подскуп $b_1, b_2, \ldots, b_k \in (b)$ је линеарно независан, тј.
$$\alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_k b_k = \theta \implies \alpha_i = 0,\ \forall i.$$
2. Вектори скупа (b) чине потпун систем вектора, тј. за произвољан вектор $v \in V$ постоји коначан подскуп вектора $b_1, b_2, \ldots, b_k \in (b)$ тако да је v представив преко њих
$$v = \alpha_1 b_1 + \alpha_2 b_2 + \cdots + \alpha_k b_k. \tag{2.1}$$

Базу која задовољава услове дефиниције 2 зовемо Хамелова база векторског простора. Коефицијенти α_i у представљању (2.1) су јединствено одређени базом (b) и зовемо их координате вектора v у односу на ову базу.

Када се база векторског простора састоји од коначно много вектора, за векторски простор кажемо да је коначно-димензионалан. Број елемената произвољне базе векторског простора V зовемо димензија тог простора.

Теорема 1. За коначно димензионалан векторски простор V димензије n скуп од n вектора представља базу ако је линеарно независан или ако је потпун.

Другим речима, ако је скуп од n вектора $n-$димензионалног простора V линеарно независан, онда тај скуп вектора представља базу у V. Ако је скуп од n вектора $n-$димензионалног простора V потпун, онда је тај скуп база у V.

Дефиниција 3. Подскуп $P \subseteq V$ векторског простора V над пољем скалара $\mathbb{K}$ представља потпростор ако је $(P, \mathbb{K}, +, \cdot)$ простор у односу на рестрикције операција $+$ и $\cdot$ на подскуп P.

Теорема 1. Подскуп $P \subseteq V$ векторског простора V је потпростор акко
$$\forall v, u \in P,\ \forall \lambda, \mu \in \mathbb{K}:\ \lambda v + \mu u \in P.$$

Линеарним комбинацијама базе векторског простора описујемо све векторе тог простора или неке његове подскупове. Линеал или разапињање је векторски простор:

$$\mathcal{L}(v_1, v_2, \ldots, v_k) = \{\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k \big| \ \lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}\}.$$

Линеал над векторима $v_1, v_2, \ldots, v_k$ је најмањи векторски простор који садржи ове векторе, тј. најмањи простор разапет над тим векторима. Линеал $\mathcal{L}(v_1, v_2, \ldots, v_k)$ садржи све линеарне комбинације вектора $v_1, v_2, \ldots, v_k$. Због тога они чине потпун систем вектора за тај линеал. Међу векторима $v_1, v_2, \ldots, v_k$ онда можемо да пронађемо бар једну базу простора $\mathcal{L}(v_1, v_2, \ldots, v_k)$. Максималан број линеарно независних вектора скупа $v_1, v_2, \ldots, v_k$ одређује димензију линеала.

Пример 2.1.1. Скуп вектора $v = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $u = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ је линеарно независан и представља базу линеала над овим векторима,

$$P = \mathcal{L}(u, v) = \{\lambda u + \mu v \mid \lambda, \mu \in \mathbb{R}\} = \left\{ \begin{bmatrix} \lambda \\ \lambda \\ \mu \end{bmatrix} \ \Big| \ \lambda, \mu \in \mathbb{R} \right\},$$

али не и базу простора $\mathbb{R}^3$. Димензије ових простора су редом

$$\dim(P) = 2, \qquad \dim(\mathbb{R}^3) = 3$$

и P је прави потпростор простора $\mathbb{R}^3$.

Вектори u, v и $w = u + v$ су линеарно зависни, па заједно не представљају базу векторских простора P или $\mathbb{R}^3$. Приметимо да важи $P = \mathcal{L}(u, v) = \mathcal{L}(u, v, w)$. Заиста,

$$\begin{aligned} \lambda u + \mu v + \nu w = \lambda u + \mu v + \nu(u + v) &= (\lambda + \nu)u + (\mu + \nu)v \in \mathcal{L}(u, v) \\ \implies \quad & \mathcal{L}(u, v, w) \subseteq \mathcal{L}(u, v), \\ \lambda u + \mu v = \lambda u + \mu v + 0 \cdot w &\in \mathcal{L}(u, v, w) \\ \implies \quad & \mathcal{L}(u, v) \subseteq \mathcal{L}(u, v, w). \end{aligned}$$

Претходни пример сугерише једно општије тврђење за потпросторе описане преко линеала. Доказ овог тврђења остављамо као лаку вежбу читаоцима.

Лема 1. Нека је $P = \mathcal{L}(v_1, v_2, \ldots, v_k)$ и $u_1, u_2, \ldots, u_m \in P$. Тада је

$$\mathcal{L}(u_1, u_2, \ldots, u_m) \subseteq P \quad \wedge \quad \mathcal{L}(v_1, v_2, \ldots, v_k) = \mathcal{L}(v_1, v_2, \ldots, v_k, u_1, u_2, \ldots, u_m).$$

Осим тога, за $\alpha \in \mathbb{R}$ важи

$$\mathcal{L}(v_1, v_2, \ldots, v_k) = \mathcal{L}(v_1, v_2, \ldots, v_k + \alpha v_1).$$

Дефиниција 4. Нека су U и V векторски простори над истим пољем скалара $\mathbb{K}$. Пресликавање $\mathcal{A} : U \to V$ назива се оператор. Уколико оператор $\mathcal{A}$ има особину

$$(\forall u, v \in U)(\forall \alpha, \beta \in \mathbb{K}) \quad \mathcal{A}(\alpha u + \beta v) = \alpha \mathcal{A} u + \beta \mathcal{A} v,$$

за њега кажемо да је линеаран.

Када је линеаран оператор $\mathcal{A}$ бијекција називамо га изоморфизам простора U и V. Изоморфизам простора у самог себе је аутоморфизам тог простора.

Теорема 2. Изоморфизам слика скуп линеарно независних вектора у линеарно независан скуп.

Теорема 3. Изоморфизам слика базу у базу.

Теорема 4. Два векторска простора над истим пољем скалара су изоморфна акко имају једнаке димензије.

У наставку упознајемо конкретне векторске просторе којима ћемо се бавити.

2.2 Уређене $n-$торке

За уређене $n-$торке реалних или комплексних бројева (поистовећиваћемо их са матрицама-колонама и матрицама-врстама) користимо покомпонентне дефиниције операција,

$$\begin{aligned} \big(x_1,\ x_2,\ \dots,\ x_n\big) + \big(y_1,\ y_2,\ \dots,\ y_n\big) &= \big(x_1+y_1,\ x_2+y_2,\ \dots,\ x_n+y_n\big), \\ \alpha\big(x_1,\ x_2,\ \dots,\ x_n\big) &= \big(\alpha\, x_1,\ \alpha\, x_2,\ \dots,\ \alpha\, x_n\big). \end{aligned}$$

Нула-вектор је $\theta = \big(0,\ 0,\ \dots,\ 0\big)$. Природну базу овог векторског простора чине вектори

$$\begin{aligned} e_1 &= \big(1,\ 0,\ \dots,\ 0,\ 0\big), \\ e_2 &= \big(0,\ 1,\ \dots,\ 0,\ 0\big), \\ &\ \vdots \\ e_{n-1} &= \big(0,\ 0,\ \dots,\ 1,\ 0\big), \\ e_n &= \big(0,\ 0,\ \dots,\ 0,\ 1\big). \end{aligned}$$

Због тога је векторски простор уређених $n-$торки димензије n. Овај векторски простор је најчешће коришћен јер се сваки други векторски простор исте димензије поистовећује са њим. Наиме, избором и фиксирањем једне базе $(b)\ :\ \{b_1, b_2, \dots, b_n\}$ у векторском простору V сви вектори $v \in V$ се идентификују кроз координате у тој бази

$$v = x_1 b_1 + x_2 b_2 + \dots + x_n b_n \equiv \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T_{(b)}.$$

Основне операције векторског простора V на тај начин постају операције над уређеним $n-$торкама координата. Резултате добијамо једнозначно одређене својим координатама у бази (b).

$$\begin{aligned} & v = x_1 b_1 + x_2 b_2 + \dots + x_n b_n, \quad u = y_1 b_1 + y_2 b_2 + \dots + y_n b_n, \\ & v + u = (x_1+y_1) b_1 + (x_2+y_1) b_2 + \dots + (x_n+y_n) b_n \\ & \iff \quad \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T_{(b)} + \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T_{(b)} = \begin{bmatrix} x_1+y_1 & x_2+y_2 & \dots & x_n+y_n \end{bmatrix}^T_{(b)}, \end{aligned}$$

$$\begin{aligned} & \alpha v = (\alpha\, x_1) b_1 + (\alpha\, x_2) b_2 + \dots + (\alpha\, x_n) b_n \\ & \iff \quad \alpha \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T_{(b)} = \begin{bmatrix} \alpha\, x_1 & \alpha\, x_2 & \dots & \alpha\, x_n \end{bmatrix}^T_{(b)}. \end{aligned}$$

Приметимо да су координате вектора базе (b) у односу на њу дате са:

$$\begin{aligned} b_1 &= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}_{(b)}^T, \\ b_2 &= \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \end{bmatrix}_{(b)}^T, \\ &\vdots \\ b_n &= \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \end{bmatrix}_{(b)}^T. \end{aligned}$$

Координате нула-вектора θ у односу на било коју базу су увек $\theta = \begin{bmatrix} 0 & 0 & \dots & 0 \end{bmatrix}_{(b)}^T$.

Испитивање линеарне независности вектора $v_1, v_2, \dots, v_k$ такође можемо свести на испитивање линеарне независности вектора њихових координата у односу на неку фиксирану базу $(b) \ : \ \{b_1, b_2, \dots, b_n\}$. Нека су

$$v_i = x_{1i}b_1 + x_{2i}b_2 + \dots + x_{ni}b_n = \begin{bmatrix} x_{1i} & x_{2i} & \dots & x_{ni} \end{bmatrix}_{(b)}^T, \quad i \in \{1, 2, \dots, k\}.$$

Тада је због јединствености репрезентације вектора преко координата

$$\begin{aligned} & \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = \theta \\ \iff \quad & \alpha_1 \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}_{(b)} + \alpha_2 \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{bmatrix}_{(b)} + \dots + \alpha_k \begin{bmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{nk} \end{bmatrix}_{(b)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(b)}. \end{aligned}$$

Сва показана израчунавања преко координата су повод да теоријску анализу у наставку везујемо управо за векторе координата у односу на неку фиксирану базу од интереса.

2.3 Матрице

Матрице реалних или комплексних бројева такође имају покомпонентно дефинисане операције,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix},$$

$$\alpha \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{bmatrix}.$$

Нула-вектор зовемо нула матрица $O = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$.

Једну базу векторског простора реалних матрица над $\mathbb{R}$ (или комплексних матрица над $\mathbb{C}$) чине матрице:

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \dots, \quad E_{1n} = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, \dots, \quad E_{2n} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

$$\vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots$$

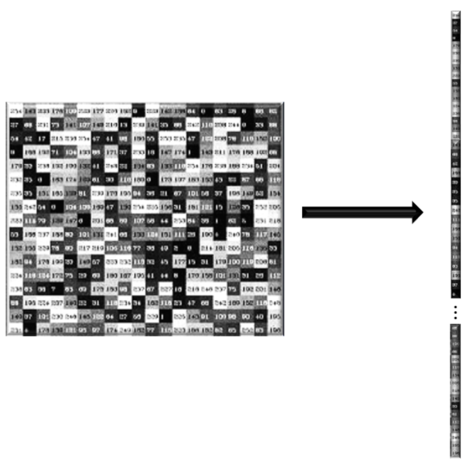
$$E_{m1} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}, \quad E_{m2} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}, \dots, \quad E_{mn} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Матрица E_{ij} има јединицу на позицији i, j и све остале елементе једнаке нули. Тако је векторски простор $\mathcal{M}_{m\times n}$ димензије mn. Због тога кажемо да су простори $\mathcal{M}_{m\times n}$ и $\mathbb{R}^{mn}$ изоморфни, што записујемо

$$\mathcal{M}_{m\times n} \cong \mathbb{R}^{mn}.$$

Ипак, да бисмо нагласили карактеристику матрица као табеле чијим се елементима приступа са два индекса, користимо ознаку $\mathcal{M}_{m\times n} = \mathbb{R}^m \otimes \mathbb{R}^n$ или $\mathcal{M}_{m\times n} = \mathbb{R}^{m\times n}$. Овакву конструкцију нових векторских простора, кроз операцију означену са $\otimes$, зовемо тензорски производ простора.

Када матрице посматрамо као елементе векторског простора користимо тзв. приступ векторизације тих матрица. Другим речима, елементе матрице A сматрамо дугачким низовима вредности $vec(A)$, слика 2.2. Овакав приступ погодује неким проблемима.



Слика 2.2: Векторизација матрице

Пример 2.3.1. Пребацивање дводимензионалне табеле бројева у један дуг низ можемо да вршимо надовезивањем врста или надовезивањем колона те матрице. Рецимо, матрицу $A = [a_{ij}] \in \mathcal{M}_{2\times 3}$ са елементима $a_{ij} = i + j$ представићемо као векторе њених вредности. Користићемо Python начин индексирања NumPy низова, тј. $A = [a_{ij}]_{i=\overline{0,1},j=\overline{0,2}}$.

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix} \overset{\text{Надовезивањем врста}}{\mapsto} A_{[0]} = \big(0,\ 1,\ 2,\ |\ 1,\ 2,\ 3\big)$$

$$\overset{\text{Надовезивањем колона}}{\mapsto} A_{[1]} = \big(0,\ 1,\ |\ 1,\ 2,\ |\ 2,\ 3\big).$$

Технички, ове векторизације се добијају нпр. секвенцијалним обилажењем елемената матрице помоћу две угњеждене петље. Разлика настаје избором распореда бројача по дубини, тј. брзине промене њихових вредности.

Надовезивање врста:

$$\begin{aligned} & pozicija := 0 \\ & for\ i = \overline{0,1}: \\ & \quad for\ j = \overline{0,2}: \\ & \qquad vec(pozicija) := a_{ij} \\ & \qquad pozicija{+}{+} \end{aligned}$$

Надовезивање колона:

$$\begin{aligned} & pozicija := 0 \\ & for\ j = \overline{0,2}: \\ & \quad for\ i = \overline{0,1}: \\ & \qquad vec(pozicija) := a_{ij} \\ & \qquad pozicija{+}{+} \end{aligned}$$

Формуле за добијање позиције елемента у новом вектору зависе од i, j и димензија матрице:

$$\begin{aligned} &\text{Надовезивањем врста:} && (i,j) \ \mapsto\ 3i + j, \\ &\text{Надовезивањем колона:} && (i,j) \ \mapsto\ i + 2j. \end{aligned}$$

У општем случају, за матрицу $A \in \mathcal{M}_{m\times n}$ векторизација је пресликавање

$$\begin{aligned} \mathcal{M}_{m\times n} &\mapsto \mathbb{R}^{mn} \\ A &\mapsto vec(A) \end{aligned}$$

где се $vec(A) = A_{[0]}$ добија надовезивањем врста или $vec(A) = A_{[1]}$ надовезивањем колона. Формуле трансформације индекса гласе (са Python начином индексирања):

$$\begin{aligned} &\text{Надовезивањем врста:} && (i,j) \ \mapsto\ n\,i + j, \\ &\text{Надовезивањем колона:} && (i,j) \ \mapsto\ i + m\,j. \end{aligned}$$

Ове формуле могу да послуже и за реконструкцију матрице A на основу вектора $vec(A)$,

$$\begin{aligned} vec(A) = A_{[0]} \ &: k \ \mapsto\ (i,j), \quad i = k//n,\ j = k\%n, \\ vec(A) = A_{[1]} \ &: k \ \mapsto\ (i,j), \quad j = k//m,\ i = k\%m, \end{aligned}$$

где су // и % целобројно дељење и остатак при целобројном дељењу, редом.

Осим третирања матрица као једног вектора, саме матрице је корисно посматрати кроз скупове вектора: вектора-колона и вектора-врста. На тај начин матрице индукују још неке веома важне векторске просторе. Тако матрицу $A = [a_{ij}] \in \mathcal{M}_{m\times n}$, можемо 'разбити' на два, у општем случају различита, скупа вектора:

$$A = \begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} & \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} & \dots & \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{bmatrix}, \qquad A = \begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix} \\ \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{bmatrix}.$$

Уведемо ознаке за назначене векторе врста и колона. За матрицу $A \in \mathcal{M}_{m\times n}$

$$A = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \dots & u_n \end{array}\right] = \left[\begin{array}{c} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_m^T \end{array}\right],$$

четири фундаментална простора су:

- простор колона $\quad \mathcal{R}(A) = \mathcal{L}(u_1, u_2, \dots, u_n) = \left\{Av \;\middle|\; v \in \mathbb{R}^n\right\} \subseteq \mathbb{R}^m,$
- простор врста $\quad \mathcal{R}(A^T) = \mathcal{L}(v_1, v_2, \dots, v_m) = \left\{A^T v \;\middle|\; v \in \mathbb{R}^m\right\} \subseteq \mathbb{R}^n,$
- (десно) језгро матрице $\quad \mathcal{N}(A) = \left\{v \in \mathbb{R}^n \;\middle|\; Av = \theta\right\},$
- лево језгро матрице $\quad \mathcal{N}(A^T) = \left\{v \in \mathbb{R}^m \;\middle|\; v^T A = \theta\right\}.$

Теорема 5. Фундаментална теорема линеарне алгебре : За матрицу $A \in \mathcal{M}_{m\times n}$ важи

$$\begin{aligned} &\operatorname{rang}(A) = \dim \mathcal{R}(A) = \dim \mathcal{R}(A^T), \\ &\operatorname{def}(A) = \dim \mathcal{N}(A), \quad \operatorname{rang}(A) + \operatorname{def}(A) = n. \end{aligned} \tag{2.2}$$

Дефиниција 5. Матрица $A \in \mathcal{M}_{m\times n}$ је пуног ранга уколико јој је ранг једнак мањој од две димензије m или n, тј. $\operatorname{rang}(A) = \min\{m, n\}$. Специјално, матрица је пуног ранга врста када је $\operatorname{rang}(A) = m \le n$, а пуног ранга колона у случају $\operatorname{rang}(A) = n \le m$.

Осим операција векторских простора матрице су снабдевене и операцијом транспоновања.

$$A = [a_{ij}]_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad A^T = [a_{ji}]_{n\times m} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Ово је операција изоморфизма између векторских простора $\mathbb{R}^m \otimes \mathbb{R}^n$ и $\mathbb{R}^n \otimes \mathbb{R}^m$ с обзиром да се слаже са основним операцијама сабирања и скалирања,

$$(\alpha A)^T = \alpha\, A^T, \qquad (A + B)^T = A^T + B^T.$$

Особине основних операција над матрицама су већ познате од раније. У овом одељку подсетићемо се различитих интерпретација операције множења две матрице. Свака од интерпретација доноси идеје које су корисне како за имплементацију алгоритама, тако и за интерпретацију добијених резултата.

Множење матрица је могуће ако је испоштовано правило унутрашњег индекса тј. ако је испоштована шема:

$$\begin{bmatrix} & & \\ m & \text{врста} & \\ n & \text{колона} & \\ & & \end{bmatrix}_{m\times \underline{n}} \begin{bmatrix} & & \\ n & \text{врста} & \\ p & \text{колона} & \\ & & \end{bmatrix}_{\underline{n}\times p} = \begin{bmatrix} & & \\ m & \text{врста} & \\ p & \text{колона} & \\ & & \end{bmatrix}_{m\times p}.$$

Најзначајнији примери множења две матрице свакако су:

1. Унутрашњи производ два вектора:

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \dots + x_ny_n.$$

   У матричним изразима препознајемо га у запису

$$v^Tu \quad \text{где су} \quad v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},\ u = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

   Матрицу димензије 1×1 третирамо као скалар.

2. Спољашњи производ два вектора:

$$vu^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ \vdots & \vdots & & \vdots \\ x_my_1 & x_my_2 & \dots & x_my_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1\begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \\ x_2\begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \\ \vdots \\ x_m\begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} \end{bmatrix} = \begin{bmatrix} y_1\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} & y_2\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} & \dots & y_n\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \end{bmatrix}.$$

   Све врсте резултата су једнаке линеарној комбинацији једног вектора, вектора u^T, тј. $\mathcal{R}(A^T) = \mathcal{L}(u)$. Такође, све колоне представљају скалирану верзију вектора v, дакле $\mathcal{R}(A) = \mathcal{L}(v)$. Закључујемо да се операцијом спољашњег производа добија матрица ранга 1. Ове матрице играју веома важну улогу у поступцима декомпозиције других матрица.

   За разумевање примене операције множења веома је битно сагледавање матрице као колекције вектора: скуп њених вектора-колона или вектора-врста. У складу са тим дати су описи наредних операција.

3. Матрица пута вектор (десна нотација) је начин записивања линеарне комбинације вектора колона матрице A :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Коефицијенти $x_1, x_2, \dots, x_n$ линеарне комбинације вектора колона налазе се у компонентама вектора v.

4. Вектор пута матрица (лева нотација):

$$\begin{bmatrix} y_1 & y_2 & \dots & y_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = y_1 \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix} +$$

$$+ y_2 \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{bmatrix} + \dots + y_m \begin{bmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

је линеарна комбинација вектора-врста матрице A. Како је $(uA)^T = A^T u^T$, све речено за десно множење вектором аналогно се преноси на лево множење. Транспоновање је изоморфизам који омогућава аналогију двеју нотација. Због тога све што будемо у наставку доказивали, а тиче се десне нотације, аналогно ће важити и за леву нотацију.

Пример 2.3.2. Одредићемо број аритметичких операција употребљених за израчунавање унутрашњег производа два реална вектора

$$v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T, \qquad u = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T.$$

Израз

$$v^T u = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

користи n множења и $n-1$ сабирања. Ова операција над векторима је операцијске сложености $2n - 1 = \mathcal{O}(2n)$. Због тога што је операцијска сложеност са линеарном зависношћу од количине података n за ову операцију се каже да је нивоа 1.

Пример 2.3.3. Потражићемо операцијску сложеност спољашњег производа два реална вектора

$$v = \begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix}^T \in \mathbb{R}^m, \qquad u = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T \in \mathbb{R}^n.$$

Спољашњи производ

$$vu^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_n \\ \vdots & \vdots & & \vdots \\ x_m y_1 & x_m y_2 & \dots & x_m y_n \end{bmatrix}$$

представља m скалирања $n-$димензионалног вектора (или n скалирања $m-$димензионалног вектора) па користи mn операција множења. Операцијска сложеност спољашњег производа је тако квадратна функција обима података и каже се да спада у операције другог нивоа. Подаци су линеарног обима $n+m$, док је број операција у поступку израчунавања квадратна функција.

Пример 2.3.4. Спољашњи производ се веома често јавља као део елементарне трансформације алгоритама над матрицама. За два вектора

$$v = \begin{bmatrix} x_1 & x_2 & \dots & x_m \end{bmatrix}^T \in \mathbb{R}^m, \qquad u = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T \in \mathbb{R}^n,$$

и матрицу $A = \begin{bmatrix} a_{ij} \end{bmatrix}_{m\times n}$ то су изрази облика

$$A + vu^T$$

и називамо их ажурирање матрице A ранга 1. Циљ оваквих трансформација јесте да се $j-$та колона матрице A промени за вредност вектора $y_j v,\ j = 1, 2, \dots, n$. Други начин тумачења израза ажурирања јесте кроз врсте матрице A : $i-$та врста матрице A мења се за вектор $x_i u^T,\ i = 1, 2, \dots, m$.

Пример 2.3.5. Одредићемо операцијску сложеност операција матрица $\cdot$ вектор. Нека је $A \in \mathcal{M}_{n\times n}$ и $v \in \mathcal{M}_{n\times 1}$. Производ Av представља израчунавање n унутрашњих производа облика:

$$\text{врста матрице } A \cdot \text{вектор } v.$$

С обзиром да сваки од ових унутрашњих производа кошта n множења и $n-1$ сабирања бројева (пример 2.3.2), то производ Av троши n^2 множења и $n(n-1)$ сабирања, што је укупно $\mathcal{O}(2n^2)$ операција. Видимо да је Av операција другог нивоа, с тим што су и подаци и обим израчунавања квадратног обима.

Специјални облици матрица смањују операцијску сложеност овог производа. Познавање распореда нула елемената унутар матрице у томе има важну улогу. Најједноставнији пример је производ

$$\operatorname{diag}(u)\, v = \begin{bmatrix} u_1 & 0 & \dots & 0 \\ 0 & u_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 \\ u_2 v_2 \\ \vdots \\ u_n v_n \end{bmatrix}.$$

Закључујемо да израз $\operatorname{diag}(u)\, v$ троши n операција множења (а не n^2). То је повод за наредну дефиницију.

Дефиниција 6. За два вектора дужине n :

$$v = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T \quad \text{и} \quad u = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}^T,$$

операција Адамаровог множења вектора, тј. производ вектора члан по члан дат је са

$$u * v = \begin{bmatrix} u_1 v_1 & u_2 v_2 & \dots & u_n v_n \end{bmatrix}^T. \tag{2.3}$$

Операција покомпонентног множења два вектора истих димензија се природно преноси на матрице истих димензија.

Дефиниција 7. За две матрице истих димензија $A, B \in \mathcal{M}_{m\times n}$, $A = [a_{ij}]$, $B = [b_{ij}]$ Адамаров производ је

$$A * B = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \dots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \dots & a_{2n}b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \dots & a_{mn}b_{mn} \end{bmatrix}. \tag{2.4}$$

Јасно је да је Адамаров производ комутативна и асоцијативна операција на скупу матрица $\mathcal{M}_{m\times n}$. Наиме ради се само о симултаном множењу парова реалних или комплексних бројева. Јединични (неутрални) елемент ове операције је матрица састављена од свих јединица, $\mathbb{1}_{m\times n} = [1]_{m\times n}$.

Постоји још интересантних и веома корисних операција 'множења' над матрицама које представљају различите комбинације производа елемената и појединих делова (блокова) матрице. Упознаћемо их кроз Python радне свеске.

Класично множење две матрице може се описати преко било које од претходно наведених 'операција' множења: унутрашњег, спољашњег производа вектора или производа вектора и матрице у једном или другом смеру.

Нека су $A \in \mathcal{M}_{m\times n}$ и $B \in \mathcal{M}_{n\times p}$ две матрице дате својим елементима,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix}.$$

Производ матрица дефинисан је у смеру AB и можемо га тумачити на следеће начине:

1. Множење матрица $C = AB$ дефинисано је задавањем појединачних елемената c_{ij} резултујуће матрице $C = [c_{ij}] \in \mathcal{M}_{n\times p}$. Вредности елемената гласе

$$\begin{aligned} c_{ij} &= \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \\ &= a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} \\ &= \sum_{k=1}^{n} a_{ik}b_{kj}, \end{aligned}$$

тј. сваки елемент резултујуће матрице је унутрашњи производ одговарајуће врсте матрице A и колоне матрице B. За овакву дефиницију множења кључно је представљање матрица чинилаца у облику:

$$AB = \begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix} \\ \begin{bmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} & \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} & \dots & \begin{bmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{np} \end{bmatrix} \end{bmatrix}.$$

Тако је производ реалних матрица AB симултано израчунавање унутрашњих производа вектора врста из A и вектора колона из B. Сви парови врста-колона унутрашњих производа чине елементе матрице $C = AB$.

2. Производ AB преко спољашњег производа вектора дат је са

$$\begin{aligned}
AB &= \begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} & \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} & \dots & \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \end{bmatrix} \\ \begin{bmatrix} b_{21} & b_{22} & \dots & b_{2p} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \end{bmatrix} + \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} & \dots & b_{2p} \end{bmatrix} + \\
&\quad \dots + \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \begin{bmatrix} b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix}.
\end{aligned}$$

Горња једнакост може се лако проверити. Наиме, у првом сабирку

$$\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1p} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & \dots & a_{11}b_{1p} \\ a_{21}b_{11} & a_{21}b_{12} & \dots & a_{21}b_{1p} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \dots & a_{m1}b_{1p} \end{bmatrix}$$

на месту ij налази се члан $a_{i1}b_{1j}$. Слично, у другом сабирку

$$\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} \begin{bmatrix} b_{21} & b_{22} & \dots & b_{2p} \end{bmatrix} = \begin{bmatrix} a_{12}b_{21} & a_{12}b_{22} & \dots & a_{12}b_{2p} \\ a_{22}b_{21} & a_{22}b_{22} & \dots & a_{22}b_{2p} \\ \vdots & \vdots & & \vdots \\ a_{m2}b_{21} & a_{m2}b_{22} & \dots & a_{m2}b_{2p} \end{bmatrix}$$

на месту ij налази се члан $a_{i2}b_{2j}$. Настављајући поступак до последњег $n-$тог сабирка

$$\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \begin{bmatrix} b_{n1} & b_{n2} & \dots & b_{np} \end{bmatrix} = \begin{bmatrix} a_{1n}b_{n1} & a_{1n}b_{n2} & \dots & a_{1n}b_{np} \\ a_{2n}b_{n1} & a_{2n}b_{n2} & \dots & a_{2n}b_{np} \\ \vdots & \vdots & & \vdots \\ a_{mn}b_{n1} & a_{mn}b_{n2} & \dots & a_{mn}b_{np} \end{bmatrix},$$

и његовог елемента на позицији ij, добијамо све елементе збира којим је дефинисан члан c_{ij} :

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Посматрано на овај начин, производ матрица AB представља збир n спољашњих производа вектора:

колона матрице $A\cdot$ врста матрице B.

Производ AB је представљен као збир n матрица ранга 1.

3. Растављањем само матрице B на колоне можемо анализирати производ AB и уз помоћ 'операције' матрица·вектор.

$$AB = \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} & \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} & \dots & \begin{bmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{np} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} A\begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} & A\begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} & \dots & A\begin{bmatrix} b_{1p} \\ b_{2p} \\ \vdots \\ b_{np} \end{bmatrix} \end{bmatrix}.$$

Прва колона резултујуће матрице добија се у производу матрице A и прве колоне матрице B, и тако редом. Множење две матрице може се тумачити као p линеарних комбинација колона матрице A чији се резултати чувају у колонама матрице $C = AB$.

4. Интерпретација леве нотације множења вектор · матрица је аналогна десној путем транспоновања. За потребе леве нотације матрица A се представља преко својих вектора-врста.

$$AB = \begin{bmatrix} [\, a_{11} & a_{12} & \dots & a_{1n}] \\ [\, a_{21} & a_{22} & \dots & a_{2n}] \\ & \vdots & & \\ [a_{m1} & a_{m2} & \dots & a_{mn}] \end{bmatrix} \begin{bmatrix} & & \\ & B & \\ & & \end{bmatrix} = \begin{bmatrix} [\, a_{11} & a_{12} & \dots & a_{1n}]\, B \\ [\, a_{21} & a_{22} & \dots & a_{2n}]\, B \\ & \vdots & & \\ [a_{m1} & a_{m2} & \dots & a_{mn}]\, B \end{bmatrix}.$$

Пример 2.3.6. Производ матрица $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ и $B = \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 3 & 1 & 2 \end{bmatrix}$ у складу са претходно наведеним тумачењима операције множења гласи:

- кроз унутрашње производе

$$AB = \begin{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \\ \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} & \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} & \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 14 & 4 & 3 \\ 32 & 10 & 3 \end{bmatrix}$$

- кроз спољашње производе

$$AB = \begin{bmatrix}1\\4\end{bmatrix} \cdot \begin{bmatrix}1 & 1 & -1\end{bmatrix} + \begin{bmatrix}2\\5\end{bmatrix} \cdot \begin{bmatrix}2 & 0 & -1\end{bmatrix} + \begin{bmatrix}3\\6\end{bmatrix} \cdot \begin{bmatrix}3 & 1 & 2\end{bmatrix}$$
$$= \begin{bmatrix}1 & 1 & -1\\4 & 4 & -4\end{bmatrix} + \begin{bmatrix}4 & 0 & -2\\10 & 0 & -5\end{bmatrix} + \begin{bmatrix}9 & 3 & 6\\18 & 6 & 12\end{bmatrix}$$
$$= \begin{bmatrix}14 & 4 & 3\\32 & 10 & 3\end{bmatrix}$$

- кроз производе матрица · вектор

$$AB = \left[\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\3\end{bmatrix} \quad \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\0\\2\end{bmatrix} \quad \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} \cdot \begin{bmatrix}-1\\-1\\2\end{bmatrix}\right] = \begin{bmatrix}14 & 4 & 3\\32 & 10 & 3\end{bmatrix}$$

- кроз производе вектор · матрица

$$AB = \begin{bmatrix}\begin{bmatrix}1 & 2 & 3\end{bmatrix} \cdot \begin{bmatrix}1 & 1 & -1\\2 & 0 & -1\\3 & 1 & 2\end{bmatrix}\\ \\ \begin{bmatrix}4 & 5 & 6\end{bmatrix} \cdot \begin{bmatrix}1 & 1 & -1\\2 & 0 & -1\\3 & 1 & 2\end{bmatrix}\end{bmatrix} = \begin{bmatrix}14 & 4 & 3\\32 & 10 & 3\end{bmatrix}$$

Написани изрази су математички еквиваленти. Ипак, различите верзије множења матрица могу да имају различите перформансе када се имплементирају на рачунару. Кључ је у начинима приступа меморији приликом прикупљања операнада или упису резултата.

Пример 2.3.7. Операцијску сложеност производа две квадратне матрице можемо да одредимо на сличан начин као што је то урађено за операцију матрица· вектор у примеру 2.3.5. Множење две реалне квадратне матрице $A, B \in \mathcal{M}_{n\times n}$ представља израчунавање n^2 унутрашњих производа вектора дужине n. С обзиром да сваки од ових унутрашњих производа кошта n множења и $n-1$ сабирања бројева, то производ матрица AB троши n^3 операција множења и $n^2(n-1)$ операција сабирања. Укупна операцијска сложеност је због тога $2n^3 - n^2 = \mathcal{O}(2n^3)$. Ово је операција трећег нивоа са квадаратним обимом података и кубним обимом посла.

Разлика у операцијској сложености два производа

матрица · вектор $\mathcal{O}(2n^2)$ и матрица · матрица $\mathcal{O}(2n^3)$

користи се у већини алгоритама за повећање њихове ефикасности. Кад год је то могуће производи матрица се замењују производом матрице и вектора. Наредним примерима описани су карактеристични случајеви.

Пример 2.3.8. Наизменична примена унутрашњег и спољашњег производа вектора се интензивно кориисти за лакше сређивање и манипулацију матричним изразима, као и описивање и валидацију алгоритамских корака. Испитивањем операцијске сложености израза uv^Tw за нека три вектора $u, v, w \in \mathbb{R}^n$ скрећемо пажњу на веома важну особину матричног множења: асоцијативност.

Израз uv^Tw можемо да рачунамо на два начина: $(uv^T)w$ и $u(v^Tw) = (v^Tw)u$. Операцијска сложеност првог начина је $2n^3$:

- Израчунавање спољашњег производа вектора u и v је сложености n^2.
- Производ матрице $uv^T \in \mathcal{M}_{n\times n}$ и вектора $w \in \mathcal{M}_{n\times 1}$ је сложености $2n^3 - n^2$

Операцијска сложеност другог начина је $3n - 1$:

- Израчунавање унутрашњег производа вектора u и v је сложености $2n - 1$.
- Скалирање вектора $w \in \mathcal{M}_{n\times 1}$ скаларом $v^T w \in \mathcal{M}_{1\times 1}$ је сложености n.

Због велике разлике у операцијској сложености два поступка израчунавања истог израза, приликом имплементације алгоритама користи се облик

$$u \underbrace{(v^T w)}_{\text{скалар}} = (v^T w)u.$$

Тумачење производа Av преко линеарних комбинација колона матрице A има важну имплементацијску карактеристику.

Пример 2.3.9. Дата је матрица $A = [a_{ij}] \in \mathcal{M}_{m\times n}$. Потребно је представити матричним множењем извлачење садржаја прве колоне те матрице. Означимо са $e_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T$ први елемент канонске базе простора $\mathbb{R}^n$. Тада је

$$Ae_1 = \begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} & \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} & \dots & \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}.$$

Слично, уколико је потребан садржај $k-$те колоне матрице A, то се може назначити множењем матрице A одговарајућим вектором канонске базе.

$$Ae_k = \begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} & \dots & \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} & \dots & \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}.$$

За извлачење врста матрице A примењују се аналогне трансформације векторима канонске базе простора $\mathbb{R}^m$.

$$e_k^T A = \begin{bmatrix} 0 & \dots & 1 & \dots & 0 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{k1} & a_{k2} & \dots & a_{kn} \end{bmatrix}.$$

Закључујемо да испитивање садржаја матрице можемо да извршимо колону по колону или врсту по врсту, и да се одговарајући поступак представља матричним множењем. Тако нпр. за $A \in \mathcal{M}_{m\times n}$ важи

$$\begin{aligned} A = O \quad &\Longleftrightarrow \quad Ae_i = \theta,\ i = 1, 2, \dots, n \quad &&\Longleftrightarrow \quad e_i^T A = \theta^T,\ i = 1, 2, \dots, m, \\ &\Longleftrightarrow \quad Av = \theta,\ \forall v \in \mathbb{R}^n &&\Longleftrightarrow \quad u^T A = \theta^T,\ \forall u \in \mathbb{R}^m. \end{aligned} \tag{2.5}$$

где су O нула матрица, θ нула-вектор одговарајућег простора $\mathbb{R}^n$ или $\mathbb{R}^m$, и e_i вектори канонске базе одговарујећег простора $\mathbb{R}^n$ или $\mathbb{R}^m$.

Пример 2.3.10. Производи

$$D_nA = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \cdot A \qquad \text{и} \qquad AD_m = A \cdot \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$$

врше скалирање редом врста и колона матрице $A \in \mathcal{M}_{n\times m}$ коефицијентима λ_k. Шематски то можемо представити на следећи начин.

$$D_nA = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \end{bmatrix} \\ \begin{bmatrix} a_{21} & a_{22} & \ldots & a_{2m} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \lambda_1 \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \end{bmatrix} \\ \lambda_2 \begin{bmatrix} a_{21} & a_{22} & \ldots & a_{2m} \end{bmatrix} \\ \vdots \\ \lambda_n \begin{bmatrix} a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \end{bmatrix},$$

$$\begin{aligned} AD_m &= \begin{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} & \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} & \ldots & \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_m \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} & \lambda_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} & \ldots & \lambda_m \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{bmatrix} \end{bmatrix}. \end{aligned}$$

Претходни примери сугеришу закључке у наставку.

Последица 1. Свака елементарна трансформација матрице A, тј. трансформација једног од следећих типова:

а) замена места двеју врста (колона);

б) множење врсте (колоне) бројем c различитим од нуле;

в) додавање једној врсти (колони) друге врсте (колоне) помножене бројем c,

може да се добије множењем матрице A неком регуларном матрицом T слева за трансформације врста, односно здесна за трансформације колона.

Доказ: а) Нека је $A \in \mathcal{M}_{n\times m}$ и A' матрица која се добија заменом места $i-$тој и $j-$тој врсти матрице A. Означимо са I_{ij} матрицу коју добијамо из јединичне матрице $I_n \in \mathcal{M}_{n\times n}$ заменом места $i-$тој и $j-$тој врсти. Тада је $A' = I_{ij}A$.

За трансформације колона најпре трансформишемо јединичну матрицу $I_m \in \mathcal{M}_{m\times m}$ у матрицу I_{ij} заменом места $i-$тој и $j-$тој колони. Тада је AI_{ij} тражена трансформисана матрица A.

Тражена трансформациона матрица T за трансформације б) и в) добија се аналогно применом исте елементарне трансформације на јединичну матрицу одговарајуће димензије. □

Доказано тврђење, последица 1 сугерише да сви поступци Гаусовог алгоритма и одређивања ранга матрице имају своје представљање кроз множење матрица. Наредно тврђење показује да и екстракција подматрица из дате матрице такође има своје представљање операцијом множења матрица.

Последица 2. Нека је $A \in \mathcal{M}_{n\times m}$ и нека су дати скупови индекса врста $1 \le i_1 < i_2 < \cdots < i_p \le m$ и индекса колона $1 \le j_1 < j_2 < \cdots < j_q \le n$. Формирамо матрице

$$E = \begin{bmatrix} e_{i_1} & e_{i_2} & \dots & e_{i_p} \end{bmatrix} \quad \text{и} \quad F = \begin{bmatrix} e_{j_1} & e_{j_2} & \dots & e_{j_q} \end{bmatrix},$$

чије су колоне вектори канонских база:

$$e_{i_k} \in \mathbb{R}^n,\ k = 1, 2, \dots, p, \qquad e_{i_k} = [0\ \dots\ 0 \quad \underbrace{1}_{i_k\text{ - позиција}} \quad 0 \dots\ 0]^T,$$

$$e_{j_k} \in \mathbb{R}^m,\ k = 1, 2, \dots, q, \qquad e_{j_k} = [0\ \dots\ 0 \quad \underbrace{1}_{j_k\text{ - позиција}} \quad 0 \dots\ 0]^T.$$

Тада је $E^T AF$ подматрица матрице A која се налази у пресеку њених врста са индексима $\{i_1, \dots, i_p\}$ и колона са индексима $\{j_1, \dots, j_q\}$. Штавише, ако је $B \in \mathcal{M}_{p\times q}$, тада $A + EBF^T$ замењује подматрицу $E^T AF$ матрицом $E^T AF + B$.

Ова последица омогућава да о Гаусовом алгоритму размишљамо и кроз симултане трансформације врста, односно колона. У наставку бавимо се особинама ранга матрице.

Лема 2. За матрице $A \in \mathcal{M}_{m\times n}$ и $B \in \mathcal{M}_{n\times p}$ важи

$$\mathcal{R}(AB) \subseteq \mathcal{R}(A).$$

ДОКАЗ: Прикажимо матрице A и B преко њихових вектора-колона:

$$A = \left[\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_n \end{array}\right], \qquad B = \left[\begin{array}{c|c|c|c} b_1 & b_2 & \dots & b_p \end{array}\right].$$

Колоне резултујуће матрице производа AB добијају се као линеарне комбинације колона матрице A где су коефицијенти тих линеарних комбинација садржани у компонентама вектора $b_1, b_2, \dots, b_p$,

$$AB = \left[\begin{array}{c|c|c|c} Ab_1 & Ab_2 & \dots & Ab_p \end{array}\right].$$

Због тога важи:

$$\mathcal{R}(AB) = \mathcal{L}(Ab_1, Ab_2, \dots, Ab_p), \qquad Ab_1, Ab_2, \dots, Ab_p \in \mathcal{L}(a_1, a_2, \dots, a_n)$$
$$\overset{\text{Лема 1}}{\Longrightarrow} \quad \mathcal{L}(Ab_1, Ab_2, \dots, Ab_p) \subseteq \mathcal{L}(a_1, a_2, \dots, a_n)$$
$$\iff \quad \mathcal{R}(AB) \subseteq \mathcal{R}(A). \quad \square$$

Последица 3. За матрице $A \in \mathcal{M}_{m\times n}$ и $B \in \mathcal{M}_{n\times p}$ важи

$$\operatorname{rang}(AB) \le \operatorname{rang}(A),\ \operatorname{rang}(B),$$

тј. $\operatorname{rang}(AB) \le \min\{\operatorname{rang}(A),\ \operatorname{rang}(B)\}$.

ДОКАЗ: Неједнакост $\operatorname{rang}(AB) \leq \operatorname{rang}(A)$ је директна последица претходне теореме. Докажимо и другу неједнакост, тј. да је $\operatorname{rang}(AB) \leq \operatorname{rang}(B)$.

$$\operatorname{rang}(AB) = \operatorname{rang}(AB)^T = \operatorname{rang}(B^T A^T) \overset{\text{Лема } 2}{\leq} \operatorname{rang}(B^T) = \operatorname{rang}(B). \quad \square$$

Наредно тврђење је окосница већине алгоритама над матрицама, а даје повод за још један начин дефинисања ранга матрице.

Теорема 6. Матрица $A \in \mathcal{M}_{m\times n}$ је ранга r акко се може представити у облику производа

$$A = BC \qquad B \in \mathcal{M}_{m\times r},\ C \in \mathcal{M}_{r\times n},\ \operatorname{rang}(B) = \operatorname{rang}(C) = r, \tag{2.6}$$

ДОКАЗ: Приметимо прво да из услова $\operatorname{rang}(A) = r$ следи да је $r \leq m, n$. Тврђење теореме је облика ако и само ако и доказујемо га у два смера. Користићемо блок матрице за поједностављен запис.

$\boxed{\Longrightarrow :}$ Матрица A ранга r има r линеарно независних колона. Променом редоследа колона унутар матрице A, тј. множењем задесна пермутационом матрицом P (последица 1) полазну матрицу A можемо довести на блок облик

$$AP = \left[\ B_{m\times r} \mid C_{m\times(n-r)}\ \right],$$

где су линеарно независне колоне у $B_{m\times r}$, тј.

$$\mathcal{R}(A) = \mathcal{R}\left(B_{m\times r}\right), \qquad \operatorname{rang}(A) = \operatorname{rang}(B_{m\times r}) = r.$$

Другим речима, ове колоне образују базу у односу на коју се колоне матрице $C_{m\times(n-r)}$ могу представити неким линеарним комбинацијама. Дакле постоји матрица $T_{r\times(n-r)} \in \mathcal{M}_{r\times(n-r)}$ за коју је

$$C_{m\times(n-r)} = B_{m\times r}T_{r\times(n-r)}.$$

Тада је

$$\begin{aligned} AP &= \left[\ B_{m\times r} \mid C_{m\times(n-r)}\ \right] = \left[\ B_{m\times r}I_r \mid B_{m\times r}T_{r\times(n-r)}\ \right] \\ &= B_{m\times r}\left[\ I_r \mid T_{r\times(n-r)}\ \right] \end{aligned}$$

Како је $\operatorname{rang}(B_{m\times r}) = r$, и $\operatorname{rang}\left(\left[\ I_r \mid T_{r\times(n-r)}\ \right]\right) = r$, то је

$$A = B_{m\times r}\left(\left[\ I_r \mid T_{r\times(n-r)}\ \right]P^{-1}\right) = B_{m\times r}C_{r\times n}, \qquad C_{r\times n} = \left[\ I_r \mid T_{r\times(n-r)}\ \right]P^{-1}.$$

Тиме је тврђење теореме 6 у директном смеру доказано.

$\boxed{\Longleftarrow :}$ Претпоставимо сада да је $A = BC$, где је $B \in \mathcal{M}_{m\times r}$ матрица пуног ранга колона и $C \in \mathcal{M}_{r\times n}$ матрица пуног ранга врста r. Како је $\operatorname{rang}(B) = r$, то матрица B има r линеарно независних врста. Избором адекватне пермутационе матрице $P \in \mathcal{M}_{m\times m}$ матрица B може да се доведе на блок облик

$$PB = \left[\begin{array}{c} B_{r\times r} \\ \hline D_{(m-r)\times r} \end{array}\right],$$

где је $B_{r\times r}$ регуларна матрица. Слично, пермутационом матрицом $Q \in \mathcal{M}_{n\times n}$ матрица C може да се доведе на блок облик

$$CQ = \left[\ C_{r\times r} \mid E_{r\times(n-r)}\ \right],$$

са регуларним блоком $C_{r\times r}$. С обзиром да пермутационе матрице само мењају места векторима оне не утичу на њихову линеарну независност, то је $\operatorname{rang}(A)=\operatorname{rang}(PAQ)$. Због тога је

$$
\begin{aligned}
PAQ=(PB)(CQ)&=\left[\begin{array}{c} B_{r\times r}\\ \hline D_{(m-r)\times r}\end{array}\right]\left[\begin{array}{c|c} C_{r\times r} & E_{r\times(n-r)}\end{array}\right]\\
&=\left[\begin{array}{c|c} B_{r\times r}C_{r\times r} & B_{r\times r}E_{r\times(n-r)}\\ \hline D_{(m-r)\times r}C_{r\times r} & D_{(m-r)\times r}E_{r\times(n-r)}\end{array}\right] \qquad (2.7)
\end{aligned}
$$

са регуларним блоком $B_{r\times r}C_{r\times r}$, закључујемо да је $\operatorname{rang}(A)\geq r$. На основу последице 3 је $\operatorname{rang}(A)\leq \operatorname{rang}(B),\operatorname{rang}(C)$, па добијамо коначно да је $\operatorname{rang}(A)=r$. $\square$

Разлагање матрице (2.6) назива се факторизација пуног ранга матрице A.

Последица 4. Матрица $A\in\mathcal{M}_{m\times n}$ је ранга 1 акко је једнака спољашњем производу нека два вектора $v\in\mathbb{R}^m$ и $u\in\mathbb{R}^n$,

$$A=v\,u^T.$$

Пример 2.3.11. Нека је позната декомпозиција пуног ранга матрице $A=BC$, где су $A\in\mathcal{M}_{n\times n}$, $B\in\mathcal{M}_{n\times r}$ и $C\in\mathcal{M}_{r\times n}$. Уведимо ознаке колона матрице B и врста матрице C :

$$
B=\left[\begin{array}{c|c|c|c} b_1 & b_2 & \dots & b_r\end{array}\right],\qquad C=\left[\begin{array}{c} c_1^T\\ \hline c_2^T\\ \hline \vdots\\ \hline c_r^T\end{array}\right].
$$

На основу правила за спољашње производе вектора, матрица A гласи:

$$A=BC=\sum_{i=1}^{r} b_i c_i^T, \qquad (2.8)$$

што је збир r матрица ранга 1. За израчунавање производа Av, где је $v\in\mathbb{R}^n$ произвољан вектор, у примеру 2.3.7 показали смо да је потребно $2n^2-n$ аритметичких операција. Коришћењем израза (2.8) за матрицу A можемо да смањимо операцијску сложеност производа Av :

$$Av=\left(\sum_{i=1}^{r} b_i c_i^T\right)v=\sum_{i=1}^{r} b_i(c_i^T v).$$

Зајључујемо да овакав поступак користи:

- r скаларних производа $n-$димензионалних вектора, што је $r(2n-1)$ операција,
- r скалирања $n-$димензионалних вектора, што је rn операција.

Коначно, то је укупно $3rn-1=\mathcal{O}(3rn)$ операција.

Последица 4 и једнакост (2.8) су повод за наредну теорему.

Теорема 7. Матрица $A\in\mathcal{M}_{m\times n}$ је ранга r акко је r најмањи број матрица $M_i\in\mathcal{M}_{m\times n}$ ранга 1 помоћу којих се A може раставити у збир,

$$A=\sum_{i=1}^{r} M_i.$$

Веома важна карактеристика декомпозиције пуног ранга (2.6) јесте очување фундаменталних потпростора матрице A.

Декомпозиција пуног ранга $A = BC$ тада даје

$$\mathcal{R}(A) \subseteq \mathcal{R}(B) \quad \text{и} \quad \operatorname{rang}(A) = \operatorname{rang}(B)$$

следи $\mathcal{R}(A) = \mathcal{R}(B)$. Познавање декомпозиције пуног ранга матрице A значи и познавање декомпозиције пуног ранга матрице A^T :

$$A = BC \quad \Longrightarrow \quad A^T = C^T B^T.$$

Због тога је

$$\mathcal{R}(A) = \mathcal{R}(B), \qquad \mathcal{N}(A) = \mathcal{N}(C),$$
$$\mathcal{R}(A^T) = \mathcal{R}(C^T), \qquad \mathcal{N}(A^T) = \mathcal{N}(B^T).$$

Увођењем ознаке за ненегативне и позитивне матрице можемо да дефинишемо и релације $\geq$ и $>$ на скупу матрица истог типа $\mathcal{M}_{m\times n}$.

Дефиниција 8. За матрице $A, B \in \mathcal{M}_{m\times n}$ је

$$\begin{aligned} A \geq B \quad &\Longleftrightarrow \quad B \leq A \quad \Longleftrightarrow \quad A - B \geq O,\\ A > B \quad &\Longleftrightarrow \quad B < A \quad \Longleftrightarrow \quad A - B > O. \end{aligned}$$

2.4 Полиноми

Скуп полинома степена не већег од n означавамо са $\mathbb{P}_n[t]$. Операције на овом скупу описују се обично кроз коефицијенте у мономима. За $P, Q \in \mathbb{P}_n[t]$ и $\alpha \in \mathbb{R}$

$$P(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0, \quad Q(t) = b_n t^n + b_{n-1}t^{n-1} + \cdots + b_1 t + b_0,$$

$$(P+Q)(t) = (a_n + b_n)t^n + (a_{n-1} + b_{n-1})t^{n-1} + \cdots + (a_1 + b_1)t + (a_0 + b_0),$$

$$(\alpha P)(t) = (\alpha a_n)t^n + (\alpha a_{n-1})t^{n-1} + \cdots + (\alpha a_1)t + (\alpha a_0).$$

Нула-вектор овог векторског простора зовемо нула-полином,

$$O(t) = 0, \forall t, \quad O(t) = 0 \cdot t^n + 0 \cdot t^{n-1} + \cdots + 0 \cdot t + 0.$$

За векторски простор полинома у примени је неколико веома корисних база. Једна од њих је база монома

$$(B)_n \ : \ \left\{t^n, t^{n-1}, \ldots, t^2, t, 1\right\}, \qquad \dim \mathbb{P}_n[t] = n + 1.$$

Координате полинома $P(t)$ у односу на мономску базу су управо коефицијенти тог полинома a_i. Због тога полиноме поистовећујемо са овим вектором коефицијената,

$$P(t) = a_n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \equiv \begin{bmatrix} a_n & a_{n-1} & \ldots & a_1 & a_0 \end{bmatrix}^T.$$

Такође, и са избором неке друге базе полиноме поистовећујемо са одговарајућим вектором коефицијената. Нпр. Тејлорова база

$$(T)_n \ : \ \left\{\frac{(t-x_0)^n}{n!}, \frac{(t-x_0)^{n-1}}{(n-1)!}, \ldots, \frac{t-x_0}{1!}, 1\right\},$$

за кординате има вредности извода полинома у тачки x_0,

$$P(t) \equiv \begin{bmatrix} P^{(n)}(x_0) & P^{(n-1)}(x_0) & \dots & P'(x_0) & P(x_0) \end{bmatrix}^T.$$

Осим сагледавања полинома кроз координате, веома је корисно и њихово разматрање као полиномских функција, тј. кроз скупове вредности које узимају у неким битним тачкама домена. У томе кључну улогу има последица основног става алгебре.

Теорема 8. (Основни став алгебре) Полином $P_n(t) \in \mathbb{P}_n[t]$ степена $n \geq 0$ има тачно n нула у пољу комплексних бројева.

Последица 5. Уколико се два полинома $P(t), Q(t) \in \mathbb{P}_n[t]$ поклапају у не мање од $n+1$ различитих тачака онда су они једнаки.

$$P(x_k) = Q(x_k),\ k = 0, 1, \dots, n \qquad \Longrightarrow \qquad P(t) = Q(t),\ \forall t.$$

ДОКАЗ: Посматрамо полиом $S(t) = P(t) - Q(t)$. Како је S разлика два полинома степена не већег од n, то је и S полином степена не већег од n. Према основном ставу алгебре полином S има највише n нула. Међутим, S има бар $n+1$ различиту нулу

$$S(x_k) = P(x_k) - Q(x_k) = 0, \quad k = 0, 1, \dots, n.$$

Закључујемо да он мора бити нула-полином.□

Последица ових разматрања јесте да је полином степена не већег од n једнозначно одређен познавањем његових вредности у бар $n+1$ различитих тачака,

$$P(t) \equiv \begin{bmatrix} P(x_0) & P(x_1) & \dots & P(x_n) \end{bmatrix}^T. \tag{2.9}$$

Приметимо да и овакав начин репрезентације полинома одговара операцијама векторског простора,

$$\begin{aligned} (P+Q)(t) &\equiv \begin{bmatrix} P(x_0)+Q(x_0) & P(x_1)+Q(x_1) & \dots & P(x_n)+Q(x_n) \end{bmatrix}^T, \\ (\alpha P)(t) &\equiv \begin{bmatrix} \alpha P(x_0) & \alpha P(x_1) & \dots & \alpha P(x_n) \end{bmatrix}^T. \end{aligned}$$

Веза између две репрезентације полинома, вектора коефицијената и вектора вредности, описује се матричним операцијама.

$$P(x_k) = a_n x_k^n + a_{n-1} x_k^{n-1} + \dots + a_1 x_k + a_0, \quad k = 0, 1, \dots, n$$

$$\Longleftrightarrow \begin{bmatrix} x_0^n & x_0^{n-1} & \dots & x_0 & 1 \\ x_1^n & x_1^{n-1} & \dots & x_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_n^n & x_n^{n-1} & \dots & x_n & 1 \end{bmatrix} \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{bmatrix} = \begin{bmatrix} P(x_0) \\ P(x_1) \\ \vdots \\ P(x_n) \end{bmatrix}.$$

Израчунаввање вредности $P(x_k)$ је узорковање полиномске функције на скупу чворова $x_k,\ k = 0, 1, \dots, n$. Одређивање коефицијената полинома на основу вектора вредности зовемо интерполација,

$$\begin{bmatrix} x_0^n & x_0^{n-1} & \dots & x_0 & 1 \\ x_1^n & x_1^{n-1} & \dots & x_1 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_n^n & x_n^{n-1} & \dots & x_n & 1 \end{bmatrix}^{-1} \begin{bmatrix} P(x_0) \\ P(x_1) \\ \vdots \\ P(x_n) \end{bmatrix} = \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_0 \end{bmatrix}.$$

Различитим избором базе простора полинома можемо операцијски да олакшамо поступак интерполације. Једна таква база је Лагранжова база, чије елементе још називамо и кардинални полиноми,

$$(L)_n \ : \ \left\{ L_k(t) = \frac{\prod\limits_{j\neq k}(t-x_j)}{\prod\limits_{j\neq k}(x_k-x_j)} \ \middle| \ k=0,1,\dots,n \right\}.$$

Лагранжова база $(L)_n$ састоји се од полинома степена не већег од n, тј. $\mathrm{st}(L_k) \leq n,\ k=0,1,\dots,n$, и има својство да за произвољан полином $P \in \mathbb{P}_n[t]$ његове вредности у чворовима x_k су управо његове координате,

$$P(t) = P(x_0)L_0(t) + P(x_1)L_1(t) + \cdots + P(x_n)L_n(t).$$

Другим речима (2.9) је представљање полинома $P(t)$ координатама у Лагранжовој бази. То је последица својства $L_k(x_j) = \begin{cases} 1, & k=j \\ 0, & k\neq j. \end{cases}$ У матричном облику ови услови гласе $\big[L_k(x_j)\big]_{k,j=\overline{0,n}} = I$, тј.

$$\begin{bmatrix} L_0(x_0) & L_1(x_0) & \dots & L_n(x_0) \\ L_0(x_1) & L_1(x_1) & \dots & L_n(x_1) \\ \vdots & \vdots & & \vdots \\ L_0(x_n) & L_1(x_n) & \dots & L_n(x_n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Дакле за полином $L_k(t)$ степена не већег од n познајемо тачно n његових нула $x_j,\ j\neq k$. Због тога је $L_k(t) = \alpha_k w_k(t)$, где је $w_k(t) = \prod\limits_{i\neq k}(t-x_i)$ монични полином са уоченим нулама. Водећи коефицијент α_k добијамо из услова $L_k(x_k) = 1:$

$$1 = L_k(x_k) = \alpha_k w_k(x_k) \quad \Longrightarrow \quad \alpha_k = \frac{1}{w_k(x_k)} = \frac{1}{\prod\limits_{i\neq k}(x_k-x_i)}.$$

На тај начин долазимо до израза за елементе Лагранжове базе.

Простор $\mathbb{P}_n[t]$ полинома степена не већег од n је коначно димензионалан. Осим њега можемо да посматрамо и простор полинома $\mathbb{P}[t]$, тј. без ограничења степена полинома. Овај векторски простор је бесконачно димензионалан. Хамелову базу чине мономи

$$(B) \ : \ \{1, t, t^2, \dots, t^n, \dots\}.$$

Било који полином је линеарна комбинација коначно много монома.

2.5 Функције

Функције једне или више промењљивих $\mathcal{F}[t],\ \mathcal{F}[t_1,t_2,\dots,t_n] \equiv \mathcal{F}[v]$ представљају елементе векторског простора захваљујући операцијама:

$$f, g \in \mathcal{F}[v],$$

$$(f+g)(v) = f(v) + g(v), \qquad (\alpha f)(v) = \alpha f(v).$$

Нула-вектор је нула-функција $O(v) = 0,\ \forall v$.

Пример 2.5.1. Испитивање линеарне независности коначног скупа диференцијабилних функција $f_0(t), f_1(t), \ldots, f_n(t)$ једне реалне променљиве t може да се олакша употребом извода. Посматрамо једнакост

$$\alpha_0 f_0(t) + \alpha_1 f_1(t) + \cdots + \alpha_n f_n(t) = 0, \quad \forall t.$$

Диференцирањем ове једнакости n пута добијамо систем једначина

$$\begin{cases} \alpha_0 f_0(t) + \alpha_1 f_1(t) + \cdots + \alpha_n f_n(t) = 0, \\ \alpha_0 f_0'(t) + \alpha_1 f_1'(t) + \cdots + \alpha_n f_n'(t) = 0, \\ \alpha_0 f_0''(t) + \alpha_1 f_1''(t) + \cdots + \alpha_n f_n''(t) = 0, & \qquad \forall t. \\ \vdots \\ \alpha_0 f_0^{(n)}(t) + \alpha_1 f_1^{(n)}(t) + \cdots + \alpha_n f_n^{(n)}(t) = 0, \end{cases}$$

Заменом неке конкретне вредности $t = t_0$ у горњи систем он постаје хомоген систем линеарних једначина по непознатим коефицијентима $\alpha_0, \alpha_1, \ldots, \alpha_n$. Решење хомогеног система ће бити јединствено ако је детерминанта система различита од нуле. Због тога, следећа детерминаната, која носи име Вронскијан, служи приликом испитивања линеарне независности функција,

$$W(f_0, f_1, \ldots, f_n) = \begin{vmatrix} f_0(t) & f_1(t) & \ldots & f_n(t) \\ f_0'(t) & f_1'(t) & \ldots & f_n'(t) \\ \vdots & \vdots & & \vdots \\ f_0^{(n)}(t) & f_1^{(n)}(t) & \ldots & f_n^{(n)}(t) \end{vmatrix}.$$

База векторског простора функција, у контексту Хамелове базе (дефиниција 2) је "превелика", има непребројиво много елемената. Због тога обично не посматрамо баш све могуће функције него неку битну класу функција, рецимо непрекидне, или довољан број пута диференцијабилне функције, или функције посебног облика. Ипак својство представљања преко коначних линеарних комбинација базе ту постаје нарушено. Линеарне комбинације у таквој бази могу да садрже пребројиво много елемената - представљају функционалне редове,

$$f(t) = \sum_{k=0}^{n} \alpha_k b_k(t).$$

Пример 2.5.2. Нека је f функција диференцијабилна произвољан број пута у околини неке тачке x_0. Тада је за вредности t довољно близу x_0 могуће разлагање у Тејлоров ред

$$f(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(t - x_0)^k.$$

За $t_0 = 0$ Тејлоров развој зовемо Маклоренов развој. Неки корисни Маклоренови развоји и области конвергенције су:

$$\frac{1}{1-t} = 1 + t + t^2 + \cdots + t^n + \ldots, \qquad |t| < 1,$$

$$e^t = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!} + \ldots \qquad t \in \mathbb{R}.$$

У случају појединих класа функција Хамелова база и даље остаје опширна. Смањење броја елемената базе могуће је помоћу појма ортогоналности, па је ортогоналност наша наредна тема. Ортогоналне базе векторских простора имају највише пребројиво много елемената. Својство представљања преко коначних линеарних комбинација и даље није испуњено. Најпознатије такве базе су ортогонални полиноми и Фуријеов ред.

Приметимо да као и у случају полиномских функција, и друге функције могу да се анализирају кроз коначне скупове вредности $f(x_k)$, $k = 0, 1, \ldots, n$. Проблем настаје када је домен функције шири од скупа чворова. Тада теорема о једнозначности представљања (последица 5) више не мора важи. Дакле, реконструкција непознате функције (која није полином) на основу коначног скупа података у општем случају није једнозначно решив проблем.

Објекти рада у свим истакнутим векторским просторима очигледно нису геометријски вектори или вектори из физике. Имају нека своја посебна својства која користимо у разне сврхе, у контексту решења различитих проблема. Па ипак, разматрање ових објеката кроз концепт векторских простора и интуиција из геометријских вектора буду корисни приликом решавања проблема. Сваки од наведених кључних векторских простора поседује и неке додатне операције које су његова особеност. Тако нпр. матрице имају операције транспоновања и множења, функције имају композицију, изводе и интеграле, итд. Саме матрице можемо посматрати као функције које делују између векторских простора. Нпр. за матрицу $A \in \mathcal{M}_{m\times n}$ можемо да разматрамо следећа пресликавања

$$A : \mathbb{R}^n \mapsto \mathbb{R}^m, \qquad v \mapsto Av, \qquad A(\alpha v + \beta u) = \alpha Av + \beta Au, \tag{2.10}$$

$$A : \mathbb{R}^m \mapsto \mathbb{R}^n, \qquad v \mapsto v^T A, \qquad (\alpha v + \beta u)^T A = \alpha v^T A + \beta u^T A, \tag{2.11}$$

$$\begin{aligned} A : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}, \quad & (u, v) \mapsto v^T Au, \quad && \begin{cases} (\alpha v + \beta w)^T Au = \alpha v^T Au + \beta w^T Au, \\ v^T A(\alpha u + \beta w) = \alpha v^T Au + \beta v^T Aw \end{cases} \\ & (u, v) \mapsto A(u, v) && \begin{cases} A(\alpha u + \beta w, v) = \alpha A(u, v) + \beta A(w, v), \\ A(u, \alpha v + \beta w) = \alpha A(u, v) + \beta A(u, w). \end{cases} \end{aligned} \tag{2.12}$$

Подсетимо се да за матрицу $A \in \mathcal{M}_{m\times n}$,

$$A = [a_{ij}] = \left[\begin{array}{c} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_m^T \end{array}\right] = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \ldots & u_n \end{array}\right],$$

и векторе канонске базе $e_i \in \mathbb{R}^m$, $e_j \in \mathbb{R}^n$, важи

$$e_i^T A = v_i^T, \quad Ae_j = u_j, \quad e_i^T Ae_j = a_{ij}.$$

Специјално, када је $A \in \mathcal{M}_{1\times n},\ B \in \mathcal{M}_{n\times 1}$,

$$A, B : \mathbb{R}^n \mapsto \mathbb{R}, \qquad v \mapsto Av, \qquad v \mapsto v^T B$$

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

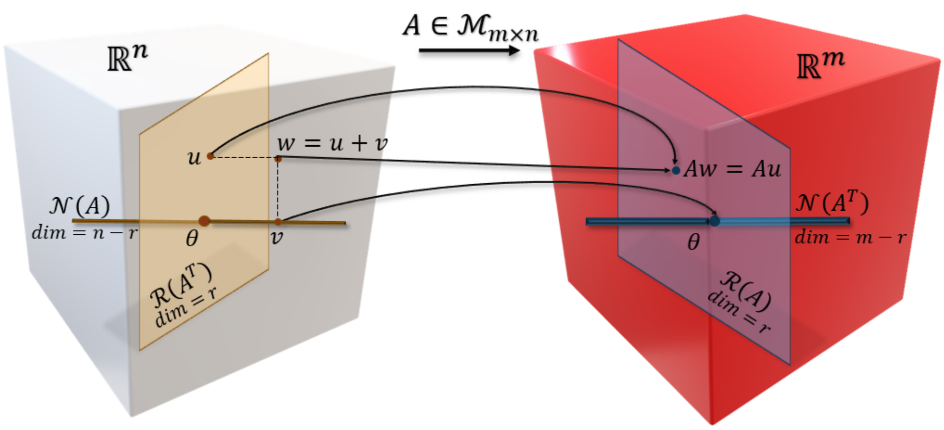
$$Ae_j = a_j, \qquad e_i^T B = b_i.$$

Матрицу $A \in \mathcal{M}_{m\times n}$ можемо да посматрамо кроз пресликавање једног вектора $v \in \mathbb{R}^n$ или кроз пресликавање два вектора $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$. Пресликавања са својством (2.10), односно (2.11), су линеарни оператори.

Дефиниција 9. Нека је $(b) :\ \{b_1, b_2, \dots, b_m\}$ база $m-$димензионалног векторског простора U, и $(s) :\ \{s_1, s_2, \dots, s_n\}$ база $n-$димензионалног векторског простора V. Затим, нека је $\mathcal{A} : U \to V$, линеаран оператор. Координате слика $\mathcal{A}(b_k),\ k = 1, 2, \dots, m$, изражене у односу на базу (s) смештене у колоне матрице формирају матрицу $A \in \mathcal{M}_{n\times m}$. Матрица A назива се матрица линеарног оператора $\mathcal{A}$ у односу на базе (b) и (s).

Теорема 9. Матрица изоморфизма је регуларна матрица.

Фундаментална теорема алгебре 5 за веома важну последицу има да свака матрица поседује потпросторе кроз које делује као бијекција, тачније на којима њена рестрикција поседује инверзно линеарно пресликавање - то су $\mathcal{R}(A)$ и $\mathcal{R}(A^T)$, слика 2.3. Ово је у вези са темом која све више добија на значају у модерним рачунарским технологијама: псеудо инверзи.



Слика 2.3: Фундаментални потпростори матрице

Пресликавања која имају особину (2.12) зовемо билинеарне функционеле - линеарна су и по једном и по другом аргументу. Билинеарно пресликавање није линеарно по вектору $\begin{bmatrix} u & v \end{bmatrix} \in \mathbb{R}^{m+n}$, већ по његовим деловима u и v. У складу са термином функције, изрази (2.10) и (2.12) могу да се анализирају кроз класична својства и алате: нуле функција, средња вредност, екстреми, изводи или интеграли. Шта висхе, на основу познавања правила за изводе и интеграле функција видимо да и ове операције поседују својства линеарности и билинеарности

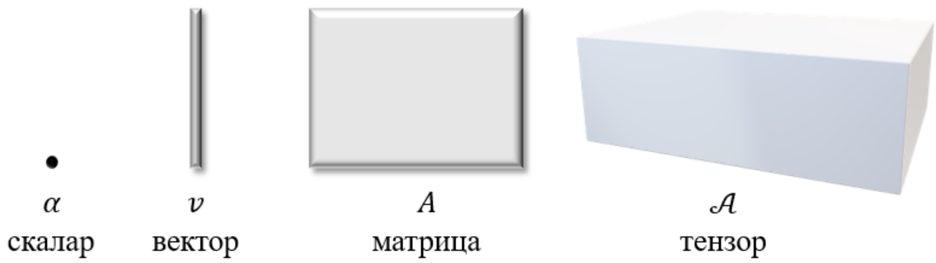
$$D(\alpha f+\beta g) = \big(\alpha f(t)+\beta g(t)\big)' = \alpha f'(t)+\beta g'(t) = \alpha\, D(f)+\beta\, D(g),$$

$$\begin{aligned} Int(\alpha f+\beta g) &= \int_a^b \big(\alpha f(t)+\beta g(t)\big)dt = \alpha\int_a^b f(t)\,dt+\beta\int_a^b g(t)\,dt\\ &= \alpha\, Int(f)+\beta\, Int(g),\\ Int(\alpha f+\beta g, h) &= \int_a^b \big(\alpha f(t)+\beta g(t)\big)h(t)\,dt = \alpha\int_a^b f(t)h(t)\,dt+\beta\int_a^b g(t)h(t)\,dt\\ &= \alpha\, Int(f,h)+\beta\, Int(g,h),\\ Int(h,\alpha f+\beta g) &= \alpha\, Int(h,f)+\beta\, Int(h,g). \end{aligned}$$

Приказани примери указују да се алати алгебре и анализе међусобно прожимају. Желимо да научимо како да ови алати сарађују у циљу решавања сложених проблема машинског учења.

2.6 Појам тензора

Због природе објеката који се појављују у проблемима машинског учења уводимо још један нов математички објекат: тензор. Интуитивно, тензоре сматрамо вишедимензионалним матрицама, слика 2.4.



Слика 2.4: Графичка илустрација скалара, вектора, матрице и тензора

Градација у димензионалности објеката са слике 2.4 описује се кроз индексирање елемената ових објеката:

- Скалар је у потпуности описан својом вредношћу и није нам потребан ниједан индекс за приступ вредности коју носи. Због тога за скалар кажемо да је објекат нула-димензионалности.

- Елементе вектора $v \in \mathbb{R}^n$ можемо да пратимо кроз један индекс, тј.

$$v : \{1,2,\dots,n\} \mapsto \mathbb{R}, \qquad v(i) = x_i,$$
$$v = \big(x_1,\ x_2,\ \dots,\ x_n\big) \quad\Longleftrightarrow\quad v : \begin{pmatrix} 1 & 2 & \dots & n\\ x_1 & x_2 & \dots & x_n \end{pmatrix}.$$

  На основу броја индекса потребних за опис "позиција" вредности које носи, векторе називамо једнодимензионалним тензорима. Тако за вектор $v \in \mathbb{R}^n$ кажемо да је једнодимензионални тензор типа n.

- Елементе матрице $A \in \mathcal{M}_{m\times n}$ пратимо распоредом помоћу два индекса,

$$A : \{1,2,\dots,m\} \times \{1,2,\dots,n\} \mapsto \mathbb{R}, \quad A(i,j) = a_{ij},$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \iff A : \begin{array}{c|cccc} & 1 & 2 & \dots & n \\ \hline 1 & a_{11} & a_{12} & \dots & a_{1n} \\ 2 & a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ m & a_{m1} & a_{m2} & \dots & a_{mn} \end{array}.$$

  Слично векторима, матрица $A \in \mathcal{M}_{m\times n}$ је дводимензионални тензор типа $m\times n$.

- Елементе тензора $\mathcal{A} \in \mathcal{M}_{m_1\times m_2\times\dots\times m_n}$ разликујемо кроз вредности n индекса,

$$\mathcal{A} : \{1,2,\dots,m_1\} \times \{1,2,\dots,m_2\} \times \dots \times \{1,2,\dots,m_n\} \mapsto \mathbb{R},$$
$$\mathcal{A}(i_1,i_2,\dots,i_n) = a_{i_1,i_2,\dots,i_n}.$$

  Тензор $\mathcal{A}$ је $n-$димензионалан, типа $m_1 \times m_2 \times \dots \times m_n$.

У наставку, због краћег записа, користићемо ознаку за скуп вредности индекса

$$[m] = \{1,2,\dots,m\}.$$

Осим колекције елемената који су индексирани са више индекса, тензори се еквивалентно дефинишу и кроз изразе сличне (2.12),

$$\begin{aligned} \mathcal{A} : \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \dots \times \mathbb{R}^{m_n} \mapsto \mathbb{R}, \qquad v_j, u_j \in \mathbb{R}^{m_j}, \text{ и } \forall i \in \{1,2,\dots,n\}, \\ \mathcal{A}(v_1,\dots,v_{i-1},\alpha v_i + \beta u_i, v_{i+1},\dots,v_n) = \alpha\mathcal{A}(v_1,\dots,v_{i-1},v_i,v_{i+1},\dots,v_n) \\ + \beta\mathcal{A}(v_1,\dots,v_{i-1},u_i,v_{i+1},\dots,v_n). \end{aligned}$$

Тако тензоре зовемо $n-$линеарна пресликавања, тј. ради се о пресликавању које је линеарно по било ком од n улазних аргумената. Аргументи (улази у функцију) овако дефинисаног тензора су вектори. Веза између вишедимензионалне матрице и $n-$линеарног пресликавања добија се помоћу вектора стандардне базе

$$a_{i_1,i_2,\dots,i_n} = \mathcal{A}(e_{i_1}, e_{i_2},\dots,e_{i_n}), \quad e_{i_k} \in \mathbb{R}^{m_k},\ k \in [n].$$

Напоменимо да тензор није линеарно пресликавање свог аргумента, тј. вектора $v \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \dots \times \mathbb{R}^{m_n} = \mathbb{R}^{m_1+m_2+\dots+m_n}$, већ његових појединачних делова v_i,

$$v = \begin{bmatrix} v_1^T & | & v_2^T & | & \dots & | & v_n^T \end{bmatrix}^T.$$

Претходна анализа обједињена је наредном дефиницијом.

Дефиниција 10. Тензор $\mathcal{A}$ димензије n, типа $m_1 \times m_2 \times \dots \times m_n$ је вишедимензионална табела бројева чији се појединачни елементи идентификују уређеном $n-$торком бројева $(i_1,i_2,\dots,i_n)$, $i_k \in [m_k]$, $k \in [n]$. Због тога, формално,

$$\mathcal{A} \ : \ [m_1] \times [m_2] \times \dots \times [m_n] \mapsto \mathbb{R}.$$

Такође, тензор је $n-$линеарно пресликавање

$$\mathcal{A} : \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \dots \times \mathbb{R}^{m_n} \mapsto \mathbb{R}.$$

Скуп тензора димензије n, типа $m_1 \times m_2 \times \dots \times m_n$ означаваћемо

$$\mathcal{M}_{m_1\times m_2\times\dots\times m_n} = \mathbb{R}^{m_1} \otimes \mathbb{R}^{m_2} \otimes \dots \otimes \mathbb{R}^{m_n} \cong \mathbb{R}^{m_1 m_2 \cdots m_n}.$$

Позиције индекса у именовању елемената тензора називамо модови или осе. Тако нпр. за

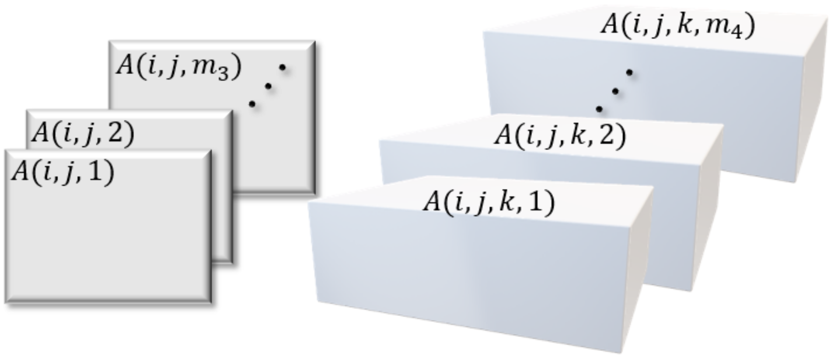
$$\mathcal{A} = [a_{i_1 i_2 \dots i_n}] \in \mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}$$

позиција индекса $i_1 \in [m_1]$ припада првом моду, $i_2 \in [m_2]$ другом, итд. Позиција индекса $i_n \in [m_n]$ носи вредност $n-$тог мода.

Напомена 1. $n-$линеарна пресликавања могу да се дефинишу у општем случају над векторским просторима $V_1, \dots, V_n$. Тада је одговарајућа вишедимензионална табела бројева одређена избором база у сваком од простора $V_1, \dots, V_n$. Због поистовећивања вектора са његовим координатама горња дефиниција је довољна за коначно димензионалне векторске просторе.

Пример 2.6.1. Шематско приказивање тензора по принципу кутија вредности радићемо према следећем договору[1]. Тензори димензија 1 и 2, тј, вектори и матрице уобичајено користе распоред елемената. То значи да је први индекс резервисан за врсте, а други за колоне. У Python нотацији то значи да је први индекс, индекс најспорије промене вредности у поступку обилажења елемената.

За тензоре димензије 3 ређамо табеле вредности према последњем индексу једну иза друге, слика 2.5. Такав приступ у приказу примењујемо и на тензоре већих димензија.



Слика 2.5: Конвенција графичког приказа вредности на примеру $3D$ и $4D$ тензора

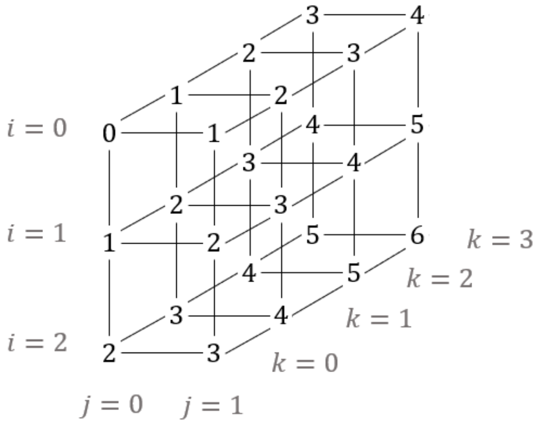
За конкретан пример посматраћемо тензор $\mathcal{A}$ димезије 3, типа $3 \times 2 \times 4$, чији су елементи описани формулом

$$a_{ijk} = i + j + k, \quad i = \overline{0,2},\ j = \overline{0,1},\ k = \overline{0,3}.$$

Дакле, користимо поново начин Python индексирања. Графички приказ овог тензора дат је на слици 2.6.

Напомена 2. Као што векторе (једнодимензионалне тензоре) можемо да сматрамо специјалним танким матрицама (дводимензионалним тензорима), тако векторе и матрице можемо да сматрамо специјалним танким тензорима виших димензија. Другим речима:

[1] Опште прихваћена конвенција око графичког представљања вредности тензора унутар кутије не постоји.



Слика 2.6: Пример тензора $3 \times 2 \times 4$,

- Вектор $v \in \mathbb{R}^n$ је тензор димензије 1, типа n, али је и тензор димензије 2 типа $n \times 1$ или $1 \times n$. Другим речима, уређена $n-$торка бројева

  $$v = \big(x_1,\ x_2,\ \ldots,\ x_n\big)$$

  као ознака је једнодимензионални тензор и одговара типу података једнодимензионалног Python NumPy низа. Вектор колона или вектор врста је тај исти објекат записан као дводимензионални тензор одређеног типа,

  $$v \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathcal{M}_{n\times 1}, \qquad v \equiv \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \in \mathcal{M}_{1\times n}.$$

  Настављајући овако, вектор можемо да посматрамо и као тензор димензије 3, а неког од типова $n \times 1 \times 1,\ 1 \times n \times 1, 1 \times 1 \times n$, итд.

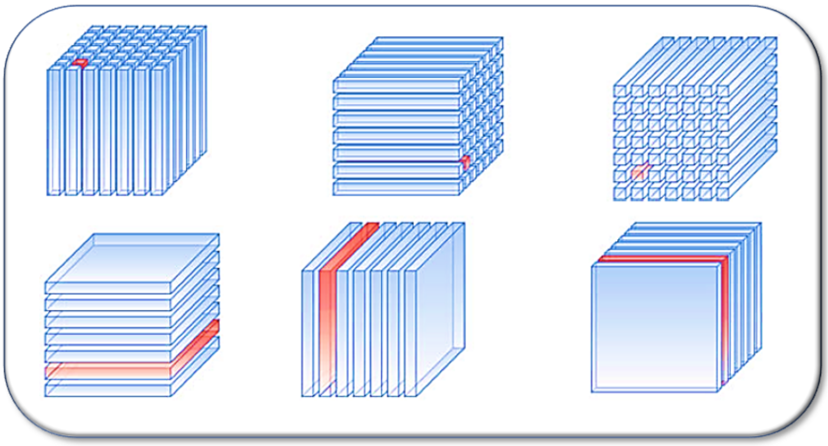
- Матрица $A \in \mathcal{M}_{m\times n}$ је тензор димензије 2 типа $m \times n$. Осим тога A је тензор димензије 3, рецимо типа $1 \times m \times n$.

Овакав начин сагледавања матрица и вектора омогућава њихово препознавање као саставних делова тензора виших димензија, слика 2.7. Ипак, мод величине 1 ($m_k = 1$) из практичних разлога никада не пишемо. Другим речима, у димензије тензора $\mathcal{M}_{m_1\times\cdots\times m_n}$ пишемо само модове са опсегом вредности $m_k > 1$.

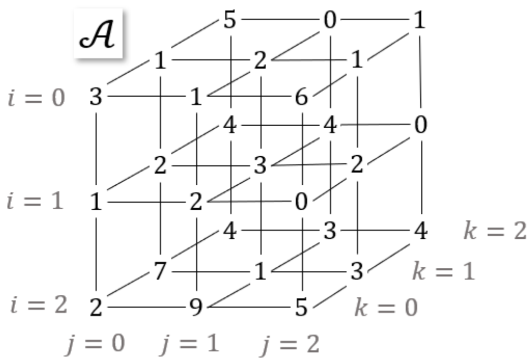
Пример 2.6.2. Нека је дат тродимензионални тензор $\mathcal{A} \in \mathcal{M}_{3\times3\times3}$ типа $3 \times 3 \times 3$ са слике 2.8. Поново је примењен Python начин индексирања.

Фронтална сечења овог тензора добијамо фиксирањем вредности трећег индекса k, док преостала два узимају све дозвољене вредности, $i, j = \overline{0, 2}$. Приликом записа пресека у матрице користимо i као индекс врсте, а j као индекс колоне.

$$\mathcal{A}(i,j,0) = \begin{bmatrix} 3 & 1 & 6 \\ 1 & 2 & 0 \\ 2 & 9 & 5 \end{bmatrix}, \quad \mathcal{A}(i,j,1) = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 7 & 1 & 3 \end{bmatrix}, \quad \mathcal{A}(i,j,2) = \begin{bmatrix} 5 & 0 & 1 \\ 4 & 4 & 0 \\ 4 & 3 & 4 \end{bmatrix}.$$



Слика 2.7: Сечење тензора на тензоре нижих димензија фиксирањем вредности једног или више индекса



Слика 2.8: Тензор $\mathcal{A} \in \mathcal{M}_{3\times3\times3}$

Хоризонтална сечења добијају се фиксирањем вредности првог индекса i, и варирањем вредности преостала два $j, k = \overline{0, 2}$. Приликом записа одговарајућих пресека у матрице користимо j као индекс врсте, а k као индекс колоне.

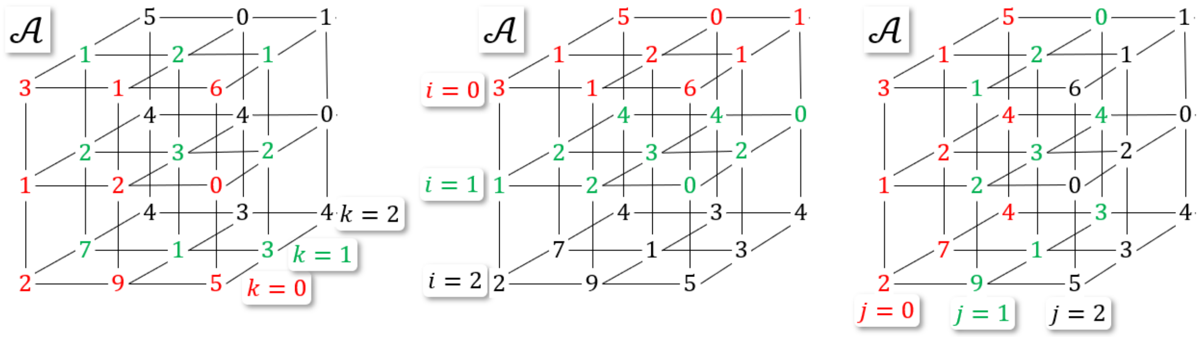
$$\mathcal{A}(0, j, k) = \begin{bmatrix} 3 & 1 & 5 \\ 1 & 2 & 0 \\ 6 & 1 & 1 \end{bmatrix}, \quad \mathcal{A}(1, j, k) = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 4 \\ 0 & 2 & 0 \end{bmatrix}, \quad \mathcal{A}(2, j, k) = \begin{bmatrix} 2 & 7 & 4 \\ 9 & 1 & 3 \\ 5 & 3 & 4 \end{bmatrix}.$$

Бочна сечења су резултат фиксирања вредности другог индекса j, и мењањем вредности преостала два $i, k = \overline{0, 2}$. Запис ових пресека у матрице користи i као индекс врсте, а k као индекс колоне.

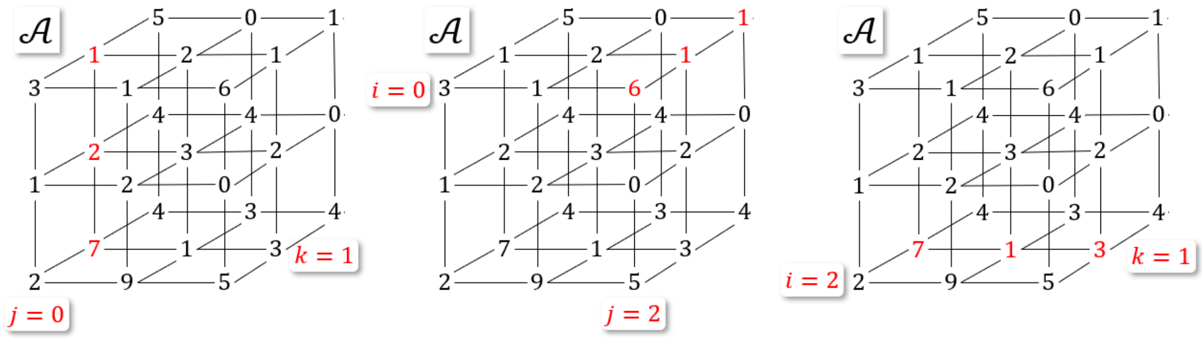
$$\mathcal{A}(i, 0, k) = \begin{bmatrix} 3 & 1 & 5 \\ 1 & 2 & 4 \\ 2 & 7 & 4 \end{bmatrix}, \quad \mathcal{A}(i, 1, k) = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 4 \\ 9 & 1 & 3 \end{bmatrix}, \quad \mathcal{A}(i, 2, k) = \begin{bmatrix} 6 & 1 & 1 \\ 0 & 2 & 0 \\ 5 & 3 & 4 \end{bmatrix}.$$

Издвајање вектора из тродимензионалног тензора $\mathcal{A} \in \mathcal{M}_{3\times3\times}$ добија се сечењем при коме се фиксирају вредности нека два индекса, слика 2.10. Тако нпр.

$$\mathcal{A}(i, 0, 1) = (1,\ 2,\ 7), \qquad \mathcal{A}(0, 2, k) = (6,\ 1,\ 1), \qquad \mathcal{A}(2, j, 1) = (7,\ 1,\ 3).$$



Слика 2.9: Фронтално, хоризонтално и бочно сечење тродимензионалног тензора



Слика 2.10: Примери сечење тензора по два фиксирана индекса

Вектори у тензору називају се влакна. Користе се за представљање вишедимензионалних матрица помоћу једне матрице - отварање или расклапање (unfolding) тензора по моду (оси), или за представљање тензора помоћу једног вектора вредности - векторизација тензора. Следи опис ове операциј применом Python начина индексирања.

Расклапање тензора $\mathcal{A} \in \mathcal{M}_{m_0 \times m_1 \times \cdots \times m_{n-1}}$ по $k-$том моду означавамо $\mathcal{A}_{[k]}$, а представља пресликавање

$$\mathcal{M}_{m_0 \times m_1 \times \cdots \times m_{n-1}} \mapsto \mathcal{M}_{m_k \times (m_0 \cdots m_{k-1} m_{k+1} \cdots m_{n-1})}$$
$$\mathcal{A} \mapsto \mathcal{A}_{[k]}.$$

Обилажење елемената $a_{i_0 i_1 \ldots i_{n-1}}$ тензора $\mathcal{A}$ обављамо са n угњеждених петљи по бројачима $i_0, i_1, \ldots, i_{n-1}$. Уколико пратимо Python конвенцију да се бројач врста најспорије мења, онда бројач који је највише постављен, тј. чије се вредности најспорије мењају је i_k. Он је индекс врста резултујуће матрице. Тако опсег вредности $0, 1, \ldots, m_k - 1$ бројача i_k дефинише висину колоне (тј. број врста) матрице $\mathcal{A}_{[k]}$. Преостали бројачи се кроз угњеждене петље ређају у дубину према њиховој основној пермутацији,

$$i_0, \ldots, i_{k-1}, i_{k+1}, \ldots, i_{n-1}.$$

Свака промена вредности најдубљег индекса означава почетак нове колоне у матрици расклапања $\mathcal{A}_{[k]}$.[2]

[2]Постоје и другачији приступи расклапању тензора, овај начин одговара имплементацији NumPy и TensorLy библиотека.

Математичка формула којом се описује овај процес позиционирања елемената тензора у матрицу је:

$$
\begin{aligned}
&(i_0, i_1, \dots, i_{n-1}) \mapsto (i_k, j), \quad \text{где је}\\
j &= i_{n-1} + m_{n-1} i_{n-2} + \cdots + i_{k+1} \prod_{s=k+2}^{n-1} m_s + i_{k-1} \prod_{s=k+1}^{n-1} m_s + i_{k-2} \prod_{k\neq s=k+1}^{n-1} m_s +\\
&\quad + \cdots + i_0 \prod_{k\neq s=1}^{n-1} m_s\\
&= \sum_{k\neq l=0}^{n-1} \left(i_l \prod_{k\neq s=l+1}^{n-1} m_s \right).
\end{aligned}
$$

Ову формулу можемо да користимо за реконструкцију тензора на основу матрице расклапања. Ову операцију зовемо склапање тензора по моду k (folding).

Векторизација тензора $\mathcal{A} \in \mathcal{M}_{m_0 \times \cdots \times m_{n-1}}$ је векторизација дуж колона матрице сечења тензора по последњем моду $\mathcal{A}_{[n-1]}$. Дакле, ради се о пресликавању

$$
\begin{aligned}
\mathcal{M}_{m_0 \times \cdots \times m_{n-1}} &\mapsto \mathbb{R}^{m_0 \cdots m_{n-1}}\\
\mathcal{A} &\mapsto vec(\mathcal{A}) = (\mathcal{A}_{[n-1]})_{[1]}.
\end{aligned}
$$

Формула којом се описује процес репозиционирања:

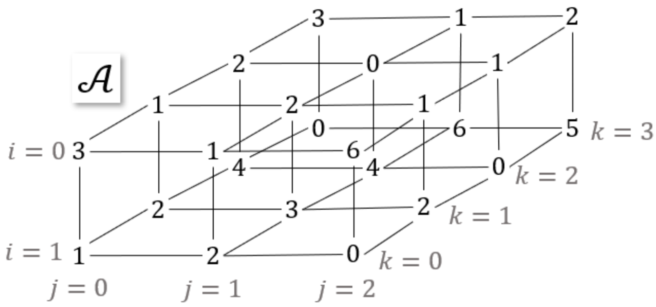
$$
(i_0, i_1, \dots, i_{n-1}) \mapsto j = \sum_{l=0}^{n-1} \left(i_l \prod_{s=l+1}^{n-1} m_s \right).
$$

Пример 2.6.3. Сликом 2.11 дата су расклапања тензора типа $3 \times 4 \times 2$ кроз влакна.

Слика 2.11: Графички приказ расклапања $3D$ тензора кроз влакна

Показаћемо расклапање тензора по моду и векторизацију на конкретном примеру.

Пример 2.6.4. Посматрамо тензор $\mathcal{A} \in \mathcal{M}_{2\times3\times4}$ са слике 2.12.



Слика 2.12: Тензор типа $2 \times 3 \times 4$

Расклапање тензора по моду 0 довешће до формирања матрице $\mathcal{A}_{[0]} \in \mathcal{M}_{2\times 12}$ на следећи начин, тј. кроз псеудо код:

$$
\begin{array}{l}
kolona = 0; \\
for\ i = \overline{0,1}: \\
\quad for\ j = \overline{0,2}: \\
\quad\quad for\ k = \overline{0,3}: \\
\quad\quad\quad matr(i, kolona) := a_{ijk} \\
\quad\quad\quad kolona{+}{+}
\end{array}
$$

$$
\mathcal{A}_{[0]} = \left[\begin{array}{cccc|cccc|cccc} 3 & 1 & 2 & 3 & 1 & 2 & 0 & 1 & 6 & 1 & 1 & 2 \\ 1 & 2 & 4 & 0 & 2 & 3 & 4 & 6 & 0 & 2 & 0 & 5 \end{array}\right] = \left[\begin{array}{c|c|c} \mathcal{A}(i,0,k) & \mathcal{A}(i,1,k) & \mathcal{A}(i,2,k) \end{array}\right]
$$

$$
= \left[\begin{array}{c} \mathcal{A}(0,j,k)_{[0]} \\ \hline \mathcal{A}(1,j,k)_{[0]} \end{array}\right].
$$

Расклапање тензора по моду 1 значи формирања матрице $\mathcal{A}_{[1]} \in \mathcal{M}_{3\times 8}$ на следећи начин:

$$
\begin{array}{l}
kolona = 0; \\
for\ j = \overline{0,2}: \\
\quad for\ i = \overline{0,1}: \\
\quad\quad for\ k = \overline{0,3}: \\
\quad\quad\quad matr(j, kolona) := a_{ijk} \\
\quad\quad\quad kolona{+}{+}
\end{array}
$$

$$
\mathcal{A}_{[1]} = \left[\begin{array}{cccc|cccc} 3 & 1 & 2 & 3 & 1 & 2 & 4 & 0 \\ 1 & 2 & 0 & 1 & 2 & 3 & 4 & 6 \\ 6 & 1 & 1 & 2 & 0 & 2 & 0 & 5 \end{array}\right] = \left[\begin{array}{c|c} \mathcal{A}(0,j,k) & \mathcal{A}(1,j,k) \end{array}\right]
$$

$$
= \left[\begin{array}{c} \mathcal{A}(i,0,k)_{[0]} \\ \hline \mathcal{A}(i,1,k)_{[0]} \\ \hline \mathcal{A}(i,2,k)_{[0]} \end{array}\right].
$$

Расклапање тензора по моду 2 даје матрицу $\mathcal{A}_{[2]} \in \mathcal{M}_{4\times 6}$:

$$\begin{aligned}
&kolona = 0;\\
&for\ k = \overline{0,3}:\\
&\quad for\ i = \overline{0,1}:\\
&\qquad for\ j = \overline{0,2}:\\
&\qquad\quad matr(k, kolona) := a_{ijk}\\
&\qquad\quad kolona++
\end{aligned}$$

$$\mathcal{A}_{[2]} = \left[\begin{array}{ccc|ccc} 3 & 1 & 6 & 1 & 2 & 0\\ 1 & 2 & 1 & 2 & 3 & 2\\ 2 & 0 & 1 & 4 & 4 & 0\\ 3 & 1 & 2 & 0 & 6 & 5 \end{array}\right] = \left[\begin{array}{c|c} \mathcal{A}(0,j,k)^T & \mathcal{A}(1,j,k)^T \end{array}\right]$$

$$= \left[\begin{array}{c} \mathcal{A}(i,j,0)_{[0]}\\ \hline \mathcal{A}(i,j,1)_{[0]}\\ \hline \mathcal{A}(i,j,2)_{[0]}\\ \hline \mathcal{A}(i,j,3)_{[0]} \end{array}\right].$$

Векторизацијом матрице $\mathcal{A}_{[2]}$ по колонама или векторизацијом матрице $\mathcal{A}_{[0]}$ по врстама добијамо векторизацију тензора $vec(\mathcal{A}) = (\mathcal{A}_{[0]})_{[0]}$

$$vec(\mathcal{A}) = (3,\ 1,\ 2,\ 3,\ 1,\ 2,\ 0,\ 1,\ 6,\ 1,\ 1,\ 2,\ 1,\ 2,\ 4,\ 0,\ 2,\ 3,\ 4,\ 6,\ 0,\ 2,\ 0,\ 5).$$

Добијена векторизација тензора $\mathcal{A}$ може се описати кроз наредни псеудо код.

$$\begin{aligned}
&pozicija = 0;\\
&for\ i = \overline{0,1}:\\
&\quad for\ j = \overline{0,2}:\\
&\qquad for\ k = \overline{0,3}:\\
&\qquad\quad vec(pozicija) := a_{ijk}\\
&\qquad\quad pozicija++
\end{aligned}$$

На сличан начин би могла да се дефинише векторизација матрице по моду k

$$vec_k(\mathcal{A}) = (\mathcal{A}_{[k]})_{[0]}.$$

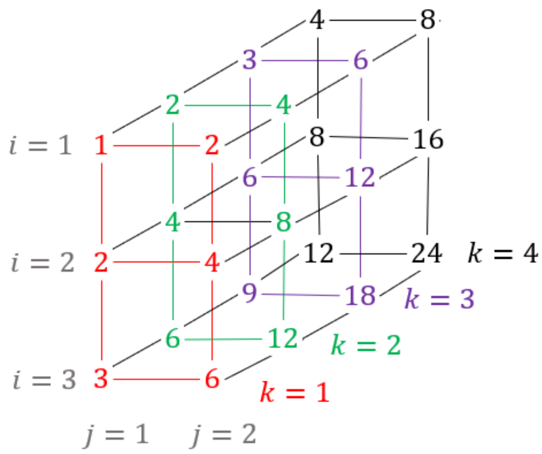
Пример 2.6.5. Сечење тензора помаже нам и да схватимо процес израчунавања у $n-$линеарном пресликавању. Посматрамо $3-$димензионални тензор типа $3 \times 2 \times 4$ са елементима

$$a_{ijk} = ijk, \quad i \in [3],\ j \in [2],\ k \in [4].$$

Како овај тензор дефинише $3-$линеарно пресликавање $\mathcal{A}(v_1, v_2, v_3)$? Томе служе базисни вектори

$$e_i^{[3]} \in \mathbb{R}^3,\ i \in [3], \quad e_j^{[2]} \in \mathbb{R}^2,\ j \in [2], \quad e_k^{[4]} \in \mathbb{R}^4,\ k \in [4]$$

и једнакост $\mathcal{A}(e_i^{[3]}, e_j^{[2]}, e_k^{[4]}) = a_{ijk}$. Наиме, фиксирањем вредности $k = 1$ пресликавање $\mathcal{A}_1(v_1, v_2) = \mathcal{A}(v_1, v_2, e_1^{[4]})$ је билинеарно, тј. описано је у облику матричног производа



Слика 2.13: Матрице $A_1,\ A_2,\ A_3,\ A_4$ сечења тензора $\mathcal{A}$

$v_1^T A_1 v_2$, где је $A_1 = [a_{ij1}]$. Приметимо да је то матрица са елементима означеним црвеном бојом на слици 2.13, тј. прво фронтално сечење.

Аналогно, билинеарна пресликавања

$$\mathcal{A}_k(v_1, v_2) = \mathcal{A}(v_1, v_2, e_k^{[4]}), \quad k \in [4]$$

описана су матричним производом

$$v_1^T A_k v_2, \quad A_k = [a_{ijk}],\ k \in [4].$$

Матрице A_k су назначене елементима различитих боја на слици 2.13 и представљају фронтална сечења овог тензора. За неке конкретне векторе $v_1 \in \mathbb{R}^3,\ v_2 \in \mathbb{R}^2,\ v_3 \in \mathbb{R}^4$ искористимо њихово представљање у одговарајућој природној бази, нпр.

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -1 \\ 2 \\ -1 \\ 1 \end{bmatrix}.$$

Тада је

$$\begin{aligned}
\mathcal{A}(v_1, v_2, v_3) &= \mathcal{A}(v_1, v_2, -1 \cdot e_1^{[4]} + 2 \cdot e_2^{[4]} - 1 \cdot e_3^{[4]} + 1 \cdot e_4^{[4]}) \\
&= -1 \cdot \mathcal{A}(v_1, v_2, e_1^{[4]}) + 2 \cdot \mathcal{A}(v_1, v_2, e_2^{[4]}) - 1 \cdot \mathcal{A}(v_1, v_2, e_3^{[4]}) + 1 \cdot \mathcal{A}(v_1, v_2, e_4^{[4]}) \\
&= -1 \cdot \mathcal{A}_1(v_1, v_2) + 2 \cdot \mathcal{A}_2(v_1, v_2) - 1 \cdot \mathcal{A}_3(v_1, v_2) + 1 \cdot \mathcal{A}_4(v_1, v_2) \\
&= -1 \cdot v_1^T A_1 v_2 + 2 \cdot v_1^T A_2 v_2 - 1 \cdot v_1^T A_3 v_2 + 1 \cdot v_1^T A_4 v_2 \\
&= v_1^T(-A_1 + 2A_2 - A_3 + A_4) v_2 \\
&= \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \left(- \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} + 2 \begin{bmatrix} 2 & 4 \\ 4 & 8 \\ 6 & 12 \end{bmatrix} - \begin{bmatrix} 3 & 6 \\ 6 & 12 \\ 9 & 18 \end{bmatrix} + \begin{bmatrix} 4 & 8 \\ 8 & 16 \\ 12 & 24 \end{bmatrix} \right) \begin{bmatrix} 2 \\ -3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 8 & 16 \\ 12 & 24 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} -16 \\ -32 \\ -48 \end{bmatrix} = 0.
\end{aligned}$$

Закључујемо да координате вектора v_3 служе као коефицијенти линеарне комбинације билинеарних пресликавања $\mathcal{A}_k$, односно одговарајућих матрица ових пресликавања.

На сличан начин смо израчунавање $\mathcal{A}(v_1, v_2, v_3)$ могли да анализирамо кроз билинеарна пресликавања која се добијају фиксирањем вредности првог или другог индекса (i или j) у бази простора $\mathbb{R}^3$ или $\mathbb{R}^2$.

Претходни пример објашњава поступак израчунавања $\mathcal{A}(v_1, v_2, \dots, v_{n-1}, v_n)$ снижавањем димензионалности тензора. Уколико је $v_n = \begin{bmatrix} x_1 & x_2 & \dots & x_{m_n} \end{bmatrix}^T$ и $e_i \in \mathbb{R}^{m_n}$, тада је разлагање по оси i_n дато са

$$\begin{aligned}
\mathcal{A}(v_1, v_2, \dots, v_{n-1}, v_n) &= x_1 \mathcal{A}(v_1, v_2, \dots, v_{n-1}, e_1) + x_2 \mathcal{A}(v_1, v_2, \dots, v_{n-1}, e_2) + \dots \\
&\qquad + x_{m_n} \mathcal{A}(v_1, v_2, \dots, v_{n-1}, e_{m_n}) \\
&= x_1 \mathcal{A}_1(v_1, v_2, \dots, v_{n-1}) + x_2 \mathcal{A}_2(v_1, v_2, \dots, v_{n-1}) + \dots \\
&\qquad + x_{m_n} \mathcal{A}_{m_n}(v_1, v_2, \dots, v_{n-1}).
\end{aligned}$$

Аналогно разлагање добија се по било којој оси тензора $\mathcal{A}$. Настављајући овај поступак снижавања димензије можемо доћи до билинеарних или линеарних тензора. Ту се поступак снижавања димензионалности завршава и примењују класична матрична израчунавања. Описани поступак је битан јер обрнутим процесом вршимо конструкцију тензора већих димензија. То је тема наредног одељка.

Пре него што било каква озбиљнија анализа рада са тензорима отпочне, треба нагласити да без обзира на једноставно продужење матрица на тензоре, операције са овим објектима нису тако једноставно продужење. Операције над матрицама су познате, детаљно испитиване и разумљиве теме. Са друге стране, операције са мултидимензионалним матрицама и даље представљају делимичну непознаницу. Потрага за адекватним алатима за проучавање тензора је веома активна област истраживања математичара, статистичара, инжењера рачунарских наука и инжењера других домена. Разлог велике ангажованости на разумевању тензора је све јаснија потреба за њима у решавању реалних проблема. Осим тога велике рачунске могућности модерних рачунара омогућавају рад са овим типом објеката у реалном времену. Ипак треба бити на опрезу са величином димензије и типа тензора с обзиром да количина података коју носе расте степено са димензијом.

2.7 Алгебра тензора

Операције које скуп тензора $\mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}$ чине векторским простором су покомпонентно сабирање тензора и скалирање, тј. покомпонентно множење скаларом.

$$\mathcal{A}, \mathcal{B} \in \mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}, \qquad \mathcal{A} = [a_{i_1 i_2 \dots i_n}], \quad \mathcal{B} = [b_{i_1 i_2 \dots i_n}], \qquad \alpha \in \mathbb{R},$$

$$\mathcal{A} + \mathcal{B} = [a_{i_1 i_2 \dots i_n} + b_{i_1 i_2 \dots i_n}] \in \mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}, \qquad \alpha\,\mathcal{A} = [\alpha\, a_{i_1 i_2 \dots i_n}] \in \mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}.$$

Димензија овог векторског простора је $m_1 \cdot m_2 \dots m_n$. По угледу на векторски простор матрица, једну базу простора $\mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}$ чине тензори одговарајућег типа чији су сви елементи једнаки 0, осим на једној позицији где је вредност 1. Означаваћемо их $\mathcal{E}_{i_1 i_2 \dots i_n}$, где је уређеном $n-$торком $(i_1, i_2, \dots, i_n)$ управо обележена позиција вредности 1. Нула-вектор је тензор који има све елементе једнаке 0, означаваћемо га са $\mathcal{O}$.

Напомена 3. Обратити пажњу на разлику између димензионалности објекта и димензије векторског простора коме овај објекат припада. За тензор $\mathcal{A} \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}$ кажемо да је $n-$димензионалан, док је векторски простор $\mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}$ коме он припада димензије $\dim(\mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}) = m_1 \cdot m_2 \cdots m_n$.

Баш као и са матрицама, кључна операција над тензорима јесте множење. Њом се описује веза између векторских простора $\mathbb{R}^{m_i}$, $i \in [n]$, и $\mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}$. Ипак, тензорски производ није продужетак класичног множења матрица. Најједноставнији начин дефинисања операције тензорског производа јесте кроз пресликавања.

Дефиниција 11. Нека су $\mathcal{A} \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}$ и $\mathcal{B} \in \mathcal{M}_{s_1 \times s_2 \times \cdots \times s_r}$ тензори. Њихов тензорски производ $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$ је тензор $\mathcal{C} \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n \times s_1 \times s_2 \times \cdots \times s_r}$ за који важи

$$\begin{aligned}
&\forall v_i \in \mathbb{R}^{m_i},\ i \in [n], \qquad \forall u_j \in \mathbb{R}^{s_j},\ j \in [r], \\
&\mathcal{C}(v_1, v_2, \ldots, v_n, u_1, u_2, \ldots, u_r) = \mathcal{A}(v_1, v_2, \ldots, v_n)\, \mathcal{B}(u_1, u_2, \ldots, u_r).
\end{aligned}$$

Елементе вишедимензионалне матрице којом је такође описан тензор $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$ тада добијам сликањем вектора елемената канонских база одговарајућих простора,

$$\begin{aligned}
\mathcal{A} = [a_{i_1 i_2 \ldots i_n}] \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}, &\qquad \mathcal{B} = [b_{j_1 j_2 \ldots j_r}] \in \mathcal{M}_{s_1 \times s_2 \times \cdots \times s_r}, \\
\mathcal{C} = [c_{i_1 i_2 \ldots i_n j_1 j_2 \ldots j_r}] \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n \times s_1 \times s_2 \times \cdots \times s_r}, \\
c_{i_1 i_2 \ldots i_n j_1 j_2 \ldots j_r} &= \mathcal{C}\big(e_{i_1}^{[m_1]}, e_{i_2}^{[m_2]}, \ldots, e_{i_n}^{[m_n]}, e_{j_1}^{[s_1]}, e_{j_2}^{[s_2]}, \ldots, e_{j_r}^{[s_r]}\big) \\
&= \mathcal{A}\big(e_{i_1}^{[m_1]}, e_{i_2}^{[m_2]}, \ldots, e_{i_n}^{[m_n]}\big)\, \mathcal{B}\big(e_{j_1}^{[s_1]}, e_{j_2}^{[s_2]}, \ldots, e_{j_r}^{[s_r]}\big) \\
&= a_{i_1 i_2 \ldots i_n}\, b_{j_1 j_2 \ldots j_r}.
\end{aligned}$$

На основу дефиниције 11 лако закључујемо особине ове операције. Пре свега, приметимо да операција тензорског производа није комутативна. Наиме, изрази $\mathcal{A} \otimes \mathcal{B}$ и $\mathcal{B} \otimes \mathcal{A}$ су дефинисани истовремено. Разлика између два пресликавања огледа се у редоследу навођења аргумената. Другим речима, резултујући тензори су различитог типа који се добијају одговарајућом пермутацијом оса тензора.

$$\begin{aligned}
&\mathcal{C} = \mathcal{A} \otimes \mathcal{B} \in \mathcal{M}_{m_1 \times \cdots \times m_n \times s_1 \times \cdots \times s_r}, \quad \mathcal{D} = \mathcal{B} \otimes \mathcal{A} \in \mathcal{M}_{s_1 \times \cdots \times s_r \times m_1 \times \cdots \times m_n}, \\
&\qquad v_i \in \mathbb{R}^{m_i},\ i \in [n], \qquad u_j \in \mathbb{R}^{s_j},\ j \in [r], \\
&\mathcal{C}(v_1, v_2, \ldots, v_n, u_1, u_2, \ldots, u_r) = \mathcal{A}(v_1, v_2, \ldots, v_n)\, \mathcal{B}(u_1, u_2, \ldots, u_r), \\
&\mathcal{D}(u_1, u_2, \ldots, u_r, v_1, v_2, \ldots, v_n) = \mathcal{B}(u_1, u_2, \ldots, u_r)\, \mathcal{A}(v_1, v_2, \ldots, v_n).
\end{aligned}$$

Тензорски производ јесте асоцијативна операција,

$$\begin{aligned}
&\mathcal{A} \in \mathcal{M}_{m_1 \times \cdots \times m_n}, \quad \mathcal{B} \in \mathcal{M}_{s_1 \times \cdots \times s_r}, \quad \mathcal{C} \in \mathcal{M}_{p_1 \times \cdots \times p_q}, \\
&\mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C}) \in \mathcal{M}_{m_1 \times \cdots \times m_n \times s_1 \times \cdots \times s_r \times p_1 \times \cdots \times p_q}, \\
&(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C} \in \mathcal{M}_{m_1 \times \cdots \times m_n \times s_1 \times \cdots \times s_r \times p_1 \times \cdots \times p_q}, \\
&\qquad \mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C}) = (\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}.
\end{aligned}$$

Ову једнакост лако добијамо кроз проверу појединачних елемената тензора.

$$\begin{aligned}
\mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C})\big(e_{i_1}^{[m_1]}, \dots, e_{i_n}^{[m_n]}, e_{j_1}^{[s_1]}, \dots, e_{j_r}^{[s_r]}, e_{k_1}^{[p_1]}, \dots, e_{k_q}^{[p_q]}\big)& \\
= \mathcal{A}\big(e_{i_1}^{[m_1]}, \dots, e_{i_n}^{[m_n]}\big)(\mathcal{B} \otimes \mathcal{C})\big(e_{j_1}^{[s_1]}, \dots, e_{j_r}^{[s_r]}, e_{k_1}^{[p_1]}, \dots, e_{k_q}^{[p_q]}\big)& \\
= \mathcal{A}\big(e_{i_1}^{[m_1]}, \dots, e_{i_n}^{[m_n]}\big)\mathcal{B}\big(e_{j_1}^{[s_1]}, \dots, e_{j_r}^{[s_r]}\big)\mathcal{C}\big(e_{k_1}^{[p_1]}, \dots, e_{k_q}^{[p_q]}\big),&
\end{aligned}$$

$$\begin{aligned}
(\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C}\big(e_{i_1}^{[m_1]}, \dots, e_{i_n}^{[m_n]}, e_{j_1}^{[s_1]}, \dots, e_{j_r}^{[s_r]}, e_{k_1}^{[p_1]}, \dots, e_{k_q}^{[p_q]}\big)& \\
= (\mathcal{A} \otimes \mathcal{B})\big(e_{i_1}^{[m_1]}, \dots, e_{i_n}^{[m_n]}, e_{j_1}^{[s_1]}, \dots, e_{j_r}^{[s_r]}\big)\mathcal{C}\big(e_{k_1}^{[p_1]}, \dots, e_{k_q}^{[p_q]}\big)& \\
= \mathcal{A}\big(e_{i_1}^{[m_1]}, \dots, e_{i_n}^{[m_n]}\big)\mathcal{B}\big(e_{j_1}^{[s_1]}, \dots, e_{j_r}^{[s_r]}\big)\mathcal{C}\big(e_{k_1}^{[p_1]}, \dots, e_{k_q}^{[p_q]}\big).&
\end{aligned}$$

Пример 2.7.1. Нека су дата два једнодимензионална тензора типа m и n редом, тј. два вектора $v \in \mathbb{R}^m$ и $u \in \mathbb{R}^n$,

$$v = \begin{bmatrix}\alpha_1 & \alpha_2 & \dots & \alpha_m\end{bmatrix}^T, \qquad u = \begin{bmatrix}\beta_1 & \beta_2 & \dots & \beta_n\end{bmatrix}^T.$$

Ови тензори дефинишу два линеарна пресликавања,

$$\begin{aligned}
& v \ : \ \mathbb{R}^m \mapsto \mathbb{R}, && u \ : \ \mathbb{R}^n \mapsto \mathbb{R}, \\
& \mathbf{x} = \begin{bmatrix}x_1 & x_2 & \dots & x_m\end{bmatrix}^T, && \mathbf{y} = \begin{bmatrix}y_1 & y_2 & \dots & y_n\end{bmatrix}^T, \\
& \mathbf{x} \mapsto v^T\mathbf{x} = \mathbf{x}^T v = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_m x_m, \\
& \mathbf{y} \mapsto u^T\mathbf{y} = \mathbf{y}^T u = \beta_1 y_1 + \beta_2 y_2 + \dots + \beta_n y_n.
\end{aligned}$$

Тада је $v \otimes u$ билинеарно пресликавање, тј. дводимензионални тензор типа $m \times n$,

$$\begin{aligned}
& A = v \otimes u \ : \ \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}, \\
& A(\mathbf{x}, \mathbf{y}) = (v^T\mathbf{x})(u^T\mathbf{y}) = (\mathbf{x}^T v)(u^T\mathbf{y}) = \mathbf{x}^T(vu^T)\mathbf{y}.
\end{aligned}$$

Закључујемо да се у случају тензорског производа два вектора ова операција поклапа са њиховим спољашњим производом. Овде није било потребно анализирати елементе тензорског производа кроз координатне векторе. Слично се добија,

$$\begin{aligned}
& B = u \otimes v \ : \ \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}, \\
& B(\mathbf{y}, \mathbf{x}) = (u^T\mathbf{y})(v^T\mathbf{x}) = (\mathbf{y}^T u)(v^T\mathbf{x}) = \mathbf{y}^T(uv^T)\mathbf{x}.
\end{aligned}$$

Претходни пример омогућава разумевање тензорског производа више вектора. Тензори овог типа имају велики значај у простору тензора баш као и спољашњи производ вектора у простору матрица. Зовемо их још елементарни тензори.

Нека су $v_i \in \mathbb{R}^{m_i}, \ i \in [n]$,

$$v_i = \begin{bmatrix}\alpha_1^{(i)} & \alpha_2^{(i)} & \dots & \alpha_{m_i}^{(i)}\end{bmatrix}^T.$$

Тензорски производ ових вектора $\mathcal{A} = v_1 \otimes v_2 \otimes \dots \otimes v_n \in \mathcal{M}_{m_1 \times m_2 \times \dots \times m_n}$ тада је

$$\begin{aligned}
\mathcal{A}(e_{i_1}^{[m_1]}, e_{i_2}^{[m_2]}, \dots, e_{i_n}^{[m_n]}) &= \big(v_1 \otimes v_2 \otimes \dots \otimes v_n\big)(e_{i_1}^{[m_1]}, e_{i_2}^{[m_2]}, \dots, e_{i_n}^{[m_n]}) \\
&= v_1(e_{i_1}^{[m_1]})\ v_2(e_{i_2}^{[m_2]}) \cdots v_n(e_{i_n}^{[m_n]}) \\
&= \big(v_1^T e_{i_1}^{[m_1]}\big)\big(v_2^T e_{i_2}^{[m_2]}\big) \cdots \big(v_n^T e_{i_n}^{[m_n]}\big) \\
&= \alpha_{i_1}^{(1)} \alpha_{i_2}^{(2)} \cdots \alpha_{i_n}^{(n)}.
\end{aligned}$$

Пример 2.7.2. Тензорски производ вектора нам омогућава једноставан опис елемената базе $\mathcal{E}_{i_1 i_2 \ldots i_n}$ простора $\mathcal{M}_{m_1\times\cdots\times m_n}$. Кренућемо од тензора малих димензија како бисмо уочили аналогију која се дешава са повећањем димензије. Као и до сада вектори канонске базе простора $\mathbb{R}^m$ означени су са $e_i^{[m]}$, $i \in [m]$. За простор матрица $\mathcal{M}_{m\times n}$ елементе базе E_{ij}, $i \in [m]$, $j \in [n]$ лако добијамо спољашњим, тј. тензорским производом вектора база $e_i^{[m]}$ и $e_j^{[n]}$,

$$E_{ij} = e_i^{[m]}(e_j^{[n]})^T = e_i^{[m]} \otimes e_j^{[n]}.$$

На пример, за $m = 2$ и $n = 3$ имамо

$$E_{11} = \begin{bmatrix}1 & 0 & 0\\0 & 0 & 0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}1 & 0 & 0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix},$$

$$E_{12} = \begin{bmatrix}0 & 1 & 0\\0 & 0 & 0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}0 & 1 & 0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\\0\end{bmatrix},$$

$$E_{13} = \begin{bmatrix}0 & 0 & 1\\0 & 0 & 0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}\begin{bmatrix}0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\0\\1\end{bmatrix},$$

$$E_{21} = \begin{bmatrix}0 & 0 & 0\\1 & 0 & 0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix},$$

$$E_{22} = \begin{bmatrix}0 & 0 & 0\\0 & 1 & 0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}0 & 1 & 0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\\0\end{bmatrix},$$

$$E_{23} = \begin{bmatrix}0 & 0 & 0\\0 & 0 & 1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}\begin{bmatrix}0 & 0 & 1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Једнакост $E_{ij} = e_i^{[m]} \otimes e_j^{[n]}$ можемо да докажемо и кроз билинеарну форму. Прво приметимо да је за $e_i, e_k \in \mathbb{R}^s$ тачно

$$e_i^T e_k = \delta_{ik} = \begin{cases}1, & i = k,\\ 0, & i \neq k.\end{cases}$$

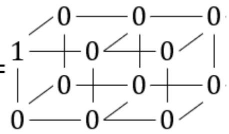
Вредност на позицији k, l тензора E_{ij} добијамо израчунавањем $E_{ij}(e_k^{[m]}, e_l^{[n]})$, тј.

$$\begin{aligned}E_{ij}(e_k^{[m]}, e_l^{[n]}) &= \big(e_i^{[m]} \otimes e_j^{[n]}\big)(e_k^{[m]}, e_l^{[n]}) = e_i^{[m]}(e_k^{[m]})\ e_j^{[n]}(e_l^{[n]})\\ &= \big((e_i^{[m]})^T e_k^{[m]}\big)\ \big((e_j^{[n]})^T e_l^{[n]}\big) = \delta_{ik}\,\delta_{jl} = 1 \iff i = k \wedge j = l.\end{aligned}$$

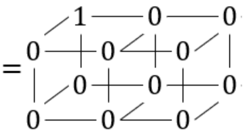
Код тродимензионалних тензора $\mathcal{M}_{m\times n\times p}$ елементе базе $\mathcal{E}_{ijk}$ лако описујемо са

$$\mathcal{E}_{ijk} = e_i^{[m]} \otimes e_j^{[n]} \otimes e_k^{[p]}.$$

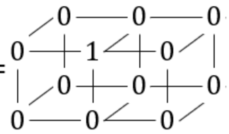
Рецимо, за $m = 2, n = 3$ и $p = 2$ имамо:

$\mathcal{E}_{111} =$ 

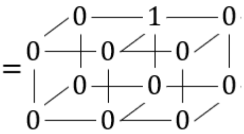
$= \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$

$\mathcal{E}_{112} =$ 

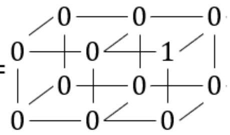
$= \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$

$\mathcal{E}_{121} =$ 

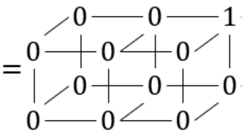
$= \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$

$\mathcal{E}_{122} =$ 

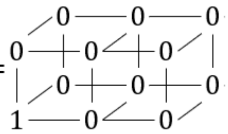
$= \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$

$\mathcal{E}_{131} =$ 

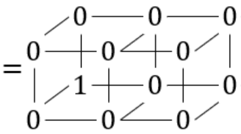
$= \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$

$\mathcal{E}_{132} =$ 

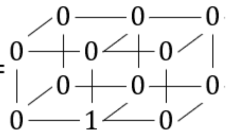
$= \begin{bmatrix}1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$

$\mathcal{E}_{211} =$ 

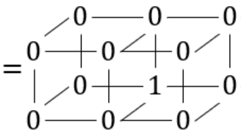
$= \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$

$\mathcal{E}_{212} =$ 

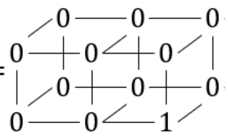
$= \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$

$\mathcal{E}_{221} =$ 

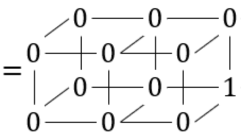
$= \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\\0\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$

$\mathcal{E}_{222} =$ 

$= \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\\0\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$

$\mathcal{E}_{231} =$ 

$= \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\0\\1\end{bmatrix} \otimes \begin{bmatrix}1\\0\end{bmatrix}$

$\mathcal{E}_{232} =$ 

$= \begin{bmatrix}0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\0\\1\end{bmatrix} \otimes \begin{bmatrix}0\\1\end{bmatrix}$

У општем случају, једнакост $\mathcal{E}_{i_1 i_2 i_3} = e_{i_1}^{[m_1]} \otimes e_{i_2}^{[m_2]} \otimes e_{i_3}^{[m_3]}$ доказујемо кроз пресликавања. За $j_k \in [m_k],\ k \in [3]$, на позицији (j_1, j_2, j_3) у тензору $\mathcal{E}_{i_1 i_2 i_3}$ налази се вредност:

$$\begin{aligned}
\mathcal{E}_{i_1 i_2 i_3}(e_{j_1}^{[m_1]}, e_{j_2}^{[m_2]}, e_{j_3}^{[m_3]}) &= \big(e_{i_1}^{[m_1]} \otimes e_{i_2}^{[m_2]} \otimes e_{i_3}^{[m_3]}\big)(e_{j_1}^{[m_1]}, e_{j_2}^{[m_2]}, e_{j_3}^{[m_3]}) \\
&= e_{i_1}^{[m_1]}(e_{j_1}^{[m_1]})\ e_{i_2}^{[m_2]}(e_{j_2}^{[m_2]})\ e_{i_3}^{[m_3]}(e_{j_3}^{[m_3]}) \\
&= \prod_{k=1}^{3} e_{i_k}^{[m_k]}(e_{j_k}^{[m_k]}) = \prod_{k=1}^{3} (e_{i_k}^{[m_k]})^T e_{j_k}^{[m_k]} \\
&= \prod_{k=1}^{3} \delta_{i_k j_k} = 1 \iff i_k = j_k,\ \forall k \in [3].
\end{aligned}$$

Приступ кроз $n-$линеарна пресликавања помоћи ће да потврдимо да за $\mathcal{E}_{i_1 i_2 \ldots i_n} \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}$, важи

$$\mathcal{E}_{i_1 i_2 \ldots i_n} = e_{i_1}^{[m_1]} \otimes e_{i_2}^{[m_2]} \otimes \cdots \otimes e_{i_n}^{[m_n]}.$$

За $j_k \in [m_k],\ k \in [n]$, на позицији $(j_1, j_2, \ldots, j_n)$ налази се вредност:

$$\begin{aligned}
\mathcal{E}_{i_1 i_2 \ldots i_n}(e_{j_1}^{[m_1]}, e_{j_2}^{[m_2]}, \ldots, e_{j_n}^{[m_n]}) &= \big(e_{i_1}^{[m_1]} \otimes e_{i_2}^{[m_2]} \otimes \cdots \otimes e_{i_n}^{[m_n]}\big)(e_{j_1}^{[m_1]}, e_{j_2}^{[m_2]}, \ldots, e_{j_n}^{[m_n]}) \\
&= \prod_{k=1}^{n} e_{i_k}^{[m_k]}(e_{j_k}^{[m_k]}) = \prod_{k=1}^{n} \delta_{i_k j_k} = 1 \iff i_k = j_k,\ \forall k \in [n].
\end{aligned}$$

У наредним примерима анализираћемо садржај вишедимензионалне табеле која је резултат тензорског производа вишедимензионалних тензора.

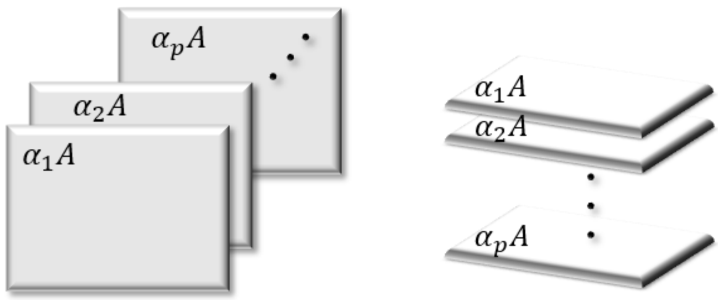
Пример 2.7.3. Нека су $A = [a_{ij}] \in \mathcal{M}_{m\times n}$ и $w = [\alpha_k] \in \mathbb{R}^p$ тензори наведених типова, и $\mathcal{B} = [b_{ijk}] = A \otimes w \in \mathcal{M}_{m\times n\times p}$ резултат њиховог тензорског производа. Потражићемо елементе одговарајуће тродимензионалне матрице.

$$b_{ijk} = \mathcal{B}\big(e_i^{[m]}, e_j^{[n]}, e_k^{[p]}\big) = A\big(e_i^{[m]}, e_j^{[n]}\big)\, w\big(e_k^{[p]}\big) = a_{ij}\alpha_k.$$

У обрнутом редоследу производ тензора A и w даје следеће:

$$\mathcal{C} = [c_{kij}] = w \otimes A \in \mathcal{M}_{p\times m\times n}, \quad c_{kij} \; = \mathcal{C}\big(e_k^{[p]}, e_i^{[m]}, e_j^{[n]}\big) = w\big(e_k^{[p]}\big)\, A\big(e_i^{[m]}, e_j^{[n]}\big) = \alpha_k a_{ij}.$$

Закључујемо да се тродимензионална матрица састоји од матрица $\alpha_k A$, $k \in [p]$, које су у зависности од редоследа операнада, $A \otimes w$ или $w \otimes A$, фронтална или хоризонтална сечења тензора, слика 2.14.

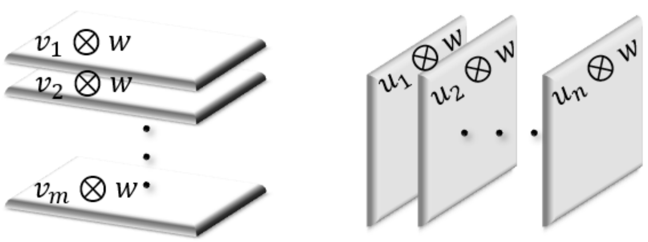


Слика 2.14: Резултат $A \otimes w$ и $w \otimes A$

Згодно је анализирати производ $\mathcal{B} = A \otimes w = [a_{ij}\alpha_k] = [b_{ijk}]$ и кроз фиксирање вредности индекса i или j. У те сврхе уведимо ознаке вектора врста и колона матрице A,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \dots & u_n \end{array}\right] = \left[\begin{array}{c} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_m^T \end{array}\right].$$

Тада нпр. за $i = 1$ елементи сечења тензора $\mathcal{B}$ су $\mathcal{B}(1, j, k) = [a_{1j}\alpha_k]$, тј елементи тензорског производа вектора $v_1 \otimes w \in \mathcal{M}_{n\times p}$. Слично, фиксирањем нпр. $j = 1$ елементи тензора $\mathcal{B}$ добијени сечењем су $\mathcal{B}(i, 1, k) = [a_{i1}\alpha_k] = u_1 \otimes w \in \mathcal{M}_{m\times p}$, слика 2.15. На сличан начин описујемо фронтална и бочна сечења тензора $w \otimes A$.



Слика 2.15: Сечење тензора $A \otimes w$ по i и j

Приметимо да када је $n = p$ матрица $A = [a_{ij}] \in \mathcal{M}_{m\times n}$ и вектор $w \in \mathbb{R}^n$ могу да се помноже класично, као две матрице. Резултујући вектор $v = Aw = [\beta_i]^T \in \mathbb{R}^m$ можемо да добијемо на основу појединих елемената тензорског производа $A \otimes w$,

$$\beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j = \sum_{j=1}^{n} b_{ijj}.$$

Пример 2.7.4. Нека су сада $A = [a_{ij}] \in \mathcal{M}_{m\times n}$ и $B = [b_{kl}] \in \mathcal{M}_{p\times q}$ два дводимензионална тензора наведеног типа. Одредићемо елементе тензора

$$\mathcal{C} = [c_{ijkl}] = A \otimes B \in \mathcal{M}_{m\times n\times p\times q},$$
$$c_{ijkl} = \mathcal{C}\left(e_i^{[m]}, e_j^{[n]}, e_k^{[p]}, e_k^{[q]}\right) = A\left(e_i^{[m]}, e_j^{[n]}\right) B\left(e_k^{[p]}, e_l^{[q]}\right) = a_{ij}b_{kl}. \tag{2.13}$$

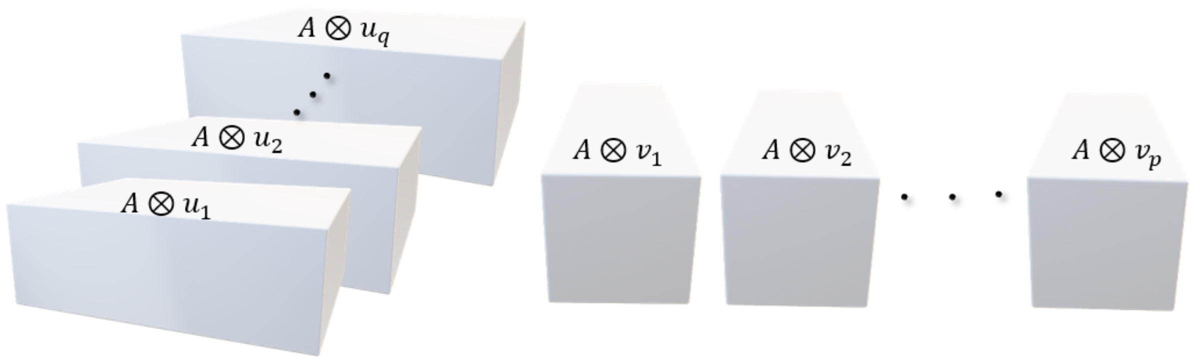
Приметимо да се елементи (класичног) производа $AB = [\alpha_{il}] \in \mathcal{M}_{m\times q}$ ове две матрице добијају када је $n = p$ са

$$\alpha_{il} = \sum_{j=1}^{n} a_{ij}b_{jl} = \sum_{j=1}^{n} c_{ijjl}.$$

Уведимо ознаке колона и врста матрице B,

$$B = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \dots & u_q \end{array}\right] = \left[\begin{array}{c} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_p^T \end{array}\right].$$

Формула (2.13) сугерише да за фиксирано l сечење тензора $\mathcal{C}$ представља тензор $A\otimes u_l$. Слично, фиксирањем вредности k сечење тензора $\mathcal{C}$ представља тензор $A\otimes v_k$, слика 2.16.



Слика 2.16: Сечење тензора $A \otimes B$

Аналогну анализу добијамо сечењем тензора $A \otimes B$ по i и j уз помоћ разлагања матрице A на врсте и колоне.

У наставку упознајемо још особина тензорског производа. Њихове доказе остављамо читаоцу као лаку вежбу.

Теорема 10. Нека су $\mathcal{A}, \mathcal{B} \in \mathcal{M}_{m_1 \times \cdots \times m_n}$ и $\mathcal{C}, \mathcal{D} \in \mathcal{M}_{s_1 \times \cdots \times s_r}$ тензори наведених типова, и $\lambda \in \mathbb{R}$ скалар. Тада за тензорски производ важе једнакости:

1. $\mathcal{A} \otimes (\mathcal{C} + \mathcal{D}) = \mathcal{A} \otimes \mathcal{C} + \mathcal{A} \otimes \mathcal{D}$.
2. $(\mathcal{A} + \mathcal{B}) \otimes \mathcal{C} = \mathcal{A} \otimes \mathcal{C} + \mathcal{B} \otimes \mathcal{C}$.
3. $(\lambda \mathcal{A}) \otimes \mathcal{C} = \lambda (\mathcal{A} \otimes \mathcal{C})$.
4. Ако је $\mathcal{A} = \mathcal{O}$ или $\mathcal{C} = \mathcal{O}$, онда је $\mathcal{A} \otimes \mathcal{C} = \mathcal{O}$.
5. Уколико је $\mathcal{A} \otimes \mathcal{C} = \mathcal{O}$ тада важи $\mathcal{A} = \mathcal{O}$ или $\mathcal{C} = \mathcal{O}$.
6. Када је $\mathcal{A} \otimes \mathcal{C} = \mathcal{B} \otimes \mathcal{D} \neq \mathcal{O}$, онда постоји $\alpha \in \mathbb{R},\ \alpha \neq 0$, тако да је $\mathcal{B} = \alpha \mathcal{A}$ и $\mathcal{D} = \dfrac{1}{\alpha} \mathcal{C}$

Осим оваквог начина множења, расклапање тензора у матрице отвара простор за примену матричног производа, то зовемо множење мода k.

Дефиниција 12. Нека су дати тензор $\mathcal{A} \in \mathcal{M}_{m_1 \times \cdots \times m_n}$ и матрица $M \in \mathcal{M}_{p \times m_k}$. Производ мода k је операција

$$\times_k \ : \ \mathcal{M}_{m_1 \times \cdots \times m_{k-1} \times m_k \times m_{k+1} \times \cdots \times m_n} \mapsto \mathcal{M}_{m_1 \times \cdots \times m_{k-1} \times p \times m_{k+1} \times \cdots \times m_n}$$

чији резултујући тензор има расклапање мода k :

$$(\mathcal{A} \times_k M)_{[k]} = M\, \mathcal{A}_{[k]},$$

где је израз $M\, \mathcal{A}_{[k]}$ класично множење матрица.

Приметимо да се ова операција може описати и кроз појединачне елементе тензора који учествују у изразу. Нека су $i_s \in [m_s],\ s \in [n], \quad j \in [p],\ l \in [m_k]$,

$$\mathcal{B} = [b_{i_1 \ldots i_{k-1}\, j\, i_{k+1} \ldots i_n}] = \mathcal{A} \times_k M, \qquad \mathcal{A} = [a_{i_1 \ldots i_{k-1}\, i_k\, i_{k+1} \ldots i_n}], \quad M = [\alpha_{j\, l}],$$

тада је

$$b_{i_1 \ldots i_{k-1}\, j\, i_{k+1} \ldots i_n} = \sum_{i_k = 1}^{m_k} \alpha_{j\, i_k} a_{i_1 \ldots i_n}.$$

Елемент $b_{i_1 \ldots i_{k-1}\, j\, i_{k+1} \ldots i_n}$ према томе представља скаларни производ $j-$те врсте матрице M и влакна мода i_k на позицији осталих модова $i_1 \ldots i_{k-1}\, i_{k+1} \ldots i_n$.

Специјалан случај множења мода k добијамо множењем тензора и вектора. Када је $\mathcal{A} \in \mathcal{M}_{m_1 \times \cdots \times m_n}$ и вектор $v^T \in \mathcal{M}_{1 \times m_k}$, резултујући тензор би био тензор типа $\mathcal{M}_{m_1 \times \cdots \times m_{k-1} \times 1 \times m_{k+1} \times \cdots \times m_n}$. Због конвенције брисања модова ширине 1 имамо да је производ мода k и вектора пресликавање којим се брише један мод, тј. губи једна димензија тензора:

$$\times_k \ : \ \mathcal{M}_{m_1 \times \cdots \times m_{k-1} \times m_k \times m_{k+1} \times \cdots \times m_n} \mapsto \mathcal{M}_{m_1 \times \cdots \times m_{k-1} \times m_{k+1} \times \cdots \times m_n}$$

Ова операција се дефинише кроз појединачне компоненте резултујућег тензора. Уколико уведемо ознаке $i_s \in [m_s],\ s \in [n], \quad j \in [m_k]$,

$$\mathcal{A} = [a_{i_1 \ldots i_n}], \qquad v = [\alpha_j],$$

тада је

$$(\mathcal{A} \times_k v)_{i_1 \dots i_{k-1} i_{k+1} \dots i_n} = \sum_{i_k=1}^{m_k} \alpha_{i_k} a_{i_1 \dots i_n}.$$

Елемент производа $\mathcal{A} \times_k v$ на позицији $i_1, \dots i_{k-1}\ i_{k+1} \dots i_n$ представља скаларни производ вектора v и влакна мода i_k на позицији осталих модова $i_1 \dots i_{k-1}\ i_{k+1} \dots i_n$.

Операција производа мода k са вектором снижава димензију тензора. Такве операције којима се губе неки модови у тензорима аргумената зовемо контракцијама тензора. Контракције тензора можемо да посматрамо и у општијем случају.

Дефиниција 13. Нека су дати тензори $\mathcal{A} \in \mathcal{M}_{m_1 \times \dots \times m_n \times p_1 \times \dots \times p_l}$ и $\mathcal{B} \in \mathcal{M}_{p_1 \times \dots \times p_l \times q_1 \times \dots \times q_s}$. Контракција заједничких модова $p_1, \dots, p_l$ ова два тензора је операција

$$\langle \cdot, \cdot \rangle_l \ : \ \mathcal{M}_{m_1 \times \dots \times m_n \times p_1 \times \dots \times p_l} \times \mathcal{M}_{p_1 \times \dots \times p_l \times q_1 \times \dots \times q_s} \ \mapsto \ \mathcal{M}_{m_1 \times \dots \times m_n \times q_1 \times \dots \times q_s}$$

задата покомпонентно

$$\mathcal{C} = [c_{i_1 \dots i_n k_1 \dots k_s}] = \langle \mathcal{A}, \mathcal{B} \rangle_l, \qquad \mathcal{A} = [a_{i_1 \dots i_n j_1 \dots j_l}], \quad \mathcal{B} = [b_{j_1 \dots j_l k_1 \dots k_s}],$$

$$c_{i_1 \dots i_n k_1 \dots k_s} = \sum_{j_1=1}^{p_1} \sum_{j_2=1}^{p_2} \dots \sum_{j_l=1}^{p_l} a_{i_1 \dots i_n j_1 \dots j_l} b_{j_1 \dots j_l k_1 \dots k_s}.$$

Операције овог типа битне су за редукцију димензионалности тензора. Приметимо да уколико се сви модови два тензора поклапају операција контракције производи скалар.

2.8 Ранг тензора

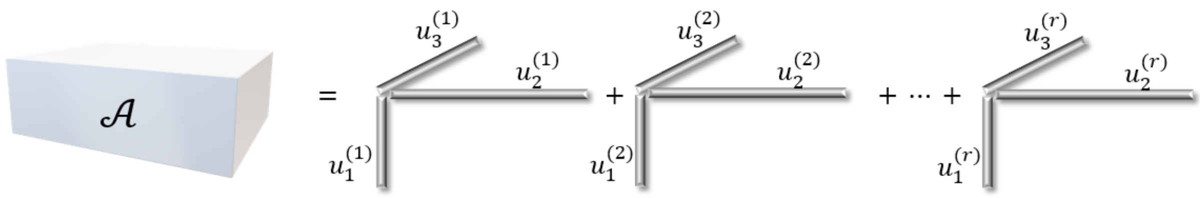
Наредна дефиниција је наставак теореме 7 за матрице.

Дефиниција 14. Тензор $\mathcal{A} \in \mathcal{M}_{m_1 \times \dots \times m_n}$, $\mathcal{A} \neq \mathcal{O}$ је ранга 1 уколико може да се прикаже као тензорски производ n вектора $v_k \in \mathbb{R}^{m_k}$, $k \in [n]$,

$$\mathcal{A} = u_1 \otimes u_2 \otimes \dots \otimes u_n \quad \iff \quad \operatorname{rang}(\mathcal{A}) = 1. \tag{2.14}$$

Тензор $\mathcal{A} \in \mathcal{M}_{m_1 \times \dots \times m_n}$, $\mathcal{A} \neq \mathcal{O}$ је ранга $r \in \mathbb{N}$ ако је то најмањи број тензора ранга 1 који у збиру дају тензор $\mathcal{A}$.

$$\mathcal{A} = \mathcal{T}_1 + \mathcal{T}_2 + \dots + \mathcal{T}_r, \quad \operatorname{rang}(\mathcal{T}_k) = 1, \ k \in [r]. \tag{2.15}$$



Слика 2.17: Илустрација ранг декомпозиције тензора димензије 3

Према договору, $\operatorname{rang}(\mathcal{O}) = 0$. Тензор ранга 1 још називамо и прост тензор, а једнакост (2.14) је његова декомпозиција. Слично, за тензор ранга r једнакост (2.15) представља његову декомпозицију - ранг декомпозиција. У литератури је још позната као CP декомпозиција (Canonical-Polyadic decomposition) или PARAFAC (Parallel Factor Analysis). Уведимо ознаке вектора у декомпозицији тензора ранга 1,

$$\mathcal{T}_k = u_1^{(k)} \otimes u_2^{(k)} \otimes \cdots \otimes u_n^{(k)}, \quad k \in [r].$$

Уколико векторе $u_i^{(k)}$ за сваки мод $i \in [n]$ сместимо у матрицу U_i по колонама,

$$U_i = \left[\begin{array}{c|c|c|c} u_i^{(1)} & u_i^{(2)} & \dots & u_i^{(r)} \end{array} \right] \in \mathcal{M}_{m_i \times r},$$

разлагање ранга r (2.15) тензора $\mathcal{A} \in \mathcal{M}_{m_1 \times \cdots \times m_n}$ тада компактно бележимо са

$$\mathcal{A} = [[U_1, U_2, \dots, U_n]].$$

Ранг тензора у многоме одступа од ранга матрице, тј. ранга дводимензионалног тензора. За матрице постоји веома ефикасан алгоритам (Гаусова елиминација) којим се утврђује њен ранг. Код тензора виших димензија аналогна процедура није позната. Такође, код матрица је познато да њихов ранг не прелази вредност мање од димензија. Са вишедимензионалним тензорима такво својство не важи. Још једна аномалија се поављује уколико дозволимо да тензори садрже комплексне бројеве. Тада ранг једног истог тензора над $\mathbb{R}$ и над $\mathbb{C}$ може да се разликује. Заправо, већина ствари у вези ранга вишедимензионалних тензора представља велику непознаницу. Због тога је ово активна област истраживања великог броја научника.

Један од начина да се делимично превазиђе проблем ранга тензора јесте његова анализа кроз матрице расклапања.

Дефиниција 15. За тензор $\mathcal{A} \in \mathcal{M}_{m_1 \times m_2 \times \cdots \times m_n}$ $k-$ранг, у ознаци rang_k, представља ранг матрице $\mathcal{A}_{[k]}$. Уколико је

$$\text{rang}_k(\mathcal{A}_{[k]}) = r_k, \quad k \in [n],$$

тада за $\mathcal{A}$ кажемо да је rang $-(r_1, r_2, \dots, r_n)$ тензор.

Глава 3

Унитарни простори

3.1 Скаларни производ и норма

Операција која уводи углове и дужине међу векторе је скаларни производ.

Дефиниција 16. Нека је V векторски простор над пољем скалара $\mathbb{R}$. Пресликавање $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ назива се реалан скаларни производ на V уколико за свака три вектора $u, v, w \in V$ и произвољан скалар $\alpha \in \mathbb{R}$, има особине:

C1: $\langle u, v \rangle = \langle v, u \rangle$;

C2: $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$;

C3: $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$;

C4: $\langle u, u \rangle > 0, u \neq \theta$.

Векторски простор снабдевен операцијом скаларног производа називамо унитаран простор.

Последица дефиниције 16 јесте да је реалан скаларни производ билинеарно пресликавање,

$$\langle \alpha u + \beta v, w \rangle = \langle \alpha u, w \rangle + \langle \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle,$$
$$\langle u, \alpha v + \beta w \rangle = \langle \alpha v + \beta w, u \rangle = \alpha \langle v, u \rangle + \beta \langle w, u \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle.$$

Избором базе у коначно димензионалном векторском простору можем да добијемо матрицу овог билинеарног пресликавања. То су матрице са веома посебним својствима. Напоменимо да није свако билинеарно пресликавање скаларни производ. Можемо да их идентификујемо управо на основу својстава придружених матрица.

Пример 3.1.1. Нека је $(b) : \{b_1, b_2, \ldots, b_n\}$ изабрана база у унитарном простору V, и $S = [s_{ij}] \in \mathcal{M}_{n \times n}$ матрица скаларног производа. С обзиром да је

$$b_1 = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \end{bmatrix}_{(b)}^T,$$
$$b_2 = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \end{bmatrix}_{(b)}^T,$$
$$\vdots$$
$$b_n = \begin{bmatrix} 0 & 0 & 0 & \ldots & 1 \end{bmatrix}_{(b)}^T,$$

матрицу билинеарног пресликавања чине скаларни производи елемената базе (b),

$$s_{ij} = \langle b_i, b_j \rangle.$$

Нека су $v, u \in V$ дати својим координатама у односу на базу (b),

$$v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}_{(b)}^T, \qquad u = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}_{(b)}^T.$$

Тада је

$$\begin{aligned}
\langle v, u \rangle &= \left\langle \sum_{i=1}^{n} x_i b_i, \sum_{j=1}^{n} y_j b_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \langle b_i, b_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j s_{ij} \\
&= \sum_{i=1}^{n} x_i \big(s_{i1} y_1 + s_{i2} y_2 + \dots + s_{in} y_n \big) \\
&= \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \dots & s_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.
\end{aligned}$$

Реалан скаларни производ је комутативна операција, па је специјално

$$s_{ij} = \langle b_i, b_j \rangle = \langle b_j, b_i \rangle = s_{ji}.$$

Због тога за матрицу S важи $S^T = S$, тј. S је симетрична матрица.

Особина С4 скаларног производа доноси још једну веома битну особину матрице скаларног производа:

$$\forall \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \neq \theta, \quad \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{n1} & s_{n2} & \dots & s_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} > 0.$$

Матрице са особином $v^T A v > 0,\ v \neq \theta$, називамо позитивно дефинитним.

На основу спроведене анализе видимо да је билинеарна функционела скаларни производ уколико је њена матрица симетрична позитивно дефинитна. Особина симетричности се лако препознаје. За позитивну дефинитност упознаћемо нешто касније критеријуме којима се проверава ова особина.

Матрице скаларног производа вектора базе су специјалан случај Грамове матрице.

Дефиниција 17. За дати скуп вектора $v_1, v_2, \dots, v_n \in V$ Грамова матрица је матрица свих скаларних производа вектора овог скупа, тј.

$$\begin{aligned}
G(v_1, v_2, \dots, v_n) &= \begin{bmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \dots & \langle v_1, v_n \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \dots & \langle v_2, v_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \dots & \langle v_n, v_n \rangle \end{bmatrix} \\
&= \big[\langle v_i, v_j \rangle \big]_{i,j=1,2,\dots n}.
\end{aligned}$$

Теорема 11. Матрица A^TA је регуларна акко $A \in \mathcal{M}_{m\times n}$ има линеарно независне колоне.

ДОКАЗ: Приметимо да је матрица A^TA квадратна реда n. Главни циљ доказа је да се покаже једнакост димензија простора слика матрица A и A^TA, тј. $\operatorname{rang}(A) = \operatorname{rang}(A^TA)$. Тада је матрица $A^TA \in \mathcal{M}_{n\times n}$ регуларна акко $\operatorname{rang}(A^TA) = n$. Како је $n = \operatorname{rang}(A) \iff$ колоне матрице A су линеарно независне, тиме би тврђење теореме било доказано.

С обзиром да је $\operatorname{rang}(A) = n - \operatorname{def}(A)$ и $\operatorname{rang}(A^TA) = n - \operatorname{def}(A^TA)$, уколико покажемо да су језгра матрица A и A^TA једнака, добићемо тражену једнакост рангова. Доказујемо најпре да је $\mathcal{N}(A) \subseteq \mathcal{N}(A^TA)$:

$$v \in \mathcal{N}(A) \iff Av = \theta \implies A^TAv = \theta \iff v \in \mathcal{N}(A^TA).$$

Следеће, показујемо $\mathcal{N}(A^TA) \subseteq \mathcal{N}(A)$:

$$v \in \mathcal{N}(A^TA) \iff A^TAv = \theta \implies v^TA^TAv = 0 \iff (Av)^T(Av) = \|Av\|^2 = 0$$

$$\iff Av = \theta \iff v \in \mathcal{N}(A). \ \Box$$

Последица 6. Када матрица A има линеарно независне колоне тада је матрица A^TA квадратна, симетрична, регуларна матрица.

Дефиниција 18. Нека је V векторски простор над пољем скалара $\mathbb{C}$. Пресликавање $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ назива се реалан скаларни производ на V уколико за свака три вектора $u, v, w \in V$ и произвољан скалар $\alpha \in \mathbb{R}$, има особине:

У1: $\langle u, v\rangle = \overline{\langle v, u\rangle}$;

У2: $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$;

У3: $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$;

У4: $\langle u, u\rangle > 0, u \neq \theta$.

Комплексан скаларни производ није билинеарна функционела јер није хомогена по другом аргументу, тј.

$$\langle u, \alpha v\rangle = \overline{\langle \alpha v, u\rangle} = \overline{\alpha\langle v, u\rangle} = \overline{\alpha}\langle u, v\rangle.$$

Ипак, и ова функционела дозвољава израчунавање кроз матричне операције. За произвољну базу $(b)\ :\ \{b_1, b_2, \ldots, b_n\}$ Грамова матрица има особину

$$G(b_1, b_2, \ldots, b_n)^H = \overline{G(b_1, b_2, \ldots, b_n)}^T = G(b_1, b_2, \ldots, b_n).$$

Норма вектора у векторском простору игра улогу апсолутне вредности међу реалним бројевима: обезбеђује неопходан садржај за мерење растојања у векторском простору. Осим свог геометријског значења дужине, норма или интензитет вектора је један број који носи пуно информација о компонентама унутар вектора.

Дефиниција 19. Нека је V векторски простор над пољем скалара $\mathbb{R}$. Пресликавање $\|\cdot\| : V \to \mathbb{R}_0^+$ назива се норма на V уколико има особине:

Н1: $\|v\| = 0 \iff v = \theta, \quad \forall v \in V$;

H2: $\|\alpha v\| = |\alpha|\,\|v\|, \quad \forall v \in V, \alpha \in \mathbb{C}$;

H3: $\|u+v\| \leq \|u\| + \|v\|, \quad \forall u, v \in V$.

Векторски простор снабдевен нормом називамо нормиран простор.

Дефиниција 20. За дате векторе $u, v \in \mathbb{R}^n$, њихово растојање је

$$d(u,v) = \|u - v\|.$$

Норма (дужина) вектора може се дефинисати преко скаларног производа,

$$\|v\| = \sqrt{\langle v, v\rangle}.$$

Тада је називамо норма индукована скаларним производом.

Две најважније неједнакости скаларног производа и норме јесу

- Шварцова[1] неједнакост[2]:
$$|\langle v, u\rangle| \leq \|v\|\,\|u\|, \tag{3.1}$$

- неједнакост троугла:
$$\|v + u\| \leq \|v\| + \|u\|. \tag{3.2}$$

Дефиниција 21. Два вектора v и u су ортогонална, у ознаци $v \perp u$, када је $\langle v, u\rangle = 0$.

Два скупа вектора S_1 и S_2 су ортогонална, у ознаци $S_1 \perp S_2$, ако је за свака два вектора

$$v \in S_1, u \in S_2 \qquad \Longrightarrow \qquad \langle v, u\rangle = 0. \tag{3.3}$$

Веома значајне базе у израчунавањима представљају ортонормиране базе векторских простора. Скаларни производ два вектора ортонормиране базе износи 0 или 1, тј. скуп вектора $(b): \ \{b_1, b_2, \ldots, b_n\}$ је ортонормиран уколико је

$$\langle b_i, b_j\rangle = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{3.4}$$

Закључујемо да је скуп вектора $v_1, v_2, \ldots, v_k \in V$ ортонормиран када је одговарајућа Грамова матрица једнака јединичној матрици $I \in \mathcal{M}_{k\times k}$.

Грамова матрица $G(b)$ вектора базе (b) биће јединична матрица уколико су вектори базе (b) ортонормирани у односу на дати скаларни производ. У том случају скларни производ вектора задатих координатама у односу на базу (b)

$$v = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}^T_{(b)}, \qquad u = \begin{bmatrix} y_1 & y_2 & \ldots & y_n \end{bmatrix}^T_{(b)},$$

рачунамо формулом за стандардни скаларни производ,

$$\langle v, u\rangle = v^T u = u^T v = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \text{ у реалном случају,} \tag{3.5}$$

$$\langle v, u\rangle = u^H v = \overline{v^H u} = \begin{bmatrix} \overline{y_1} & \overline{y_1} & \ldots & \overline{y_n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \text{ у комплексном случају.} \tag{3.6}$$

[1] Karl Hermann Amandus Schwarz, (1843–1921), немачки математичар

[2] у литератури позната и као Коши-Шварц-Буњаковски неједнакост

За њега користимо једноставнију ознаку операције, тј. уместо $\langle v, u\rangle$ писаћемо $v \cdot u$ да специјално нагласимо овај скаларни производ.

Операција транспоновања и коњугованог транспоновања матрица 'слаже' се са операцијама множења над матрицама као и остале операције степене нотације.

$$\begin{array}{ccc} (A+B)^T = A^T + B^T & (A+B)^H = A^H + B^H & (A^T)^{-1} = (A^{-1})^T \\ (\lambda A)^T = \lambda A^T & (\lambda A)^H = \overline{\lambda} A^H & (\lambda A)^{-1} = \dfrac{1}{\lambda} A^{-1} \\ (AB)^T = B^T A^T & (AB)^H = B^H A^H & (AB)^{-1} = B^{-1} A^{-1} \end{array}$$

Захваљујући томе, матрична нотација стандардног скаларног производа је у потпуности сагласна са његовим особинама из дефиниција 16 и 18.

Пример 3.1.2. Нека су колоне матрице A вектори $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$,

$$A = \left[\begin{array}{c|c|c|c} v_1 & v_2 & \ldots & v_k \end{array}\right].$$

Ортогоналност вектора $u \in \mathbb{R}^n$ на векторе $v_1, v_2, \ldots, v_k$ проверавамо матричним изразом:

$$u \perp \{v_1, v_2, \ldots, v_k\} \quad \iff \quad A^T u = \theta. \tag{3.7}$$

У случају да су $u, v_1, v_2, \ldots, v_k \in \mathbb{C}^n$, еквивалентан матрични израз гласи

$$u \perp \{v_1, v_2, \ldots, v_k\} \quad \iff \quad A^H u = \theta. \tag{3.8}$$

У наставку бавимо се реалним векторима. Претпоставимо да је у векторском простору V димензије n, дата ортонормирана база

$$(b): \ \{b_1, b_2, \ldots, b_n\},$$

тј. важи

$$b_i \cdot b_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Координате произвољног вектора v у односу на базу (b)

$$v = \lambda_1 b_1 + \lambda_2 b_2 + \cdots + \lambda_n b_n$$

се називају Фуријеове[3] координате и важи да је $\lambda_k = v \cdot b_k$, тј.

$$v = (v \cdot b_1) b_1 + (v \cdot b_2) b_2 + \cdots + (v \cdot b_n) b_n. \tag{3.9}$$

Уколико уведемо ознаку

$$V = \left[\begin{array}{c|c|c|c} b_1 & b_2 & \ldots & b_n \end{array}\right], \quad \Lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix},$$

[3]Jean-Baptiste Joseph Fourier (1768– 1830), француски математичар

Фуријеове координате вектора v, $\lambda_k = v \cdot b_k$, можемо истовремено да одредимо матричним производом $\Lambda = V^T v$ у реалном случају, и $\Lambda = V^H v$ у случају комплексних вектора. Уколико уведемо ознаке $v_k = (v \cdot b_k) b_k$ добијамо ортогонално разлагање вектора v преко његових ортогоналних пројекција на координатне осе P_k

$$v = v_1 + v_2 + \cdots + v_n.$$

Тада је једнакошћу

$$\begin{aligned} \|v\|^2 &= \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 = \|v_1\|^2 + \|v_2\|^2 + \cdots + \|v_n\|^2, \\ &= (v \cdot b_1)^2 + (v \cdot b_2)^2 + \cdots + (v \cdot b_n)^2, \end{aligned} \tag{3.10}$$

дато уопштење Питагорине теореме.

Угао α између два реална не-нула вектора v и u дефинишемо релацијом

$$\cos\alpha = \frac{\langle v, u \rangle}{\|v\| \, \|u\|}. \tag{3.11}$$

Која неједнакост омогућава ову дефиницију?

Косинус угла између вектора $v, u \in \mathbb{R}^n$ прилично добро може да опише сличности у серијама вредности које носе координате ова два вектора. Таква анализа садржаја вектора се нарочито корисно показала у ситуацијама када су све компоненте вектора ненегативни бројеви.

Приметимо да је $\cos\alpha > 0$ када је $\alpha \in (-\pi/2, \pi/2)$. Вредност $\cos\alpha = 1$ означава колинеарност два вектора и њихово исто усмерење, тј. најмање угловно одступање између вектора $\alpha = 0$. Вектори су супротно усмерени када је $\cos\alpha = -1$, а међусобно су ортогонални за $\cos\alpha = 0$.

Нумерички посматрано $\cos\alpha = \dfrac{v \cdot u}{\|v\| \, \|u\|} > 0$, означава да израз

$$v \cdot u = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n, \qquad v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T, \; u = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T \in \mathbb{R}^n,$$

садржи више позитивних сабирака $x_i y_i$, тј. координата истог знака. Друга опција је да постоји мањи број кордината истог знака које су доминантне по величини, тј. постојање екстрема. Због могућности описивања сличности у понашању вредности компоненти вектора, у реалним проблемима инжењерства постоји веома честа потреба за мерењем тзв. косинусне сличности вектора:

$$sim(v, u) = \frac{v \cdot u}{\|v\| \, \|u\|} = \frac{v}{\|v\|} \cdot \frac{u}{\|u\|}. \tag{3.12}$$

Једнакост (3.12) описује косинусну сличност као сличност између нормираних вектора $\dfrac{v}{\|v\|}$ и $\dfrac{u}{\|u\|}$. Због тога ова мера указује на сличност у "оријентацији"вектора, а не на блискост у смислу њиховог растојања (норме). Косинусна сличност није норма, што потврђује и њено име.

Напомена 4. Пораст вредности растојања вектора указује на њихову различитост, док повећање вредности косинусне сличности указује на блискост тих вектора по правцима. Због тога се веома често уместо косинусне сличности користи косинусна удаљеност за описивање различитости вектора

$$D_C(v, u) = 1 - sim(v, u).$$

За стандардни скаларни производ (3.5), одговарајућу норму зовемо Еуклидова норма. Експлицитни израз за Еуклидову норму преко координата вектора $v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T \in \mathbb{R}^n$ гласи:

$$\|v\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Најзначајније норме вектора су тзв. $p-$норме или L_p норме.

За реалан или комплексан вектор $v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ је

L_p: $\|v\|_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p}$, $p \geq 1$, чији су специјални случајеви:

L_1: $\|v\|_1 = \sum_{k=1}^{n} |x_k|$, позната још и као Менхетн норма или такси норма,

L_2: $\|v\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2}$, Еуклидова норма,

L_∞: $\|v\|_\infty = \max_k \{|x_k|\}$, макс норма или норма Чебишева.

На основу самих дефиниција L_p норми јасно је да важи

$$\|v\|_\infty \leq \|v\|_2 \leq \|v\|_1, \qquad \forall v \in \mathbb{R}^n, \quad \text{тј.}$$
$$\max_k \{|x_k|\} \leq \sqrt{\sum_{k=1}^{n} |x_k|^2} \leq \sum_{k=1}^{n} |x_k|.$$

Свака од наведених норми омогућава дефиницију растојања над векторима које називамо метрика Минковског или $p-$ метрика

$$d_p(v, u) = \|v - u\|_p. \tag{3.13}$$

У општем случају, норма не мора да буде индукована скаларним производом. Препознавање индукованих норми омогућава једнакост паралелограма

$$2\|u\|^2 + 2\|v\|^2 = \|u + v\|^2 + \|u - v\|^2. \tag{3.14}$$

Уколико норма задовољава (3.14) тада је она индукована реалним скаларним производом

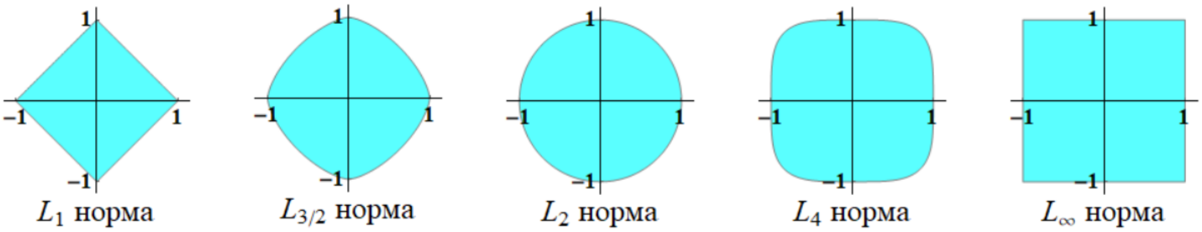
$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4}.$$

Од наведених L_p норми једино Еуклидова норма L_2 представља норму индуковану скаларним производом, подразумевајући да је димензија простора $n \geq 2$.

Пример 3.1.3. За наредне теме, као и за теорију грешака, појам растојања између вектора је веома битан. Повремено је корисно посматрати норму као меру сличности информација које вектори носе. Да бисмо боље разумели разлике у геометирији коју производе L_p−норме $p \geq 1$, приказаћемо у $2D$ и $3D$ геометрији јединичне сфере у односу на L_1, L_2 и L_∞− три норме најчешће у практичној употреби. Приказане су и $L_{3/2}$, L_4 норме као прелазни облици између ове три најбитније норме.

За $2D$ примере користићемо ознаку координата вектора $v = \begin{bmatrix} x & y \end{bmatrix}^T$.

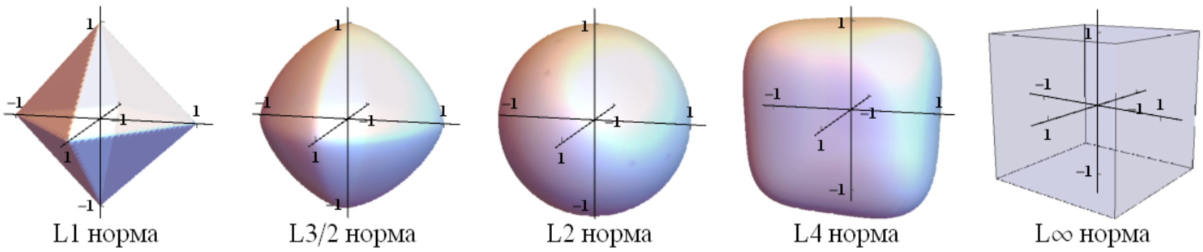
L_1 јединична сфера:	$S_1 = \{v \mid \|v\|_1 = 1\} = \{v \mid \ \|x\| + \|y\| = 1\}$
$L_{3/2}$ јединична сфера:	$S_{3/2} = \{v \mid \|v\|_{3/2} = 1\} = \{v \mid \ \left(\|x\|^{3/2} + \|y\|^{3/2}\right)^{2/3} = 1\}$
L_2 јединична сфера:	$S_2 = \{v \mid \|v\|_2 = 1\} = \{v \mid \ \left(x^2 + y^2\right)^{1/2} = 1\}$
L_4 јединична сфера:	$S_4 = \{v \mid \|v\|_4 = 1\} = \{v \mid \ \left(\|x\|^4 + \|y\|^4\right)^{1/4} = 1\}$
L_∞ јединична сфера:	$S_\infty = \{v \mid \|v\|_\infty = 1\} = \{v \mid \ \max\{\|x\|, \|y\|\} = 1\}$



Слика 3.1: Јединичне сфере у $2D$

За $3D$ примере користићемо ознаку за координате вектора $v = \begin{bmatrix} x & y & z \end{bmatrix}^T$.

L_1 јединична сфера:	$S_1 = \{v \mid \|v\|_1 = 1\} = \{v \mid \ \|x\| + \|y\| + \|z\| = 1\}$
$L_{3/2}$ јединична сфера:	$S_{3/2} = \{v \mid \|v\|_{3/2} = 1\} = \{v \mid \ \left(\|x\|^{3/2}+\|y\|^{3/2}+\|z\|^{3/2}\right)^{2/3} = 1\}$
L_2 јединична сфера:	$S_2 = \{v \mid \|v\|_2 = 1\} = \{v \mid \ \left(x^2 + y^2 + z^2\right)^{1/2} = 1\}$
L_4 јединична сфера:	$S_4 = \{v \mid \|v\|_4 = 1\} = \{v \mid \ \left(\|x\|^4 + \|y\|^4 + \|z\|^4\right)^{1/4} = 1\}$
L_∞ јединична сфера:	$S_\infty = \{v \mid \|v\|_\infty = 1\} = \{v \mid \ \max\{\|x\|, \|y\|, \|z\|\} = 1\}$



Слика 3.2: Јединичне сфере у $3D$

Да ли можете да претпоставите како би изгледали скупови S_p за $p \in (0,1)$?

Пример 3.1.4. Једна од веома важних пдеудо-норми је L_0 норма којом се враћа број ненула елемената вектора. На тај начин L_0 представља неку врсту густине, тј. ретке поседнутости, података унутар вектора. За меру L_0 кажемо да је псеудо-норма јер не задовољава услов хомогености норме H2. Норма која је по карактеристикама најближа L_0 псеудо-норми је L_1 за велике димензије вектора. Показано је да минимизација по L_1 норми максимизује ретку поседнутост вектора, [17].

3.2 Афина геометрија

Множење вектора v скаларом $\lambda \in (0,1]$ даје вектор $u = \lambda v$ истог правца и смера али не веће дужине у односу на вектор v,

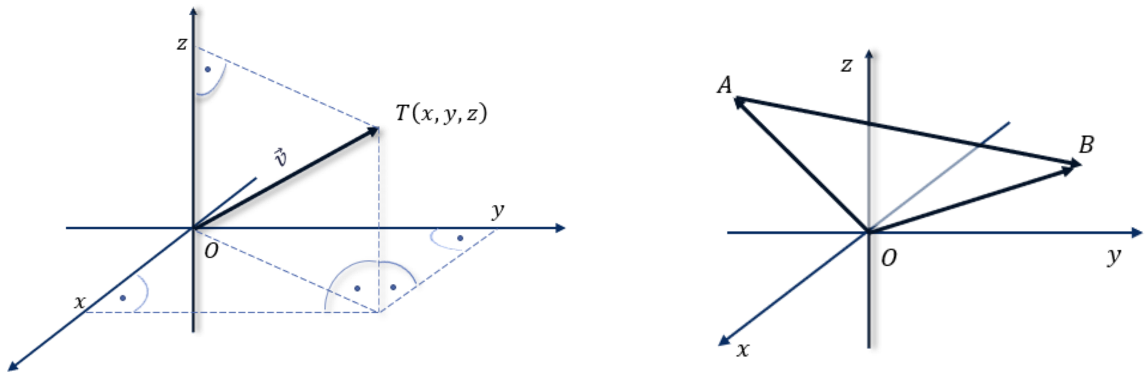
$$\cos\angle(v,\lambda v) = \frac{v\cdot(\lambda v)}{\|v\|\,\|\lambda v\|} = \frac{\lambda(v\cdot v)}{|\lambda|\,\|v\|^2} = \frac{\lambda\|v\|^2}{\lambda\|v\|^2} = 1, \tag{3.15}$$

$$\|\lambda v\| = |\lambda|\,\|v\| \overset{\lambda\geq 0}{=} \lambda\|v\| \overset{\lambda\leq 1}{\leq} \|v\|. \tag{3.16}$$

Ова особина, као и неједнакости троугла

$$\big|\|v\| - \|u\|\big| \leq \|v+u\| \leq \|v\| + \|u\|,$$

користе се за описивање унутрашњости и спољашњости скупова вектора. У томе важну улогу имају специјалне линеарне комбинације. Да бисмо развијали геометријску интуицију ових појмова, модел који користимо је аналитичка геометрија у $2D$ и $3D$ (изучавана у оквиру предмета Математика 1). Да бисмо назначили рад у $2D$ и $3D$ обично користимо ознаку $\vec{v}$ за векторе. Аналитичка геометрија поред појмова вектора и базе векторског простора користи појам тачке и координатног система. Сваки вектор векторског простора са почетком у координатном почетку O својим



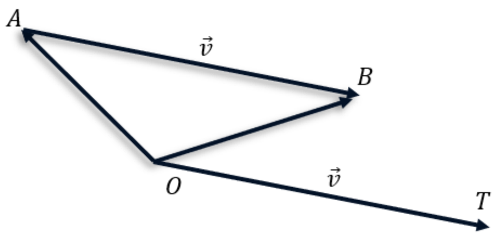
Слика 3.3: Афина геометрија $3D$

крајем одређује јединствену тачку. Због тога вршимо поистовећивање вектора са вектором положаја тачака. Такве векторе зовемо везаним. Другим речима, док радимо искључиво са везаним векторима, моделирамо (скицирамо) векторски простор. Када уведемо слободне векторе и тачке, тада моделирамо афине просторе.

Додавањем тачака векторском простору ширимо полазни скуп објеката, самим тим и операција и пресликавања над њима (афине трансформације). Векторски простор проширен скупом тачака и операцијом тачка+вектор (транслација) представља афини простор. Особине операција у афином простору су природна последица операција над векторима:

$$T + \theta = T, \quad (T+\vec{u}) + \vec{v} = T + (\vec{u}+\vec{v}),$$
$$\lambda T = \lambda\overrightarrow{OT}.$$

За произвољне две тачке A и B постоји јединствен вектор v такав да је $B = A + \vec{v}$. Вектор $\vec{v}$ обично зовемо слободан вектор и означавамо га још и са $\overrightarrow{AB}$. Координате слободног вектора $\vec{v} = \overrightarrow{AB}$ могу се одредити као разлика координата вектора положаја крајње и почетне тачке, слика 3.4.



$$
\begin{aligned}
\vec{v} &= \overrightarrow{OT} \equiv T \\
&= \overrightarrow{OB} - \overrightarrow{OA} \\
&= B - A \\
B &= A + \vec{v}
\end{aligned}
$$

Слика 3.4: Слободни и везани вектор $\vec{v}$

Дефиниција 22. Нека је $P \subseteq \mathbb{R}^n$ векторски потпростор и $v \in \mathbb{R}^n$ произвољан вектор (тачка). Транслацију

$$L = v + P = \{v + u \mid u \in P\}$$

зовемо афини потпростор простора $\mathbb{R}^n$. Димензија афиног потпростора L је димензија векторског потпростора P.

Афини потпростор је векторски потпростор акко садржи нула-вектор θ. Аналитичка геометрија са тачком, правом и равни је модел геометрије афиних простора. Тачка је тако афини потпростор димензије нула, права је димензије 1 и раван је афини потпростор димензије 2.

Операције над векторима (линеарне комбинације) су средство кретања кроз простор. Њима можемо описивати неке карактеристичне скупове простора. Најзначајнији подскупови су тзв. геометријске примитиве. У наставку одељка подсетићемо се неких геометријских примитива и видети како се оне могу описати линеарним комбинацијама.

Пример 3.2.1. Једначина праве кроз две тачке $A \equiv \vec{v}$ и $B \equiv \vec{u}$, са вектором положаја текуће тачке $X \equiv \vec{r}$, гласи:

$$
\begin{array}{lll}
 & \overrightarrow{AX} = \lambda \overrightarrow{AB}\ , & \lambda \in \mathbb{R}, \\
\iff & \vec{r} - \vec{v} = \lambda(\vec{u} - \vec{v})\ , & \lambda \in \mathbb{R} \\
\iff & \vec{r} = \vec{v} + \lambda(\vec{u} - \vec{v})\ , & \lambda \in \mathbb{R} \\
\iff & X = A + \lambda \overrightarrow{AB}\ , & \lambda \in \mathbb{R}.
\end{array}
$$

Добијена једначина праве сугерише да се ради о једнодимезионалном (афином) објекту

$$\vec{r} \in \vec{v} + \mathcal{L}(\vec{u} - \vec{v}).$$

Димензионалност афиног потпростора можемо да пратимо и преко броја слободних параметара којима је одговарајући скуп описан. С обзиром да се све тачке са праве могу описати преко једног слободног параметра $\lambda \in \mathbb{R}$, још једном потврђујемо да се ради о једнодимензионалном афином потпростору.

Наредна теорема потврђује геометријска својства афиних потпростора које имају тачка, права и раван у $3D$.

Теорема 12. Нека је $L \subseteq \mathbb{R}^n$ подскуп афиног простора. L је афини потпростор акко за произвољне две тачке $A, B \in L$ важи да права која пролази кроз тачке A и B цела лежи у L.

ДОКАЗ: $\boxed{\Longrightarrow:}$ Нека је $L = v + P$, где је $P \subseteq \mathbb{R}^n$ векторски потпростор. С обзиром да $A, B \in L$ то постоје вектори $u_A, u_B \in P$ за које је $A = v + u_A$ и $B = v + u_B$. Покажимо да за произвољну вредност $\lambda \in \mathbb{R}$ тачка $X = A + \lambda\overrightarrow{AB}$ такође припада афином потпростору L. То је еквивалентно тврђењу да постоји вектор $u \in P$ за који је $X = v + u$.

$$X = A + \lambda\overrightarrow{AB} = v + u_A + \lambda\big((v + u_B) - (v + u_A)\big) = v + \big(u_A + \lambda(u_B - u_A)\big).$$

С обзиром да вектор $u_A + \lambda(u_B - u_A) = (1-\lambda)u_A + \lambda u_B$ представља линеарну комбинацију вектора векторског потпростора P, на основу теореме 1 закључујемо да важи $u_A + \lambda(u_B - u_A) = u \in P$. Због тога је $X = v + u \in L$.

$\boxed{\Longleftarrow:}$ Нека је $A \equiv v \in L$ произвољна тачка. Означимо скуп

$$P = L - v = \{u - v \mid u \in L\} \quad \Longrightarrow \quad L = v + P.$$

Показаћемо да је P векторски потпростор коришћењем теореме 1. За два вектора скупа P имамо

$$\begin{aligned}
w_1, w_2 \in P \quad &\Longrightarrow \quad v, v + w_1, v + w_2 \in L \\
&\Longrightarrow \begin{cases} \lambda(v + w_1) + (1-\lambda)v = v + \lambda w_1 \in L, \ \lambda \in \mathbb{R} \\ \mu(v + w_2) + (1-\mu)v = v + \mu w_2 \in L, \ \mu \in \mathbb{R} \end{cases} \\
&\Longrightarrow \quad \theta(v + \lambda w_1) + (1-\theta)(v + \mu w_2) = \theta\lambda w_1 + (1-\theta)\mu w_2 + v \in L, \ \theta \in \mathbb{R} \\
&\overset{\theta=\frac{1}{2}}{\Longrightarrow} \quad \frac{\lambda}{2}w_1 + \frac{\mu}{2}w_2 + v \in L \quad \Longrightarrow \quad \frac{\lambda}{2}w_1 + \frac{\mu}{2}w_2 \in P.
\end{aligned}$$

С обзиром да λ и μ могу да узму било коју реалну вредност закључујемо да је скуп P заиста затворен за линеарне комбинације. $\square$

Теорема 13. Скуп $L \subseteq \mathbb{R}^n$ је афини потпростор акко је скуп решења неког система линеарних једначина.

ДОКАЗ: $\boxed{\Longrightarrow:}$ Нека је $L = v + P \subseteq \mathbb{R}^n$ афини потпростор. Нека матрица $A \in \mathcal{M}_{n\times n}$ има језгро $\mathcal{N}(A) = P$ и означимо вектор $Av = b$. Тада L представља скуп решења R_S система линеарних једначина $Ax = b$. Заиста, v је једно решење овог система. Уколико је $v_1 \in R_S$ још једно решење, тада је

$$A(v_1 - v) = b - b = \theta \quad \Longrightarrow \quad v_1 - v \in \mathcal{N}(A) \quad \Longleftrightarrow \quad v_1 \in v + P = L.$$

Слично, за $u \in P = \mathcal{N}(A)$ важи

$$A(v + u) = Av + Au = b + \theta = b \quad \Longrightarrow \quad v + u \in R_S.$$

$\boxed{\Longleftarrow:}$ Показали смо да је $R_S = v + \mathcal{N}(A)$ што даје и доказ теореме 13 у супротном смеру. $\square$

Последица 7. Пресек два афина потпростора је афини потпростор.

Дефиниција 23. Нека су $v_1, v_2, \ldots, v_k$ елементи афиног простора $\mathbb{R}^n$ и $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$ скалари. Линеарна комбинација

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k, \qquad \alpha_1 + \alpha_2 + \cdots + \alpha_k = 1$$

је афина комбинација.

Дефиниција 24. За скуп $C \subset \mathbb{R}^n$ афино затварање представља скуп свих афиних комбинација коначно много елемената скупа C,

$$\operatorname{aff}(C) = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \mid v_1, v_2, \ldots, v_k \in C,\ \alpha_1 + \alpha_2 + \cdots + \alpha_k = 1,\ \alpha_i \in \mathbb{R}\}.$$

Наредним тврђењем показујемо да афино затварање у афином простору игра улогу линеала у векторским просторима.

Теорема 14. Афино затварање скупа $C \subset \mathbb{R}^n$ је најмањи афини потпростор који садржи скуп C.

ДОКАЗ: Покажимо прво да је $\operatorname{aff}(C)$ афини потпростор. На основу теореме 12 довољно је да покажемо да је права кроз произвољне две тачке скупа $\operatorname{aff}(C)$ цела садржана у овом скупу. Због тога означимо са A и B произвољне две тачке из $\operatorname{aff}(C)$. Тада је

$$\begin{aligned}
A \in \operatorname{aff}(C) \iff \quad & A = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k,\ v_1, v_2, \ldots, v_k \in C,\\
& \alpha_1 + \alpha_2 + \cdots + \alpha_k = 1,\\
B \in \operatorname{aff}(C) \iff \quad & B = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_s u_s,\ u_1, u_2, \ldots, u_s \in C,\\
& \beta_1 + \beta_2 + \cdots + \beta_s = 1,\\
\implies \lambda A + (1-\lambda)B &= \lambda(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) + (1-\lambda)(\beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_s u_s)\\
& v_1, v_2, \ldots, v_k, u_1, u_2, \ldots, u_s \in C,\\
& \lambda(\alpha_1 + \alpha_2 + \cdots + \alpha_k) + (1-\lambda)(\beta_1 + \beta_2 + \cdots + \beta_s) = 1,\\
& \implies \quad \lambda A + (1-\lambda)B \in \operatorname{aff}(C).
\end{aligned}$$

Остаје још да се покаже да је $\operatorname{aff}(C)$ најмањи афини потпростор који садржи скуп C. Другим речима, за произвољни афини потпростор $L \subseteq \mathbb{R}^n$ показаћемо импликацију

$$C \subseteq L \quad \implies \quad \operatorname{aff}(C) \subseteq L.$$

Нека је $u \in \operatorname{aff}(C)$ произвољан вектор. Тада за неке векторе $v_1, v_2, \ldots, v_k \in C$ важи

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k, \quad \alpha_1 + \alpha_2 + \cdots + \alpha_k = 1.$$

С обзиром да је $v_1 \in C \subseteq L$, тада је $P = L - v_1$ векторски потпростор. Имајући у виду $v_2, \ldots, v_k \in C \subseteq L$, то сви вектори $v_2 - v_1, \ldots, v_k - v_1 \in P$. Приметимо да је

$$u \in L \quad \iff \quad u - v_1 \in P.$$

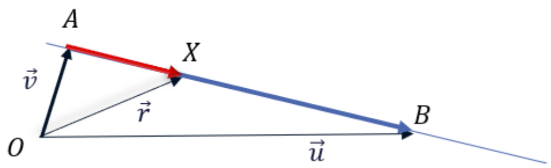
Због тога покажимо да $u - v_1 \in P$.

$$\begin{aligned}
u - v_1 &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k - v_1\\
&= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k - (\alpha_1 + \alpha_2 + \cdots + \alpha_k) v_1\\
&= \alpha_2(v_2 - v_1) + \cdots + \alpha_k(v_k - v_1) \in P. \quad \square
\end{aligned}$$

У наставку бавићемо се специјалном врстом афиних комбинација. Као увод у причу користимо пример из $3D$ геометрије.

Пример 3.2.2. Уколико желимо да опишемо само тачке са дужи AB потребно је ограничити опсег параметра λ у опису тачака са праве p. Наиме, све тачке X дужи AB образују векторе $\overrightarrow{AX}$ истог смера као $\overrightarrow{AB}$ али не веће дужине од $\|\overrightarrow{AB}\|$. Дакле, за њих важи:

$$\overrightarrow{AX} = \lambda\overrightarrow{AB}, \quad \text{за } \lambda \in [0, 1].$$



Слика 3.5: Тачке дужи AB

Приметимо да на основу (3.15) и (3.16) то значи исти правац и смер вектора $\overrightarrow{AX}$ и $\overrightarrow{AB}$, и одговарајући однос њихових дужина

$$\frac{\|\overrightarrow{AX}\|}{\|\overrightarrow{AB}\|} = \lambda. \tag{3.17}$$

Због тога су све тачке X дужи AB дате са

$$\begin{aligned} X - A = \lambda\overrightarrow{AB},\ \lambda \in [0,1],\ \text{тј.}& \\ X = A + \lambda(B - A) = (1-\lambda)A + \lambda B, &\quad \lambda \in [0,1], \\ \iff \vec{r} = (1-\lambda)\vec{v} + \lambda\vec{u}, &\quad \lambda \in [0,1]. \end{aligned}$$

За $\lambda = 0$ тачка X поклапа се са A, за $\lambda = 1$ је $X = B$, док све остале вредности $\lambda \in (0, 1)$ описују тачке између A и B. Нпр., на основу (3.17), X је средиште (центар, тежиште) T дужи AB за $\lambda = \frac{1}{2}$, тј.

$$T = \frac{1}{2}\vec{v} + \frac{1}{2}\vec{u} = \frac{\vec{v}+\vec{u}}{2}.$$

Приметимо да за коефицијенте (тежине) линеарне комбинације

$$\vec{r} = (1-\lambda)\vec{v} + \lambda\vec{u},\ \lambda \in [0,1] \tag{3.18}$$

важи:

$$0 \le \lambda, 1-\lambda \le 1 \quad \text{и} \quad \lambda + 1 - \lambda = 1.$$

Другим речима, све афине комбинације облика

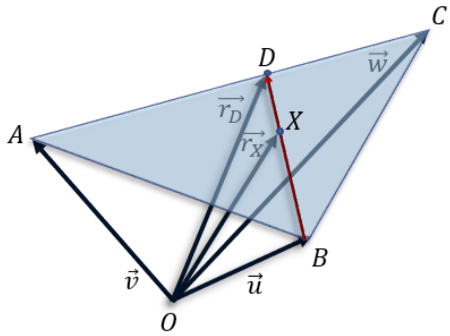
$$\alpha\vec{v} + \beta\vec{u}, \quad 0 \le \alpha, \beta, \quad \alpha + \beta = 1,$$

описују тачке дужи AB.

Уколико упоредимо дужине дужи AX и XB, тај однос биће једнак односу одговарајућих норми вектора:

$$\frac{|AX|}{|XB|} = \frac{\|\overrightarrow{AX}\|}{\|\overrightarrow{XB}\|} = \frac{\|\overrightarrow{AX}\|}{\|\overrightarrow{AB}\| - \|\overrightarrow{AX}\|} = \frac{\frac{\|\overrightarrow{AX}\|}{\|\overrightarrow{AB}\|}}{1 - \frac{\|\overrightarrow{AX}\|}{\|\overrightarrow{AB}\|}} \overset{(3.17)}{=} \frac{\lambda}{1-\lambda}. \tag{3.19}$$

Пример 3.2.3. Посматрамо сада три вектора положаја $A \equiv \vec{v}$, $B \equiv \vec{u}$ и $C \equiv \vec{w}$. Они представљају темена троугла ΔABC. Користећи претходно разматрање за дужи можемо описивати тачке $D \equiv \vec{r}_D$ са нпр. странице AC или тачке $X \equiv \vec{r}_X$ унутар троугла ΔABC.



На основу (3.18) знамо да је

$$\begin{aligned} D &= (1-\lambda)A + \lambda C, \ \lambda \in [0,1] \\ &= (1-\lambda)A + 0\,B + \lambda C, \ \lambda \in [0,1] \\ \iff \vec{r}_D &= (1-\lambda)\vec{v} + 0 \cdot \vec{u} + \lambda\vec{w} \end{aligned}$$

Добили смо још једну линеарну комбинацију, сада три вектора, чији су коефицијенти из сегмента $[0,1]$ и збир коефицијената линеарне комбинације је 1.

$$1 - \lambda + 0 + \lambda = 1$$

Слично, за произвољну унутрашњу тачку $X \equiv \vec{r}_X$ троугла ΔABC, с обзиром да је тачка са дужи BD, на основу (3.18) важи:

$$\begin{aligned} X &= \mu B + (1-\mu)D, && \mu \in [0,1], \\ &= \mu B + (1-\mu)\big((1-\lambda)A + \lambda C\big), && \lambda, \mu \in [0,1] \\ &= (1-\mu)(1-\lambda)A + \mu B + (1-\mu)\lambda C, && \lambda, \mu \in [0,1], \\ \iff \quad \vec{r}_X &= (1-\mu)(1-\lambda)\vec{v} + \mu\vec{u} + (1-\mu)\lambda\vec{w}, && \lambda, \mu \in [0,1]. \end{aligned}$$

За коефицијенте и ове линеарне комбинације важе особине да су ненегативни и у збиру дају 1 :

$$0 \le (1-\mu)(1-\lambda),\ \mu, (1-\mu)\ \lambda \le 1 \qquad \text{и} \qquad (1-\mu)(1-\lambda) + \mu + (1-\mu)\lambda = 1.$$

Закључујемо да линеарне комбинације облика

$$\alpha\vec{v} + \beta\vec{u} + \gamma\vec{w}, \qquad 0 \le \alpha, \beta, \gamma, \quad \alpha + \beta + \gamma = 1,$$

описују све тачке фигуре ΔABC.

Уколико је D средиште дужи AC, за њега важи $\vec{r}_D = \frac{1}{2}\vec{v} + \frac{1}{2}\vec{u}$. Тежиште T троугла ΔABC налази се на тежишној линији BD, на трећини њене дужине од средишта D. На основу (3.19) онда је

$$\vec{r}_T = \frac{2}{3}\vec{r}_D + \frac{1}{3}\vec{u} = \frac{2}{3}\left(\frac{1}{2}\vec{v} + \frac{1}{2}\vec{w}\right) + \frac{1}{3}\vec{u} = \frac{1}{3}\vec{v} + \frac{1}{3}\vec{u} + \frac{1}{3}\vec{w} = \frac{\vec{v} + \vec{u} + \vec{w}}{3}.$$

Претходни примери дају повод за увођење наредних дефиниција. Сваки вектор задат листом од n бројева $\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$, посматраћемо и као тачку (вектор положаја тачке) одговарајућег $n-$димензионалног афиног простора.

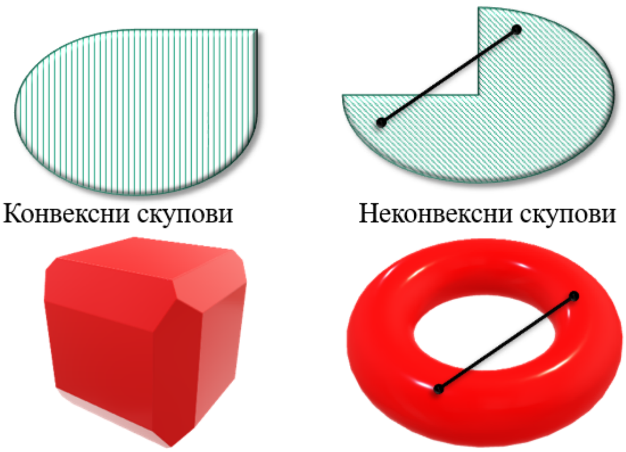
Дефиниција 25. Скуп вектора (тачака) S (афиног простора) је конвексан уколико за свака два вектора (тачке) скупа S дуж која их спаја цела припада скупу S, тј.

$$\forall X, Y \in S \text{ и } \lambda \in [0,1], \quad (1-\lambda)X + \lambda Y \in S.$$

Последица 8. Афини потпростори су конвексни скупови.

Један од термина који се користи за описивање спајања тачака неког скупа дужима кроз тај скуп, јесте да су такве тачке видљиве једна другој. У том контексту, скуп је конвексан ако су било које две тачке скупа видљиве једна другој.

Конвексне скупове коначно димензионалног векторског простора замишљамо по угледу на њихове $2D$ и $3D$ моделе, слика 3.6.



Слика 3.6: Примери конвексних и неконвексних скупова у $2D$ и $3D$

Теорема 15. Пресек два конвексна скупа је конвексан скуп.

ДОКАЗ: Нека су A и B два конвексна скупа и $C = A \cap B$. Покажимо да је и C конвексан, тј. да важи

$$a, b \in C \quad \Longrightarrow \quad (\forall \lambda \in [0,1])\ \lambda a + (1-\lambda) b \in C.$$

$$\begin{aligned} a, b \in C \iff a, b \in A \cap B \ &\iff \ (a \in A \cap B \wedge b \in A \cap B) \\ &\Longrightarrow (\forall \lambda \in [0,1]) \big(\lambda a + (1-\lambda) b \in A \wedge \lambda a + (1-\lambda) b \in B\big) \\ &\iff (\forall \lambda \in [0,1]) \big(\lambda a + (1-\lambda) b \in A \cap B = C\big). \quad \square \end{aligned}$$

На сличан начин може се показати да је пресек произвољне фамилије конвексних скупова поново конвексан скуп.

Теорема 16. Нека су $A, B \subset \mathbb{R}^n$ конвексни скупови. Тада је њихова сума

$$A + B = \{u + v \mid u \in A,\ v \in B\},$$

такође конвексан скуп.

ДОКАЗ: Нека су $u, v \in A + B$ произвољни вектори, тј. постоје $a_u, a_v \in A$ и $b_u, b_v \in B$ за које је

$$u = a_u + b_u, \qquad v = a_v + b_v.$$

Покажимо да конвексна комбинација $\lambda u + (1-\lambda) v,\ \lambda \in [0,1]$ вектора u и v такође припада скупу $A + B$, тј. потребно је да пронађемо векторе $a \in A$ и $b \in B$ за које је $\lambda u + (1-\lambda) v = a + b$.

$$\begin{aligned} \lambda u + (1-\lambda) v &= \lambda(a_u + b_u) + (1-\lambda)(a_v + b_v) \\ &= \big(\lambda a_u + (1-\lambda) a_v\big) + \big(\lambda b_u + (1-\lambda) b_v\big). \end{aligned}$$

Означимо $a = \lambda a_u + (1-\lambda)a_v$. Вектор a је добијен конвексном комбинацијом два вектора a_u и a_v конвексног скупа A. Због тога је $a \in A$. Аналогно закључујемо за вектор $b = \lambda b_u + (1-\lambda)b_v \in B$. Следи да је заиста

$$\lambda u + (1-\lambda)v = a + b \in A + B. \qquad \square$$

Последица 9. Транслација конвексног скупа је конвексан скуп.

Дефиниција 26. Конвексна комбинација вектора $v_1, v_2, \dots, v_k$ је линеарна комбинација

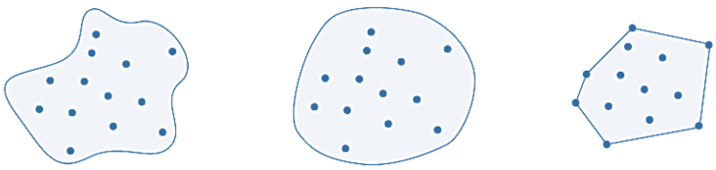
$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k$$

у којој су сви скалари ненегативни ($\lambda_i \geq 0$) и важи да је $\lambda_1 + \lambda_2 + \dots + \lambda_k = 1$.

Дефиниција 27. Нека је $S \subset \mathbb{R}^n$ непразан скуп. Конвексно затварање скупа S је скуп свих конвексних комбинација вектора овог скупа,

$$\operatorname{conv}(S) = \{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k \mid v_1, v_2, \dots, v_k \in S, \quad \lambda_i \geq 0,\ \lambda_1 + \lambda_2 + \dots + \lambda_k = 1\}.$$

Дефиниција 28. За две тачке (вектора) $u, v \in \mathbb{R}^n$ скуп $\operatorname{conv}(u, v)$ зовемо дуж са крајњим тачкама u и v.



Слика 3.7: Пример тачака коначног скупа S са надскупом који је 1^o неконвексан, 2^o конвексан и 3^o конвексно затварање

Теорема 17. Нека је $S \subset \mathbb{R}^n$ непразан скуп. Конвексно затварање $\operatorname{conv}(S)$ је конвексан скуп.

ДОКАЗ: $\operatorname{conv}(S) = \{\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k \mid v_1, v_2, \dots, v_k \in S, \quad \lambda_i \geq 0,\ \lambda_1 + \lambda_2 + \dots + \lambda_k = 1\}$

Нека су $a, b \in S$ произвољни елементи скупа S. Тада постоје скалари $\alpha_i, \beta_i \geq 0$ за које је

$$\begin{aligned} a &= \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k, \qquad && v_1, v_2, \dots, v_k \in S,\ \alpha_1 + \alpha_2 + \dots + \alpha_k = 1, \\ b &= \beta_1 u_1 + \beta_2 u_2 + \dots + \beta_m u_m, && u_1, u_2, \dots, u_m \in S,\ \beta_1 + \beta_2 + \dots + \beta_m = 1. \end{aligned}$$

Нека је $\lambda \in [0,1]$ произвољна вредност. Онда важи

$$\begin{aligned} \lambda a + (1-\lambda)b &= \lambda(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k) + (1-\lambda)(\beta_1 u_1 + \beta_2 u_2 + \dots + \beta_m u_m) \\ &= \gamma_1 v_1 + \gamma_2 v_2 + \dots + \gamma_k v_k + \gamma_{k+1} u_1 + \dots + \gamma_{k+m} u_m, \\ &\text{где су} \quad \gamma_i = \lambda\alpha_i,\ i \leq k,\ \text{ и } \gamma_{k+j} = (1-\lambda)\beta_j. \end{aligned}$$

Покажимо да овај израз представља конвексну комбинацију вектора $v_1, \dots, v_k \in S$ и вектора $u_1, \dots, u_m \in S$, тј. да лежи у $\operatorname{conv}(S)$.

$$\lambda, 1-\lambda, \alpha_i, \beta_j \geq 0 \quad \Longrightarrow \quad \lambda\alpha_i, (1-\lambda)\beta_j \geq 0,$$

$$\begin{aligned}\gamma_1+\gamma_2+\cdots+\gamma_{k+m} &= \lambda\alpha_1+\lambda\alpha_2+\cdots+\lambda\alpha_k+(1-\lambda)\beta_1+(1-\lambda)\beta_2+\cdots+(1-\lambda)\beta_m\\ &= \lambda(\alpha_1+\alpha_2+\cdots+\alpha_k)+(1-\lambda)(\beta_1+\beta_2+\cdots+\beta_m)\\ &= \lambda+1-\lambda=1. \quad \square\end{aligned}$$

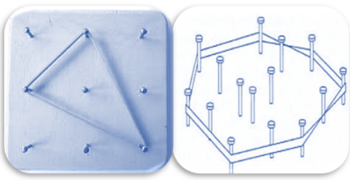
Овим доказом смо оправдали име конвексног затварања.

Теорема 18. Нека је S непразан скуп. Тада је $\text{conv}(S)$ најмањи конвексни скуп (у односу на релацију $\subseteq$) који садржи све тачке скупа S.

ДОКАЗ: Нека је M скуп који се налази у пресеку свих конвексних скупова који садрже S. Дакле, пресек свих конвексних надскупова од S означимо са M. Према теореми 15 скуп M је конвексан, па представља најмањи конвексан скуп који садржи S, јер је подскуп свих таквих конвексних скупова. Покажимо $\text{conv}(S) \subseteq M$, што ће значити $M = \text{conv}(S)$.

Нека је $v \in \text{conv}(S)$. То је v конвексна комбинација неких елемената скупа S. С обзиром да је $S \subseteq M$ и M је конвексан, то M мора да садржи било коју конвексну комбинацију својих елемената. Самим тим мора да садржи и v. $\square$

Конвексно затварање скупа S је најмањи конвексан подскуп (у односу на релацију $\subseteq$) који садржи све елементе скупа S, видети слику 3.7. Интуитивно, конвексно затварање у равни можемо замислити уз помоћ ексера и гумене траке, слика 3.8. Тачке скупа S представимо фиксираним ексерима. Тада ће границу конвексног затварања представљати гумица најмањег обима разапета око свих ових тачака. Такође, ограничена област гумицом имаће најмању површину у односу на све остале конвексне скупове који садрже све тачке скупа S. То су особине које се преносе и у веће димензије уз адекватну терминологију: обим омотача у $2D$ у $3D$ постаје површина омотача области; површина области у $2D$ је запремина конвексног затварања у $3D$. Ове мере се могу уопштити и за просторе више димензија.



Слика 3.8: Модел ексера и гумене траке

У равни конвексно затварање $\text{conv}(S)$ коначног скупа S је конвексан полигон (многоугао) чија су темена неке од тачака скупа S док се преостале тачке скупа S у унутрашњости тог полигона или на његовим ивицама. У $3D$ конвексно затварање представља конвексан полиедар. У више димензија, одговарајући термин је конвексни политоп.

Слика 3.9: Конвексни политопи у $2D$ (полигон) и $3D$ (полиедар)

У претходним примерима показан је начин одређивања средишта дужи и тежишта троугла. Ове тачке имају заједничке особине које су последице њихових алгебарских својстава. Због тога у наставку представићемо ове специјалне конвексне комбинације коначног скупа вектора.

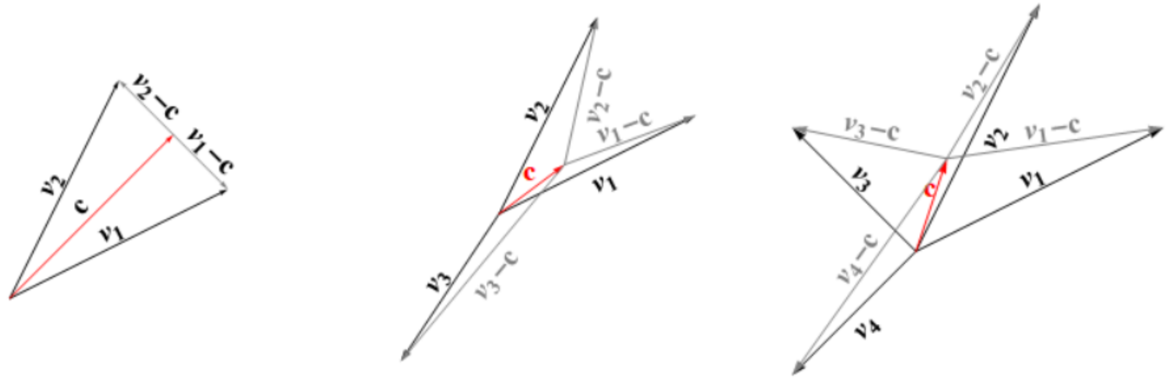
Дефиниција 29. Центар или средина скупа вектора $v_1, v_2, \ldots, v_k$, је вектор у ознаци $c = c(v_1, v_2, \ldots, v_k)$ дат линеарном комбинацијом:

$$c = \frac{1}{k}v_1 + \frac{1}{k}v_2 + \cdots + \frac{1}{k}v_k = \frac{v_1 + v_2 + \cdots + v_k}{k}. \tag{3.20}$$

Једнакост (3.20) је представљање вектора c преко вектора $v_1, v_2, \ldots, v_k$. Центар скупа вектора c је уопштење аритметичке средине бројева. Представља конвексну комбинацију вектора $v_1, v_2, \ldots, v_k$, у којој сваки вектор има исти "удео"или тежину. Дакле, центар скупа вектора је у конвексном затварању овог скупа. Претходни примери фигура аналитичке геометрије, и примери дати на слици 3.10, нам кажу да се центар скупа вектора налази "унутар"или између вектора $v_1, v_2, \ldots, v_k$, што нам описује и следећа неједнакост.

На основу неједнакости троугла важи

$$\|c\| = \left\|\frac{1}{k}v_1 + \frac{1}{k}v_2 + \cdots + \frac{1}{k}v_k\right\| \leq \frac{1}{k}\|v_1\| + \frac{1}{k}\|v_2\| + \cdots + \frac{1}{k}\|v_k\| \leq \max_{j=\overline{1,k}} \|v_j\|.$$



Слика 3.10: Центар скупа од 2, 3 и 4 вектора

Посебно својство центра $c = c(v_1, v_2, \ldots, v_k)$ описано је изразом:

$$\sum_{i=1}^{k}(v_i - c) = (v_1 - c) + (v_2 - c) + \cdots + (v_k - c) = \theta. \tag{3.21}$$

Дакле, вектори разлика $v_i - c$ се међусобно анулирају. Због тога се центар скупа вектора још и назива центар масе или тежиште у одговарајућим приликама.

Центар $c = c(v_1, v_2, \ldots, v_k)$ скупа вектора $v_1, v_2, \ldots, v_k$ реалног векторског простора V има још једно занимљиво својство. То је својство минималности збира квадрата растојања центра до скупа $v_1, v_2, \ldots, v_k$.

Теорема 19. Центар $c = c(v_1, v_2, \ldots, v_k)$ скупа вектора $v_1, v_2, \ldots, v_k$ задовољава једнакост

$$\sum_{i=1}^{k} \|v_i - c\|^2 = \min_{v \in V} \sum_{i=1}^{k} \|v_i - v\|^2. \tag{3.22}$$

ДОКАЗ: Једнакост (3.22) еквивалентна је неједнакости коју ћемо доказати:

$$\forall v \in V, \qquad \sum_{i=1}^{k} \|v_i - v\|^2 \geq \sum_{i=1}^{k} \|v_i - c\|^2.$$

$$\begin{aligned}
\sum_{i=1}^{k} \|v_i - v\|^2 &= \sum_{i=1}^{k} (v_i - v)\cdot(v_i - v) = \sum_{i=1}^{k} (v_i - c + c - v)\cdot(v_i - c + c - v) \\
&= \sum_{i=1}^{k} \left(\|v_i - c\|^2 + (v_i - c)\cdot(c - v) + (c - v)\cdot(v_i - c) + \|c - v\|^2\right) \\
&= \sum_{i=1}^{k} \|v_i - c\|^2 + k\|c - v\|^2 + \left(\sum_{i=1}^{k}(v_i - c)\right)\cdot(c - v) + (c - v)\cdot\left(\sum_{i=1}^{k}(v_i - c)\right) \\
&\overset{(3.21)}{=} \sum_{i=1}^{k} \|v_i - c\|^2 + k\|c - v\|^2 + \theta\cdot(c - v) + (c - v)\cdot\theta \\
&= \sum_{i=1}^{k} \|v_i - c\|^2 + k\|c - v\|^2 \geq \sum_{i=1}^{k} \|v_i - c\|^2.
\end{aligned}$$

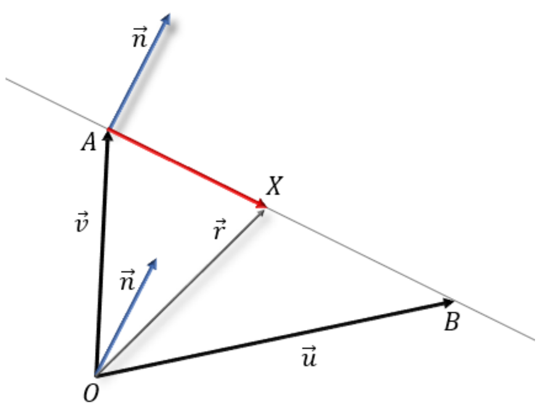
У последњем реду, једнакост очигледно важи акко $\|c - v\|^2 = 0$, тј. када је $v = c$. Тиме је особина (3.22) доказана. $\square$

За алгоритме класификације, кластеризације и оптимизације политопи играју веома важну улогу. Формална дефиниција политопа у $\mathbb{R}^n$ даје неке веома важне класе конвексних скупова. Да бисмо их разумели крећемо од геоматрија малих димензија.

Пример 3.2.4. Једначина праве у равни постављена кроз тачке $A \equiv \vec{v}$ и $B \equiv \vec{u}$, осим линеарним комбинацијама (пример 3.2.2), може се добити и помоћу операције скаларног производа. Нека је $\vec{n}$ произвољан вектор нормалан на $\overrightarrow{AB}$, тј. важи

$$\vec{n}\cdot\overrightarrow{AB} = 0.$$

Тада за сваку тачку $X \equiv \vec{r}$ праве $\overline{AB}$ важи ортогоналност одговарајућих вектора,



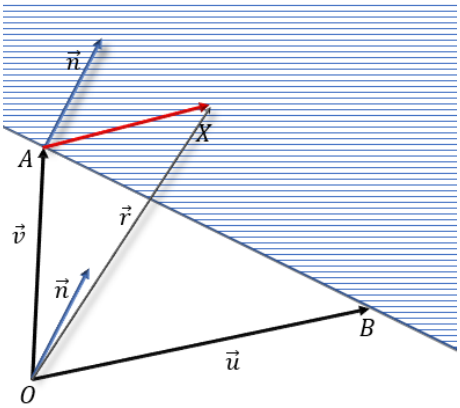
$$\begin{aligned}
\vec{n} \perp \overrightarrow{AX} \quad &\iff \quad \vec{n}\cdot\overrightarrow{AX} = 0 \\
&\iff \quad \vec{n}\cdot(\vec{r} - \vec{v}) = 0 \\
&\iff \quad \vec{n}\cdot\vec{r} = \vec{n}\cdot\vec{v}
\end{aligned}$$

Слика 3.11: Права и нормални вектор

Уколико означимо $b = \vec{n}\cdot\vec{v}$, тада је једначина праве $\overline{AB}$ са вектором нормале $\vec{n}$ дата са:

$$\vec{n}\cdot\vec{r} = b. \tag{3.23}$$

Овај облик једначине праве је веома погодан за описивање две полуравни на које је подељена раван дуж праве $\overline{AB}$. Једну од полуравни чиниће све тачке X такве да вектори $\vec{n}$ и $\overrightarrow{AX}$ образују угао у опсегу $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, видети слику 3.12.



$$
\begin{aligned}
\cos\angle(\vec{n}, \overrightarrow{AX}) > 0 \quad &\iff \quad \vec{n}\cdot\overrightarrow{AX} > 0 \\
\iff \quad & \vec{n}\cdot(\vec{r}-\vec{v}) > 0 \\
\iff \quad & \vec{n}\cdot\vec{r} > \vec{n}\cdot\vec{v} \\
\iff \quad & \vec{n}\cdot\vec{r} > b
\end{aligned}
$$

Слика 3.12: Полураван чија је граница права $\overline{AB}$

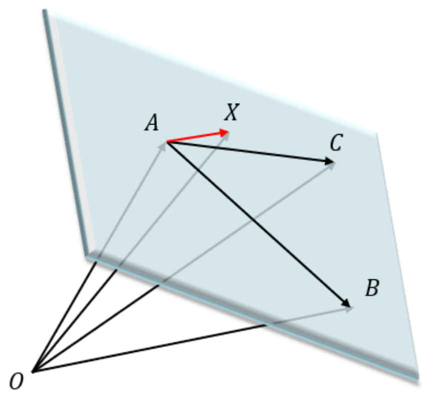
Друга полураван се на сличан начин описује. Њу ће чинити све тачке X такве да вектори $\vec{n}$ и $\overrightarrow{AX}$ образују угао у опсегу $\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$. Тачке X тада задовољавају неједнакост

$$\vec{n}\cdot\vec{r} < b.$$

Уколико овим полуравнима додамо њихову границу говоримо о затвореним полуравнима. Њихове неједначине су тада

$$\vec{n}\cdot\vec{r} \geq b \qquad \text{и} \qquad \vec{n}\cdot\vec{r} \leq b$$

Пример 3.2.5. Афина раван је дводимензионални објекат јер се све тачке равни могу описати помоћу два слободна параметра.



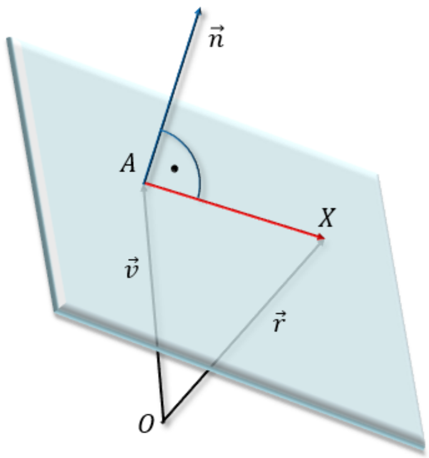
$$\alpha: \ \overrightarrow{AX} = \lambda\overrightarrow{AB} + \mu\overrightarrow{AC}, \qquad \lambda, \mu \in \mathbb{R}.$$

$$X \in \alpha \iff X \in A + \mathcal{L}(\overrightarrow{AB}, \overrightarrow{AC}).$$

Слика 3.13: Афина раван у $3D$ кроз три тачке

Осим тога, слично претходном примеру, једначина равни у простору $3D$ може се добити помоћу једне тачке те равни и вектора нормале на раван, слика 3.14.
Како свака раван дели простор $3D$ на два полупростора, овај други начин задавања равни омогућава нам да опишемо тачке полупростора на сличан начин као и њихову границу. Сваки од два полупростора описан је једном од следећих неједнакости:

$$\vec{n}\cdot\vec{r} < b, \qquad \vec{n}\cdot\vec{r} > b.$$



$$\alpha : \vec{n} \cdot \overrightarrow{AX} = 0 \iff \vec{n} \cdot (\vec{r} - \vec{v}) = 0$$
$$\iff \vec{n} \cdot \vec{r} = \vec{n} \cdot \vec{v}$$
$$\iff \vec{n} \cdot \vec{r} = b, \quad b = \vec{n} \cdot \vec{v}.$$

Слика 3.14: Афина раван у $3D$ и вектор нормале

Одговарајући затворени полупростори добијају се када укључимо границу у полупросторе. Све тачке ових затворених скупова задовољавају једну од неједнакости:

$$\vec{n} \cdot \vec{r} \leq b, \qquad \vec{n} \cdot \vec{r} \geq b.$$

По угледу на претходне примере геометрија малих димензија уводим наредне дефиниције.

Дефиниција 30. Афина хиперраван у $\mathbb{R}^n$ кроз тачку $A \equiv \vec{v}$ ортогонална на вектор $\vec{n}$ је скуп тачака $X \equiv \vec{r}$ за које важи

$$\vec{n} \cdot \vec{r} = \vec{n} \cdot \vec{v} = b.$$

Свака хиперраван дели афини простор $\mathbb{R}^n$ на два полупростора:

$$\vec{n} \cdot \vec{r} < b \qquad \text{и} \qquad \vec{n} \cdot \vec{r} > b.$$

Затворени полупростори укључују границу, тј. афину хиперраван, па све тачке ових скупова задовољавају једну од неједнакости:

$$\vec{n} \cdot \vec{r} \leq b \qquad , \qquad \vec{n} \cdot \vec{r} \geq b.$$

Доказ да је полупростор конвексан скуп остављамо као лаку вежбу читаоцу.

Дефиниција 31. Дат је скуп вектора $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ и бројеви $b_1, b_2, \ldots, b_k \in \mathbb{R}$. Скуп

$$P = \{r \in \mathbb{R}^n \mid v_j \cdot u \leq b_j, \quad j = 1, 2, \ldots, k\}$$
$$= \bigcap_{j=1}^{k} \{r \in \mathbb{R}^n \mid v_j \cdot u \leq b_j\}$$

назива се полиедар или политоп у $\mathbb{R}^n$.

Дакле, политопи представљају пресек неколико полупростора у $\mathbb{R}^n$. Сходно томе политопи могу бити ограничени и неограничени скупови, али су као пресек конвексних скупова и сами конвексни. Ове особине могу се уочити већ у $2D$ геометрији.

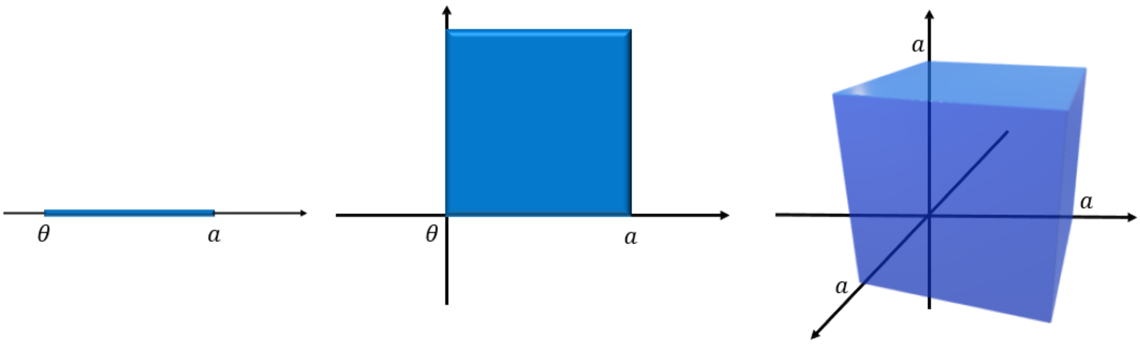
Свако конвексно затварање коначног скупа тачака представља конвексан ограничен политоп. Важи и обрат, конвексан ограничен политоп је конвексно затварање коначног скупа тачака које представљају темена тог политопа. Темена политопа зову се екстремне тачке овог скупа.

Дефиниција 32. Нека је $A \subset \mathbb{R}^n$. Тачка $a \in A$ је екстремна тачка скупа A уколико за било које две тачке $b, c \in A$ за које је $a = \dfrac{b+c}{2}$ следи да је $a = b = c$.

Пример 3.2.6. Скуп тачака $C \in \mathbb{R}^n$ зовемо коцка странице $a > 0$ (слика 3.15) када је

$$C = \left\{ \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T \mid 0 \leq x_i \leq a,\ i = 1, 2, \dots, n \right\} = [0, a]^n .$$

Јединична коцка је коцка странице дужине 1, $[0, 1]^n$.



Слика 3.15: Коцке у $1D$, $2D$ и $3D$

Веома важну класу конвексних ограничених политопа чине скупови које називамо симплексима.

Дефиниција 33. Нека су $v_0, v_1, \dots, v_k \in \mathbb{R}^n,\ k \leq n$ такви да су $v_1 - v_0, v_2 - v_0, \dots, v_k - v_0$ линеарно независни вектори. $k-$симплекс је скуп

$$\Delta_k = \operatorname{conv}(v_0, v_1, \dots, v_k).$$

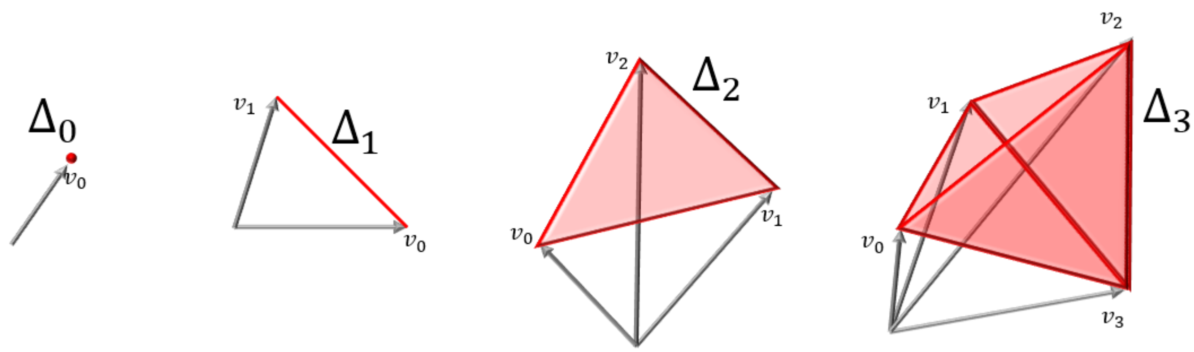
Екстремне тачке, тј. темена симплекса $\Delta_k = \operatorname{conv}(v_0, v_1, \dots, v_k)$ су одређена векторима положаја $v_0, v_1, \dots, v_k$. Да бисмо лакше разумели ове скупове, њихови примери у $3D$ дати су на слици 3.16. Симплекси $\Delta_k = \operatorname{conv}(v_0, v_1, \dots, v_k),\ k = 0, 1, 2, 3$, означцхени су црвеном бојом.

Пример 3.2.7. Користећи везу између скаларног производа вектора и операције множења матрица, дефиниција политопа (дефиниција 31)

$$\begin{aligned} P &= \{x \in \mathbb{R}^n \mid v_j \cdot x \leq b_j, \quad j = 1, 2, \dots, m\} \\ &= \left\{x \in \mathbb{R}^n \mid v_j^T x \leq b_j, \quad j = 1, 2, \dots, m\right\}, \end{aligned} \tag{3.24}$$

може се описати релацијом делимичног уређења $\geq$ на скупу матрица и вектора, дефиниција 8. Нека врсте матрице A садрже координате вектора v_j из дефниције политопа (3.24) тј,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_m^T \end{bmatrix},$$



Слика 3.16: 0−симплекс, 1−симплекс, 2−симплекс и 3−симплекс

и уведимо ознаку вектора $b = \begin{bmatrix} b_1 & b_2 & \dots & b_m \end{bmatrix}^T$. Долазимо до еквивалентне дефиниције политопа

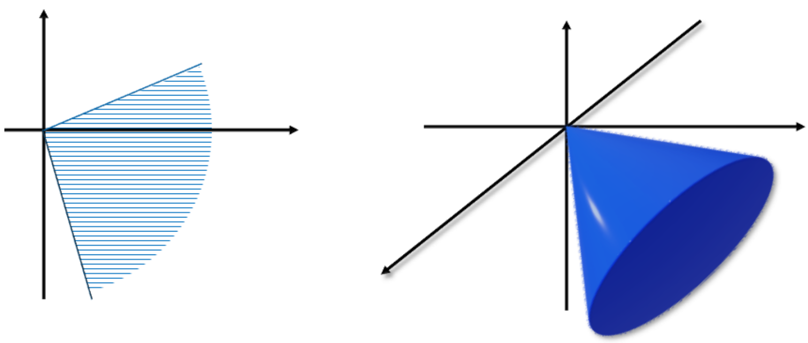
$$P = \{x \in \mathbb{R}^n \mid Ax \leq b\}.$$

Ова дефиниција политопа представља део стандарног записа проблема линеарног програмирања.

Осим политопа, у афиној геометрији значајну улогу имају и конуси.

Дефиниција 34. Скуп тачака $C \subseteq \mathbb{R}^n$ зовемо конус уколико има особину

$$(\forall v \in C)(\forall \lambda \geq 0) \quad \lambda v \in C.$$



Слика 3.17: Примери конуса у $2D$ и $3D$

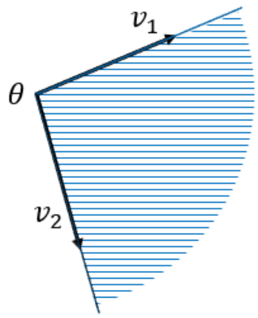
С обзиром да за $\lambda = 0$ је $\lambda v = \theta$ сви конуси садрже координатни почетак. Због тога θ називамо теме конуса. Конус је конвексан уколико представља конвексан скуп. Дакле, за произвољне $v_1, v_2 \in C$ и $\alpha \in [0, 1]$ конвексан конус садржи и све тачке облика

$$\alpha v_1 + (1 - \alpha) v_2.$$

С обзиром де конус садржи све тачке облика $\lambda_1 v_1$ и $\lambda_2 v_2,\ \lambda_1, \lambda_2 \geq 0$, закључујемо да конвексни конус садржи све тачке облика

$$\alpha_1 v_1 + \alpha_2 v_2, \quad \alpha_1, \alpha_2 \geq 0. \tag{3.25}$$

Скуп тачака описан изразом (3.25) зовемо дводимензионални сектор са теменом у θ и границама одређеним векторима v_1 и v_2, слика 3.18.



Слика 3.18: Дводимензионални сектор

3.3 Грешке

У овом одељку бавимо се елементима процене квалитета неког нумеричког резултата.

Дефиниција 35. За реалну вредност a апроксимација је нека вредност $x \in \mathbb{R}$. Грешка апроксимације је разлика

$$\varepsilon_a = x - a.$$

Апсолутна грешка апроксимације је

$$\Delta a = |x - a| = |\varepsilon_a|. \tag{3.26}$$

Релативна грешка апроксимације је тада

$$\delta a = \frac{\Delta a}{|a|} = \frac{|x - a|}{|a|}. \tag{3.27}$$

У реалним ситуацијама тачна вредност a нам углавном није позната, па самим тим нисмо у могућности да одредимо апсолутно и релативно одступање апроксимације x. Тада баратамо само границама апсолутне и релативне грешке.

Дефиниција 36. Број Δ је граница апсолутне грешке апроксимације x за a, уколико је $\Delta a \leq \Delta$. Слично, δ је граница релативне грешке када важи $\delta a \leq \delta$.

Напоменимо да ће за ознаке грешака наизменично бити коришћено

$$\varepsilon_x = \varepsilon_a, \quad \Delta x = \Delta a \quad \text{и} \quad \delta x = \delta a,$$

или нека слична скраћеница. Апсолутна грешка се обично изражава у јединицама апроксимиране величине, док се релативна грешка као безимена величина врло често изаржава у процентима. Апсолутна грешка даје меру одступања приближне вредности x од a, док релативна грешка показује меру колико је то одступање битно у односу на апроксимирану величину a. Приметимо да израз (3.27) није дефинисан за $a = 0$. Другим речима, у односу на $a = 0$ посматрамо само апсолутно одступање апроксимације x. Због тога се веома често користи еквивалентан опис релативне грешке:

$$\delta a = |r| \text{ је релативна грешка уколико је } x = a(1 + r).$$

Релативна грешка није осетљива на скалирање вредности, самим тим није осетљива на ред величине апроксимиране вредности,

$$a \mapsto \alpha a \text{ и } x \mapsto \alpha x \quad \Longrightarrow \quad \delta a = \frac{|a-x|}{|a|} = \frac{|\alpha a - \alpha x|}{|\alpha a|} = \delta(\alpha a).$$

Претходне скаларне дефиниције нам помажу да основне појмове теорије грешака проширимо на серије приближних вредности, тј. векторе. Улогу апсолутне вредности преузима норма вектора. Квалитет истовремене апроксимација серије вредности описујемо нормом вектора. Напоменимо да скуп матрица $\mathcal{M}_{m\times n}$ у односу на стандардне операције сабирања матрица и множења матрице скаларом представља векторски простор. Самим тим под појмом вектора можемо да размишљамо и о матрицама. Норма матрица је ипак нешто сложенији појам због додатне операције која прати матрице - множење матрица.

Дефиниција 37. За вектор вредности $a = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}^T$ апроксимација је неки вектор вредности $v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$. Апсолутна грешка апроксимације гласи

$$\Delta a = \|v - a\|. \tag{3.28}$$

Релативна грешка апроксимације је

$$\delta a = \frac{\Delta a}{\|a\|}. \tag{3.29}$$

Дефиниција граница грешака остаје иста и у случају апроксимације вектора.

За квалификацију грешке, али и уопште за дубље разумевање информација које неки вектор носи, осим уобичајених норми могу се користити још неке карактеристике вектора, по природи статистичке, тј. оне добијају на значају са порастом броја димензија n. У дефиницијама које следе користићемо ознаке

$$\mathbb{1} = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T \in \mathbb{R}^n \qquad \text{и} \qquad v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T \in \mathbb{R}^n.$$

- Еуклидова норма вектора:

$$\|v\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{v^T v}\,.$$

- Средње-квадратна норма:

$$\|v\|_{sr} = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}} = \frac{\|v\|}{\sqrt{n}} = \sqrt{\frac{v^T v}{n}} = \sqrt{\frac{v^T v}{\mathbb{1}^T \mathbb{1}}}\,.$$

- Средња вредност:

$$\mu(v) = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\mathbb{1}^T v}{n} = \frac{\mathbb{1}^T v}{\mathbb{1}^T \mathbb{1}}.$$

- Стандардно одступање[4]:

$$\sigma(v) = \sqrt{\frac{(x_1 - \mu(v))^2 + (x_2 - \mu(v))^2 + \cdots + (x_n - \mu(v))^2}{n}}\,.$$

[4]Веома често се користи и кориговано стандардно одступање, тј. Беселова корекција, облика

$$\sigma(v) = \sqrt{\frac{(x_1 - \mu(v))^2 + (x_2 - \mu(v))^2 + \cdots + (x_n - \mu(v))^2}{n-1}}\,.$$

$$\sigma(v) = \|v - \mu(v)\mathbb{1}\|_{sr} = \left\|v - \frac{\mathbb{1}^T v}{\mathbb{1}^T\mathbb{1}} e\right\|_{sr} = \frac{\|v - \frac{1}{n}(\mathbb{1}^T v)\mathbb{1}\|}{\sqrt{n}}.$$

Приметимо да је

$$\begin{aligned}
n\sigma(v)^2 &= \|v - \mu(v)\mathbb{1}\|^2 = (v - \mu(v)\mathbb{1})^T(v - \mu(v)\mathbb{1}) \\
&= v^T v - 2\mu(v)\mathbb{1}^T v + \mu(v)^2\mathbb{1}^T\mathbb{1} = \|v\|^2 - 2n\mu(v)^2 + n\mu(v)^2 \\
&= \|v\|^2 - n\mu(v)^2 \\
\implies \quad & \sigma(v)^2 = \|v\|_{sr}^2 - \mu(v)^2. \qquad (3.30)
\end{aligned}$$

Квадрат стандардног одступања $\sigma(v)^2$, тј. девијације, зовемо варијанса.

За вектор $v \in \mathbb{R}^n$ кажемо да је вектор стандардизованих вредности уколико има карактеристике

$$\mu(v) = 0 \qquad \text{и} \qquad \sigma(v) = 1. \qquad (3.31)$$

Постоји поступак стандардизације вектора, тј. довођења вектора на облик такав да задовољава услове (3.31).

- вектор центрираних података (вредности) вектора v дефинишемо са
$$\dot{v} = v - \mu(v)\mathbb{1},$$
- нормиран вектор центрираних вредности вектора v дефинишемо са
$$\dot{v}^* = \frac{\dot{v}}{\|\dot{v}\|},$$
- вектор стандардизованих, или нормализованих, вредности вектора v гласи:
$$\widetilde{v} = \frac{\dot{v}}{\sigma(v)} = \frac{v - \mu(v)\mathbb{1}}{\sigma(v)}.$$

Пример 3.3.1. За векторе $v, u \in \mathbb{R}^{10}$,

$$v = \begin{bmatrix}1 & 1 & -3 & 2 & 5 & 0 & -8 & -1 & 2 & 0\end{bmatrix}^T, \; u = \begin{bmatrix}2 & 1 & 3 & 2 & 0 & 0 & 6 & 3 & 2 & 5\end{bmatrix}^T$$

налазимо:

$$\begin{aligned}
\mu(v) &= \frac{1+1-3+2+5+0-8-1+2+0}{10} = -\frac{1}{10} = -0.1, \\
\mu(u) &= \frac{2+1+3+2+0+0+6+3+2+5}{10} = \frac{12}{5} = 2.4.
\end{aligned}$$

Добијене средње вредности вектора v и u указују да је вектор v скоро центрираних вредности, али да u није такав. Потражимо и остале карактеристике ових вектора.

$$\dot{v} = v - (-0.1)\,\mathbb{1} = \begin{bmatrix}1.1 & 1.1 & -2.9 & 2.1 & 5.1 & 0.1 & -7.9 & -0.9 & 2.1 & 0.1\end{bmatrix}^T,$$

$$\|v\| = \sqrt{109} \approx 10.44, \qquad \|v\|_{sr} = \frac{\sqrt{109}}{\sqrt{10}} = \sqrt{10.9} \approx 3.30,$$

$$\sigma(v)^2 = \|v\|_{sr}^2 - \mu(v)^2 = 10.9 - 0.01 = 10.89, \quad \implies \quad \sigma(v) = \sqrt{10.89} = 3.3,$$

$$\widetilde{v} = \frac{\dot{v}}{\sigma(v)} \approx \begin{bmatrix}0.33 & 0.33 & -0.88 & 0.64 & 1.54 & 0.03 & 2.45 & -0.27 & 0.64 & 0.03\end{bmatrix}^T,$$

$$\dot u = u - 2.4\,\mathbb{1} = \begin{bmatrix}-0.4 & -1.4 & 0.6 & -0.4 & -2.4 & -2.4 & 3.6 & 0.6 & -0.4 & 2.6\end{bmatrix}^T,$$
$$\|u\| = \sqrt{92} \approx 9.59, \qquad \|u\|_{sr} = \frac{\sqrt{92}}{\sqrt{10}} \approx 3.03,$$
$$\sigma(u)^2 = \|u\|_{sr}^2 - \mu(u)^2 = 9.2 - 5.76 = 3.44, \quad \Longrightarrow \quad \sigma(u) = \sqrt{3.44} \approx 1.85,$$
$$\widetilde{u} = \frac{\dot u}{\sigma(u)} \approx \begin{bmatrix}-0.22 & -0.75 & 0.32 & -0.22 & -1.29 & -1.29 & 1.94 & 0.32 & -0.22 & 1.40\end{bmatrix}^T.$$

Како бисмо што боље разумели управо одређене карактеристике вектора v и u, и трансформације центрирања и нормализације вредности компоненти, приказаћемо графички компоненте вектора $v, u, \dot v, \dot u, \widetilde{v}$ и $\widetilde{u}$, видети слику 3.19. У те сврхе векторе $n-$димензионалног простора посматраћемо као функције скупа индекса, тј.

$$v : \{1, 2, \dots, 10\} \mapsto \mathbb{R} \qquad \text{тако да} \qquad v(i) = v_i.$$

Центрирани вектор $\dot v = v - \mu(v)\mathbb{1}$ погодан је за опис девијације (одступања) компоненти вектора v од њихове просечне вредности $\mu(v)$. Центрирани вектор $\dot v$ је једнак нула-вектору θ акко је v вектор константних вредности, тј. када је $v = \alpha\,\mathbb{1},\ \alpha \in \mathbb{R}$. Стандардно одступање $\sigma(v)$ вектора v може се веома лако описати помоћу центрираног вектора
$$\sigma(v) = \|\dot v\|_{sr}.$$
Тада је
$$\widetilde{v} = \frac{\dot v}{\sigma(v)} = \frac{\dot v}{\|\dot v\|_{sr}},$$
тј. нормализован вектор је нормиран центрирани вектор у односу на средње-квадратну норму.

Нека су $v, u, \mathbb{1} \in \mathbb{R}^n,\ \mathbb{1} = \begin{bmatrix}1 & 1 & \dots & 1\end{bmatrix}^T$ вектори и $\alpha \in \mathbb{R}$ скалар. Показаћемо особине уведених мера и појмова:

а) $\mu(v+u) = \mu(v) + \mu(u)$; б) $\mu(\alpha v) = \alpha\mu(v)$; в) $\mu(\mathbb{1}) = 1$;

г) $\sigma(v+\alpha\mathbb{1}) = \sigma(v)$; д) $\sigma(\alpha v) = |\alpha|\sigma(v)$; ђ) $\sigma(\mathbb{1}) = 0$

е) $\mu(\dot v) = 0$; ж) $\sigma(\dot v) = \sigma(v)$;

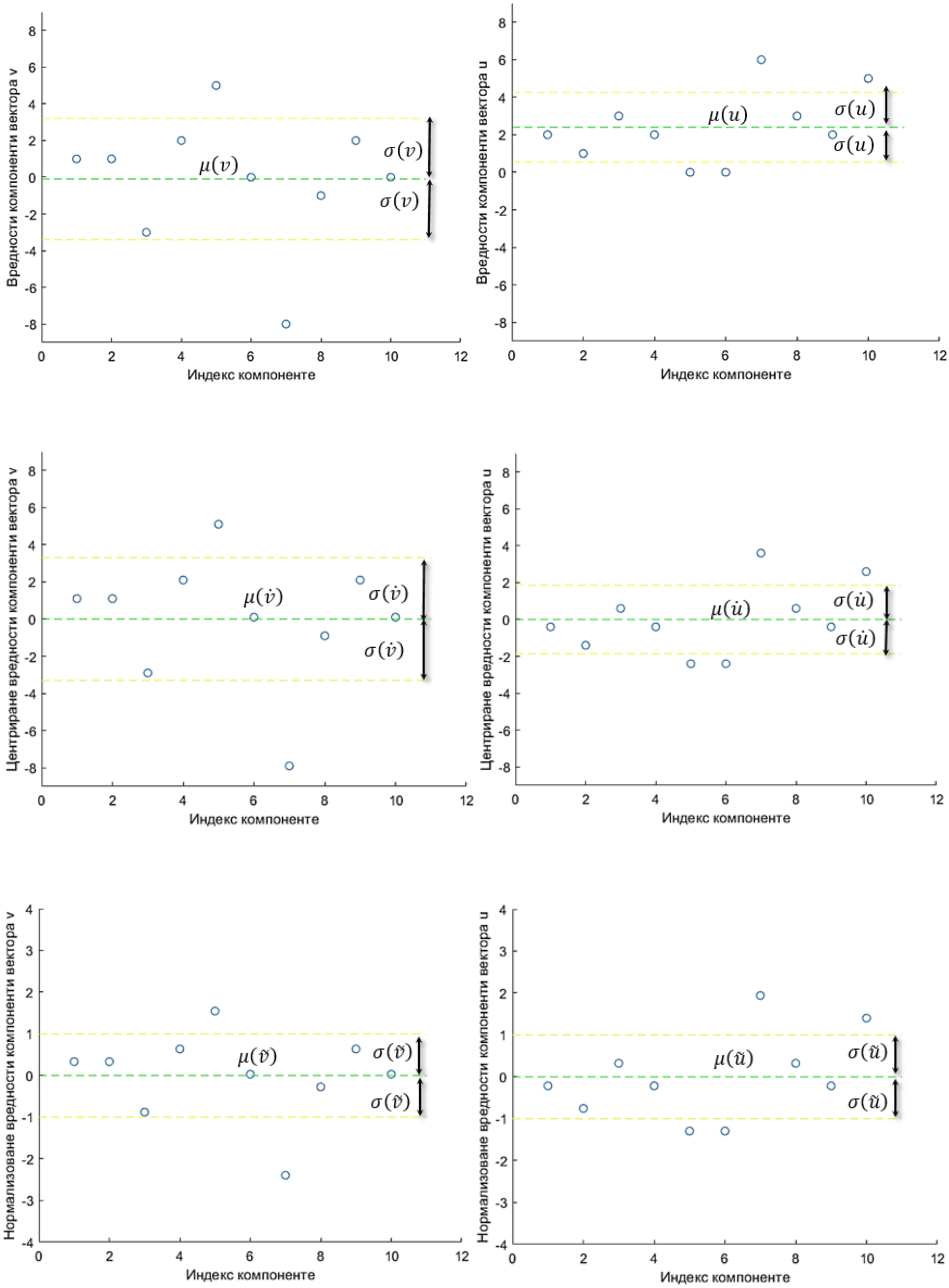
з) $\mu(\widetilde{v}) = 0$; и) $\sigma(\widetilde{v}) = 1$; ј) $\|\widetilde{v}\|^2 = n$.

а) $$\mu(v+u) = \frac{\mathbb{1}^T(v+u)}{\mathbb{1}^T\mathbb{1}} = \frac{\mathbb{1}^T v}{\mathbb{1}^T\mathbb{1}} + \frac{\mathbb{1}^T u}{\mathbb{1}^T\mathbb{1}} = \mu(v) + \mu(u).$$

б) $$\mu(\alpha v) = \frac{\mathbb{1}^T(\alpha v)}{\mathbb{1}^T\mathbb{1}} = \alpha\frac{\mathbb{1}^T v}{\mathbb{1}^T\mathbb{1}} = \alpha\mu(v).$$

в) $$\mu(\mathbb{1}) = \frac{\mathbb{1}^T\mathbb{1}}{\mathbb{1}^T\mathbb{1}} = 1.$$

г) $$\sigma(v+\alpha\mathbb{1}) = \frac{\left\|(v+\alpha\mathbb{1}) - \frac{\mathbb{1}^T(v+\alpha\mathbb{1})}{\mathbb{1}^T\mathbb{1}}\mathbb{1}\right\|}{\sqrt{n}} = \frac{\left\|v - \frac{\mathbb{1}^T v}{\mathbb{1}^T\mathbb{1}}\mathbb{1} + \alpha\mathbb{1} - \alpha\frac{\cancel{\mathbb{1}^T\mathbb{1}}}{\cancel{\mathbb{1}^T\mathbb{1}}}\mathbb{1}\right\|}{\sqrt{n}}$$
$$= \frac{\left\|v - \frac{\mathbb{1}^T v}{\mathbb{1}^T\mathbb{1}}\mathbb{1} + \cancel{\alpha\mathbb{1}} - \cancel{\alpha\mathbb{1}}\right\|}{\sqrt{n}} = \frac{\|v - \mu(v)\mathbb{1}\|}{\sqrt{n}} = \|v - \mu(v)\mathbb{1}\|_{sr} = \sigma(v).$$



Слика 3.19: Графичко поређење различитих трансформација вредности компоненти вектора v и u

д) $\sigma(\alpha v) = \dfrac{\left\|\alpha v - \alpha \frac{\mathbb{1}^T v}{\mathbb{1}^T \mathbb{1}} \mathbb{1}\right\|}{\sqrt{n}} = |\alpha| \dfrac{\left\|v - \frac{\mathbb{1}^T v}{\mathbb{1}^T \mathbb{1}} \mathbb{1}\right\|}{\sqrt{n}} = |\alpha| \sigma(v).$

ђ) $\sigma(\mathbb{1}) = \|\mathbb{1} - \mu(\mathbb{1})\mathbb{1}\|_{sr} = \|\theta\|_{sr} = 0.$

е) На основу доказаних особина под а), б) и в) важи:

$$\mu(\dot{v}) = \mu(v - \mu(v)\mathbb{1}) = \mu(v) - \mu\big(\mu(v)\mathbb{1}\big) = \mu(v) - \mu(v)\mu(\mathbb{1}) = \mu(v) - \mu(v) = 0.$$

ж) На основу доказане особине под г) важи:

$$\sigma(\dot{v}) = \sigma\big(v - \mu(v)\mathbb{1}\big) = \sigma(v).$$

Као последицу доказане једнакости добијамо де је $\widetilde{v} = \dfrac{\dot{v}}{\sigma(v)} = \dfrac{\dot{v}}{\sigma(\dot{v})}$.

з,и,ј) На основу претходно доказаних особина и (3.30), имамо

$$\begin{aligned}
\mu(\widetilde{v}) &= \mu\left(\frac{\dot{v}}{\sigma(v)}\right) = \frac{\mu(\dot{v})}{\sigma(v)} = 0,\\
\sigma(\widetilde{v}) &= \sigma\left(\frac{\dot{v}}{\sigma(v)}\right) = \frac{\sigma(\dot{v})}{\sigma(v)} = \frac{\sigma(v)}{\sigma(v)} = 1,\\
\|\widetilde{v}\|^2 &= n\|\widetilde{v}\|_{sr}^2 = n\big(\mu(\widetilde{v})^2 + \sigma(\widetilde{v})^2\big) = n.
\end{aligned}$$

Средње-квадратна норма $\|v\|_{sr} = \sqrt{\dfrac{x_1^2 + \cdots + x_n^2}{n}}$ показује колика је 'типична' вредност компоненте $|x_i|$. Тако нпр. Еуклидова норма вектора $\mathbb{1}$ је $\|\mathbb{1}\| = \sqrt{n}$, док је средње-квадратна норма једнака $\|\mathbb{1}\|_{sr} = 1$. Средње-квадратна норма, због n у имениоцу израза, погодна је за поређење просечних карактеристика вектора различитих димензија. Дакле, фактор $\dfrac{1}{n}$ обезбеђује да димензија вектора уђе у карактеристике мерења 'типичности', док квадратни корен у изразу за $\|v\|_{sr}$ омогућава да типичност буде мерена на истом реду величине као и компоненте вектора v.

У односу на средње-квадратну норму израчунава се и одговарајуће растојање. За два вектора $v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ и $u = \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}^T$ средње-квадратно одступање вредности два вектора или средње-квадратна девијација се дефинише са

$$\mathrm{rms}(v, u) = \|v - u\|_{sr} = \frac{\|v - u\|}{\sqrt{n}} = \sqrt{\frac{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}{n}}.$$

Средње-квадратно одступање описује просечан удео разлика компоненти $(x_i - y_i)^2$ у Еуклидовом растојању два вектора v и u.

С обзиром да је стандардно одступање

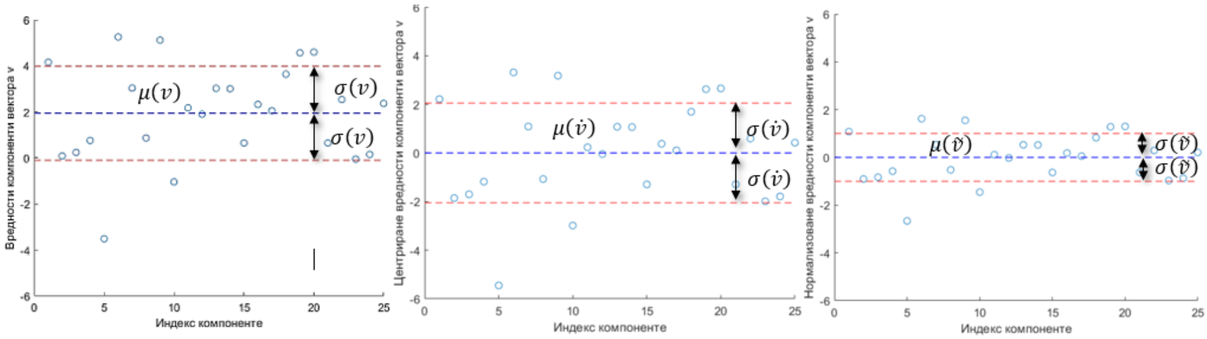
$$\sigma(v) = \|\dot{v}\|_{sr} = \mathrm{rms}\left(v, \mu(v)\mathbb{1}\right),$$

ова карактеристика вектора показује просечно одступање компоненти вектора v од њихове средње вредности $\mu(v)$. Стандардно одступање вектора v је 0 акко је вектор v са константним вредностима, тј. $v = \alpha \cdot \mathbb{1},\ \alpha \in \mathbb{R}$. Стандардно одступање вектора је близу 0 уколико су компоненте вектора међусобно блиске по вредностима, тј. уколико су вредности компоненти довољно уједначене.

Нормализован вектор је центриран (средња вредност је 0) и има стандардну девијацију једнаку 1. Како је

$$\widetilde{v} = \frac{v - \mu(v)\mathbb{1}}{\sigma(v)} = \frac{\dot{v}}{\sigma(v)} \qquad \Longrightarrow \qquad v = \sigma(v)\cdot\widetilde{v} + \mu(v)\mathbb{1}, \quad \dot{v} = \sigma(v)\cdot\widetilde{v},$$

то стандардизовани вектор $\widetilde{v}$ даје увид у однос појединачних компоненти вектора v у односу на просечно понашање свих компоненти вектора v, видети слику 3.20.



Слика 3.20: Центрирање и стандардизација вектора са 25 насумично изабраних компоненти

Пример 3.3.2. Нека су

$$v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T \qquad \text{и} \qquad \widetilde{v} = \begin{bmatrix} \widetilde{x}_1 & \widetilde{x}_2 & \dots & \widetilde{x}_n \end{bmatrix}^T,$$

Ако је $\widetilde{x}_1 = 0.7$, то значи да се компонента x_1 налази у просечном опсегу девијације од средње вредности $\mu(v)$. Уколико је нпр. $\widetilde{x}_1 = 1.5$, то означава да компонента x_1 има девијацију од средње врдности $\mu(v)$ већу од просека свих компоненти.

На основу примера 3.3.1, 3.3.2 и слика 3.19, 3.20 можемо приметити да се највећи број вредности компоненти вектора кретао у опсегу $[\mu - \sigma, \mu + \sigma]$ за сваки од анализираних вектора. Ову појаву објашњава наредна неједнакост.

Дат је вектор $v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T \in \mathbb{R}^n$. Уколико је k компоненти овог вектора са апсолутним вредностима не мањим од $a > 0$, нпр. $|x_1|, |x_2|, \dots, |x_k| \geq a > 0$, тада важи:

$$\|v\|^2 = x_1^2 + x_2^2 + \dots + x_k^2 + \dots + x_n^2 \geq x_1^2 + x_2^2 + \dots + x_k^2 \geq ka^2,$$

тј.

$$\frac{\|v\|^2}{n} \geq \frac{k}{n}a^2 \iff \|v\|_{sr}^2 \geq \frac{k}{n}a^2 \iff \frac{k}{n} \leq \left(\frac{\|v\|_{sr}}{a}\right)^2.$$

С обзиром да је $k < n$, тј. $\dfrac{k}{n} < 1$, неједнакост

$$\frac{k}{n} \leq \left(\frac{\|v\|_{sr}}{a}\right)^2 \tag{3.32}$$

пружа нетривијалне информације у случају када је $a > \|v\|_{sr}$, тј. када је $\dfrac{\|v\|_{sr}}{a} < 1$. Добијена неједнакост (3.32) позната је као Чебишевљева[5] неједнакост. Служи за процену броја компоненти вектора v чије су апсолутне вредности веће или једнаке од

[5]Пафнутиј Чебишев (1821–1894) - руски математичар

броја $a > \|v\|_{sr}$. На тај начин Чебишевљева неједнакост оправдава опис средње-квадратне норме као показатеља просечног тренда компоненти вектора. На основу неједнакости Чебишева закључујемо да вектор v не може имати много компоненти које су по апсолутној вредности веће од $\|v\|_{sr}$. Уколико је граница одступања изражена са $a = \alpha\|v\|_{sr}$, $\alpha > 1$, то (3.32) гласи $\frac{k}{n} \leq \frac{1}{\alpha^2}$. Другим речима, део компоненти вектора v за које је девијација мања од $\alpha\|v\|_{sr}$ је

$$1 - \frac{k}{n} \geq 1 - \frac{1}{\alpha^2}.$$

Уколико желимо да пратимо одступања компоненти вектора v од средње вредности $\mu(v)$ Чебишевљева неједнакост може да послужи у процени максималног броја екстремних одступања. У том случају Чебишевљева неједнакост (3.32) примењује се на вектор центрираних вредности $\dot{v}$. Како је $\|\dot{v}\|_{sr} = \sigma(v)$, то (3.32) гласи

$$\frac{k}{n} \leq \left(\frac{\sigma(v)}{a}\right)^2, \tag{3.33}$$

где је k број компоненти вектора $v = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$ за које важи да је

$$|x_i - \mu(v)| \geq a.$$

Дефиниција 38. За два вектора $v, u \in \mathbb{R}^n$, и одговарајуће векторе центрираних вредности $\dot{v}$ и $\dot{u}$, дефинише се коваријанса вектора

$$Cov(v, u) = \frac{\dot{v} \cdot \dot{u}}{n}. \tag{3.34}$$

За скуп вектора $v_1, v_2, \dots, v_m \in \mathbb{R}^n$, означимо са $A \in \mathcal{M}_{n\times m}$ матрицу чије су колоне вектори центрираних вредности $\dot{v}_1, \dot{v}_2, \dots, \dot{v}_m$. Матрица коваријансе скупа вектора $v_1, v_2, \dots, v_m$ је

$$Cov(v_1, v_2, \dots, v_m) = \left[Cov(v_i, v_j)\right]_{i,j=1,\dots,n} = \frac{A^T A}{n}. \tag{3.35}$$

Коваријансом вектора се изражава обим синхронизације варијабилности (одступања од њихове просечне вредности) у подацима које носе вектори v и u. Уколико је коваријанса позитивна, то указује да су одступања покомпонентно гледано чешће истог знака. Другим речима, компоненте вектора су често имале истовремено повећања или умањења у односу на своју просечну вредност. Таква мера може да укаже на међузависност у промени вредности компоненти два вектора. Што је вредност коваријансе два вектора ближа нули то више указује на слабу међузависност вредности унутар вектора v и u. Матрица коваријансе (3.35) такве податке међузависности даје за све парове вектора из једног скупа. Веома често се уместо димензије n у изразима (3.34) и (3.35), користи $n - 1$ као коефицијент скалирања. Коваријанса дата тим изразом назива се кориговане (према Беселовој корекцији). За велике димензије n ова разлика у коефицијенту скалирања је занемарљива.

Проблем са коваријансом као оценом повезаности смера промена компоненти јесте што њена вредност може зависити и од малог броја доминантних сабирака, тј. од малог броја довољно великих екстрема међу компонентама вектора. Скалирањем, тј. нормализацијом вектора се тај утицај донекле ублажава.

Дефиниција 39. Дати су вектори $v, u \in \mathbb{R}^n$ и њихови центрирани вектори $\dot{v}$ и $\dot{u}$. Коефицијент корелације вектора v и u дефинише се са

$$\rho(v,u) = \frac{\dot{v} \cdot \dot{u}}{\|\dot{v}\| \, \|\dot{u}\|} = \cos\angle(\dot{v},\dot{u}) = sim(\dot{v},\dot{u}). \tag{3.36}$$

За скуп вектора $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$, означимо са $A_* \in \mathcal{M}_{n\times m}$ матрицу чије су колоне $\dot{v}_1^*, \dot{v}_2^*, \ldots, \dot{v}_m^*$ нормирани вектори центрираних вредности вектора $v_1, v_2, \ldots, v_m$. Матрица корелације скупа вектора $v_1, v_2, \ldots, v_m$ је

$$Corr(v_1, v_2, \ldots, v_m) = \big[\rho(v_i, v_j)\big]_{i,j=1,\ldots,n} = A_*^T A_*. \tag{3.37}$$

Коефицијент корелације има следеће особине:

а) $\rho(v,u) = \dfrac{\widetilde{v} \cdot \widetilde{u}}{n}$, где су $\widetilde{v}$ и $\widetilde{u}$ стандардизовани вектори v и u, редом.

б) $\sigma(v+u)^2 = \sigma(v)^2 + 2\rho(v,u)\sigma(v)\sigma(u) + \sigma(u)^2$.

а) Подсетимо се да је $\widetilde{v} = \dfrac{\dot{v}}{\sigma(v)} \quad \Longrightarrow \quad \dot{v} = \sigma(v)\cdot\widetilde{v} \quad \Longrightarrow \quad \|\dot{v}\| = \sigma(v)\|\widetilde{v}\|$.

$$\rho(v,u) = \frac{\dot{v}^T\dot{u}}{\|\dot{v}\| \, \|\dot{u}\|} = \frac{\sigma(v)\sigma(u)\left(\widetilde{v}^T\widetilde{u}\right)}{\sigma(v)\sigma(u)\|\widetilde{v}\|\|\widetilde{u}\|} = \frac{\widetilde{v}^T\widetilde{u}}{\|\widetilde{v}\|\|\widetilde{u}\|} = \frac{\widetilde{v}^T\widetilde{u}}{\sqrt{n}\sqrt{n}} = \frac{\widetilde{v}^T\widetilde{u}}{n}.$$

б) Имајући у виду $(v \dot{+} u) = \dot{v} + \dot{u}$, тада важи

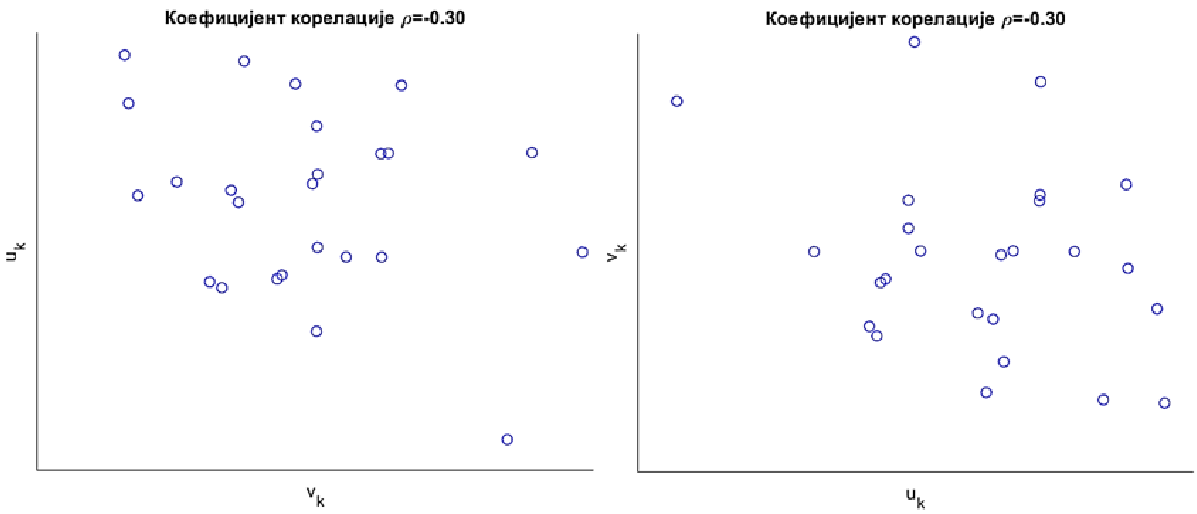
$$\begin{aligned}\sigma(v+u)^2 &= \frac{\|(v \dot{+} u)\|^2}{n} = \frac{\|\dot{v}+\dot{u}\|^2}{n} = \frac{\|\dot{v}\|^2 + \|\dot{u}\|^2 + 2\dot{v}^T\dot{u}}{n} = \frac{\|\dot{v}\|^2}{n} + \frac{\|\dot{u}\|^2}{n} + 2\frac{\dot{v}^T\dot{u}}{n}\\ &= \sigma(v)^2 + \sigma(u)^2 + 2\frac{\dot{v}^T\dot{u}}{\|\dot{v}\| \, \|\dot{u}\|}\frac{\|\dot{v}\| \, \|\dot{u}\|}{n} = \sigma(v)^2 + \sigma(u)^2 + 2\rho(v,u)\sigma(v)\sigma(u).\end{aligned}$$

Коефицијент корелације је симетрична функција, тј. $\rho(v,u) = \rho(u,v)$. Описује косинус угла између центрираних вектора, тј. показује усаглашеност компоненти вектора у смеру промена вредности у њиховим девијацијама од одговарајуће средње вредности. За векторе v и u кажемо да су некорелисани када је

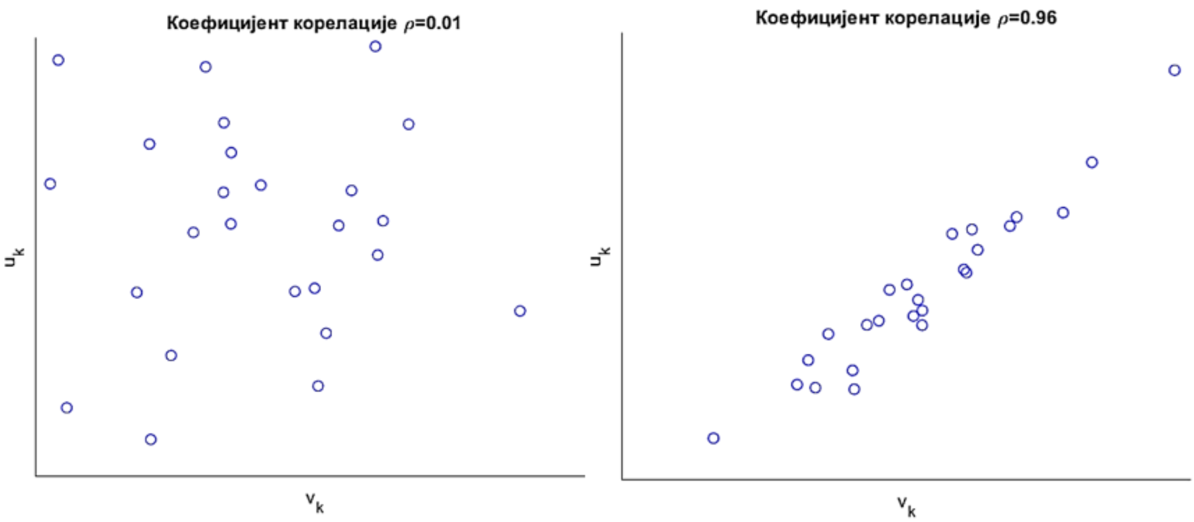
$$\rho(v,u) = 0 \quad \Longleftrightarrow \quad \dot{v} \perp \dot{u}.$$

Екстремни случај $\rho(v,u) = 1$ показује да су вектори $\dot{v}$ и $\dot{u}$ поравнани, тј. колинеарни и истог усмерења. Слично, $\rho(v,u) = -1$ указује на колинеарност и супротно усмерење вектора $\dot{v}$ и $\dot{u}$. Уобичајено је да се коефицијент корелисаности изражава у процентима. Нпр. $\rho(v,u) = 0.25$ читамо да су вектори v и u са 25% корелисаности. Коефицијент корелисаности описује ниво линеарне међузависности између компоненти вектора.

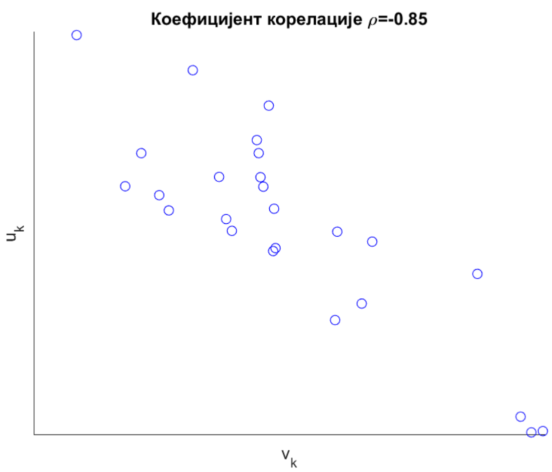
На сликама 3.21-3.23 приказани су вектори v и u као (x,y), односно (y,x) координате тачака у равни. Вектори од 25 вредности су случајно генерисани. На сликама 3.21-3.23 приказана је повезаност просторног распоред компоненти вектора као тачака (v_k, u_k), са коефицијентом корелисаности вектора $\rho(v,u)$.



Слика 3.21: Компоненте вектора v и u као координате $2D$ тачака (v_k, u_k) и (u_k, v_k)



Слика 3.22: Пример слабе и јаке корелисаности вектора

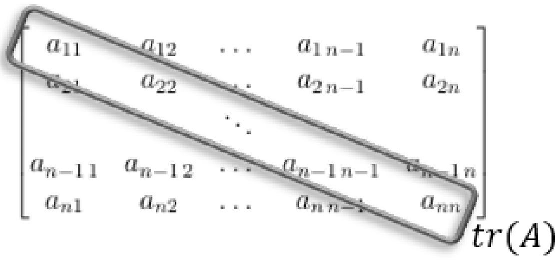


Слика 3.23: Пример значајне негативне корелисаности

3.4 Траг матрице

Над квадратним матрицама дефинисане су веома корисне скаларне функције као што су детерминанта, норма и траг. Све наведене скаларне функције су у тесној вези са сопственим вредностима матрице - њеним скаларним карактеристикама. Представљамо особине функције траг матрице. Од свих скаларних функција над квадратном матрицом њена дефиниција и поступак израчунавања су најједноставнији. Свеједно, траг матрице носи неке веома битне информације о матрицама.

Дефиниција 40. Траг квадратне матрице $A = [a_{ij}] \in \mathcal{M}_{n\times n}$ је збир њених дијагоналних елемената, тј. $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$.



Пример 3.4.1. Траг матрице $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \\ 2 & 0 & 3 \end{bmatrix}$ износи $\mathrm{tr}(A) = 1 + 2 + 3 = 6$.

Одредићемо траг јединичне и нула-матрице реда n.

$$\mathrm{tr}(I) = \mathrm{tr}(\mathrm{diag}(1, 1, \dots, 1)) = n,$$
$$\mathrm{tr}(O) = \mathrm{tr}(\mathrm{diag}(0, 0, \dots, 0)) = 0.$$

Теорема 20. За квадратну матрицу $A = [a_{ij}]$ реда n са сопственим вредностима $\lambda_1, \lambda_2, \dots, \lambda_n$, важи

$$a_{11} + a_{22} + \cdots + a_{nn} = \mathrm{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n. \tag{3.38}$$

ДОКАЗ: Карактеристични полином матрице $A \in \mathcal{M}_{n\times n}$ гласи

$$P_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} - \lambda \end{vmatrix}. \tag{3.39}$$

Сопствене вредности $\lambda_1, \lambda_2, \dots, \lambda_n$, матрице A су нуле карактеристичног полинома. Због тога је факторизација полинома $P_A(\lambda)$

$$P_A(\lambda) = \det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda). \tag{3.40}$$

Развијањем десне стране у једнакости (3.40) добијамо Вијетове формуле

$$\det(A - \lambda I) = (-\lambda)^n + (-\lambda)^{n-1}(\lambda_1 + \lambda_2 + \cdots + \lambda_n) + \cdots + \lambda_1\lambda_2\cdots\lambda_n.$$

Потражимо коефицијент уз $(-\lambda)^{n-1}$ из (3.39). Лапласовим развојем по првој колони детерминанте у (3.39) налазимо

$$P_A(\lambda) = (a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & a_{23} & \dots & a_{2\,n-1} & a_{2n} \\ a_{32} & a_{33}-\lambda & \dots & a_{3\,n-1} & a_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1\,2} & a_{n-1\,3} & \dots & a_{n-1\,n-1}-\lambda & a_{n-1\,n} \\ a_{n2} & a_{n3} & \dots & a_{n\,n-1} & a_{nn}-\lambda \end{vmatrix} + \sum_{k=2}^{n} a_{k1}A_{k1}(\lambda),$$

тј.

$$P_A(\lambda) = (a_{11}-\lambda)P_{A_{11}}(\lambda) + \sum_{k=2}^{n} a_{k1}A_{k1}(\lambda). \tag{3.41}$$

Како су сви кофактори $A_{k1}(\lambda)$, $k = 2,\dots,n$, полиноми степена не већег од $n-2$ (недостају им чланови $a_{11}-\lambda$ и $a_{kk}-\lambda$), то се тражени коефицијент уз $(-\lambda)^{n-1}$ налази у првом сабирку Лапласовог развоја (3.41). Настављајући Лапласов развој на детерминанти $P_{A_{11}}(\lambda)$, аналогно претходном објашњењу долазимо до закључка да је коефицијент уз $(-\lambda)^{n-1}$ карактеристичног полинома исти као коефицијент уз $(-\lambda)^{n-1}$ у производу дијагоналних елемената $(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$. Дакле коефицијент износи $a_{11}+a_{22}+\cdots+a_{nn} = \mathrm{tr}(A)$. $\square$

Последица 10. Сличне матрице имају једнак трагове.

ДОКАЗ: Нека су A и B сличне матрице и важи да је $A = T^{-1}BT$. С обзиром да је

$$\begin{aligned} P_A(\lambda) &= \det(A-\lambda I) = P(T^{-1}BT-\lambda I) = \det(T^{-1}BT-\lambda T^{-1}T) \\ &= \det(T^{-1})\det(B-\lambda I)\det(T) = P_B(\lambda), \end{aligned}$$

то сличне матрице имају исте спектре. Због тога су и трагови сличних матрица једнаки. $\square$

Ову особину још формулишемо у облику: траг матрице је инваријантан у односу на промену базе векторског простора.

Теорема 21. (Силвестерова[6] теорема) Нека су $A \in \mathcal{M}_{m\times n}$ и $B \in \mathcal{M}_{n\times m}$ произвољне матрице и за њихове димензије важи $m \geq n$. Тада је

$$\det(AB-\lambda I_m) = (-\lambda)^{m-n}\det(BA-\lambda I_n),$$

где I_m и I_n означавају јединичне матрице одговарајућег реда.

За дате матрице $A \in \mathcal{M}_{m\times n}$ и $B \in \mathcal{M}_{n\times m}$ на основу теореме Силвестера 21, матрице $AB \in \mathcal{M}_{m\times m}$ и $BA \in \mathcal{M}_{n\times n}$ имају исте спектре (до на вишеструкост сопствене вредности $\lambda = 0$). Другим речима

$$\sum_{\lambda\in\mathrm{Sp}(AB)} \lambda = \sum_{\mu\in\mathrm{Sp}(BA)} \mu.$$

Због тога ове матрице имају и једнаке трагове. Закључујемо:

$$\mathrm{tr}(AB) = \mathrm{tr}(BA). \tag{3.42}$$

Следећим теоремама дате су још неке особине трага матрице. Њихови докази остављени су као лака вежба за читаоца.

[6] Џејмс Џозеф Силвестер (1814–1897) амерички математичар

Теорема 22. Нека су $A, B \in \mathcal{M}_{n\times n}$ и α, β скалари. Тада важе следеће једнакости.

а) $\mathrm{tr}(A^T) = \mathrm{tr}(A)$.

б) $\mathrm{tr}(A^H) = \overline{\mathrm{tr}(A)}$.

в) $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

г) $\mathrm{tr}(\alpha A) = \alpha\,\mathrm{tr}(A)$.

д) $\mathrm{tr}(\alpha A + \beta B) = \alpha\,\mathrm{tr}(A) + \beta\,\mathrm{tr}(B)$.

Теорема 23. Нека су $A \in \mathcal{M}_{m\times n}$ и $B \in \mathcal{M}_{m\times n}$. Тада важи

$$\mathrm{tr}(A^T B) = \mathrm{tr}(B^T A).$$

Показану особине теореме 23 издвајамо као веома важну. На основу ње траг матрице сматра се скаларним производом (дефиниција 16) на векторском простору реалних матрица $\mathbb{R}_{m\times n}$

$$\langle A, B\rangle_{\mathbb{R}} = \mathrm{tr}(A^T B) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}. \tag{3.43}$$

3.5 Норма матрица

Дефинисање норме на простору матрица носи специфичне захтеве интерпретације самих матрица. Уколико матрице посматрамо као елементе векторског простора $\mathcal{M}_{m\times n}$ димензије $mn \times 1$, онда се норме вектора, нпр. L_p или било које друге, лако примењују на матрице. У таквој ситуацији матрице посматрамо као дугачке листе елемената, а сам поступак представља векторизацију матрица, видети слику 2.2. Норме овог типа називамо матричне члан-по-члан норме.

Норма матрица индукована скаларним производом (3.43), је веома лака за израчунавање. Ову норму називамо Фробенијусова норма матрице $A = [a_{ij}] \in \mathcal{M}_{m\times n}$.

$$\|A\|_F = \sqrt{(A, A)_{\mathbb{R}}} = \sqrt{\mathrm{tr}(A^T A)} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}. \tag{3.44}$$

Уочавамо аналогију у изразима између Фробенијусове норме матрица и Еуклидове норме вектора.

Пример 3.5.1. За матрицу $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{bmatrix}$ имамо да је $A^T A = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 6 \\ 1 & 6 & 13 \end{bmatrix}$. Тада је

$$\|A\|_F = \sqrt{\mathrm{tr}(A^T A)} = \sqrt{19}.$$

Фробенијусова норма се слаже са стандардном (Еуклидовом) нормом вектора на следећи начин

$$\|Av\| \leq \|A\|_F \|v\|. \tag{3.45}$$

Ову особину зовемо субмултипликативност матричне и векторске норме.

Да бисмо показали да особина (3.45) заиста важи означимо елементе матрице $A = [a_{ij}]_{n\times n}$ и вектора $v = [x_i]_{n\times 1}$. Тада је $i-$та компонента вектора Av дата са $(Av)_i = \sum_{j=1}^{n} a_{ij}x_j$, тј. вектор Av гласи

$$Av = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \sum_{j=1}^{n} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{bmatrix}.$$

Због тога имамо да је $\|Av\|^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij}x_j \right)^2$. Имајући у виду Шварцову неједнакост (3.1) за све $i = 1, 2, \ldots, n$,

$$\left(\sum_{j=1}^{n} a_{ij}x_j \right)^2 \leq \left(\sum_{j=1}^{n} a_{ij}^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right) = \left(\sum_{j=1}^{n} a_{ij}^2 \right) \|v\|^2,$$

лако долазимо до неједнакости

$$\|Av\|^2 \leq \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \right) \|v\|^2 = \|A\|_F^2 \|v\|^2.$$

Неједнакост (3.45) одатле једноставно следи.

Последица 11. За Фробенијусову норму матрице важи

$$\|AB\|_F \leq \|A\|_F \|B\|_F. \tag{3.46}$$

Уколико означимо колоне реалне матрице A са $a_1, a_2, \ldots, a_n$,

$$A = \left[\begin{array}{c|c|c|c} a_1 & a_2 & \ldots & a_n \end{array} \right],$$

Фробенијусова норма може описати и L_2 норму ове колекције вектора. Грамова матрица A^TA на главној дијагонали има елементе $a_i^T a_i$, тј. садржи квадрате норми вектора колона a_i. Због тога је

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \|a_i\|^2}.$$

Фробенијусова норма је дакле Еуклидова норма вектора Еуклидових норми колона матрице A :

$$v = \begin{bmatrix} \|a_1\| & \|a_2\| & \ldots & \|a_n\| \end{bmatrix}, \qquad \Longrightarrow \qquad \|A\|_F = \|v\|. \tag{3.47}$$

Описане особине Фробенијусове норме повод су за увођење следеће дефиниције.

Дефиниција 41. Функција $\|\cdot\| : \mathcal{M}_{m\times n} \mapsto \mathbb{R}_0^+,\ m,n \in \mathbb{N}$ је матрична норма ако $\forall A \in \mathcal{M}_{m\times n},\ B \in \mathcal{M}_{n\times p}$ и $\forall c \in \mathbb{R}$ задовољава следеће особине:

МН1 $\|A\| \geq 0$ и $\|A\| = 0 \iff A = O$,

МН2 $\|cA\| = |c|\,\|A\|$,

МН3 $\|A+B\| \leq \|A\| + \|B\|$,

МН4 $\|A\,B\| \leq \|A\|\,\|B\|$.

Дефиниција 42. За матрицу A спектрални радијус (полупречник), у ознаци $\rho(A)$ је

$$\rho(A) = \max\left\{|\lambda| \;\middle|\; \lambda \in \mathrm{Sp}(A)\right\}.$$

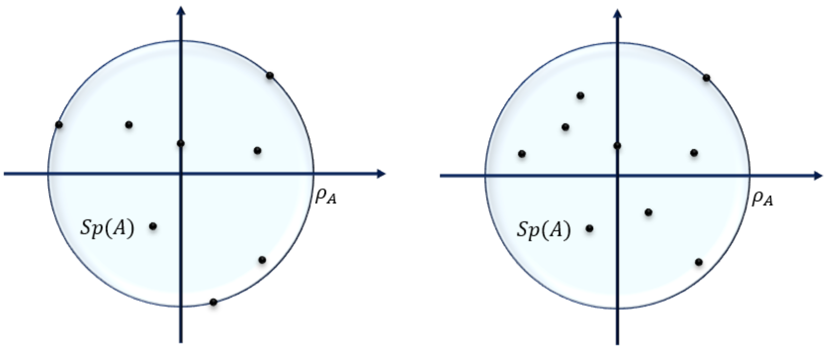
Спектрални радијус матрице A је полупречник минималне централне кружнице која садржи све сопствене вредности матрице A. Спектрални полупречник представља горњу границу модула сопствених вредности.

Дефиниција 43. За сопствену вредност λ матрице A кажемо да је доминантна уколико је по модулу већа или једнака од свих осталих сопствених вредности те матрице.

Напомена 5. Доминантна сопствена вредност неке матрице је проста нула карактеристичног полинома када је строго већа од свих осталих нула. Све доминантне сопствене вредности имају једнаке модуле.

Пример 3.5.2. Сопствене вредности матрице $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & -4 & 3 \\ 0 & 0 & 3 \end{bmatrix}$ су: $\mathrm{Sp}(A) = \{1, -4, 3\}$. Спектрални полупречник је $\rho(A) = \max\left\{|1|, |-4|, |3|\right\} = 4$. Доминантна сопствена вредност је $\lambda_{\max} = -4$.

Спектрални радијус представља полупречник централне кружнице (кружнице са центром у координатном почетку комплексне равни) која садржи све сопствене вредности матрице A. Ова карактеристика матрица игра веома важну улогу у процесима конвергенције низова матрица. Уколико означимо неку доминантну сопствену вредност матрице са $\lambda_{\max}$, онда је $\rho(A) = |\lambda_{\max}|$.



Слика 3.24: Кружница спектралног полупречника матрице са више доминантних сопствених вредности и са једном доминантном сопственом вредношћу

Теорема 24. Ако је $\|\cdot\| : \mathcal{M}_{m\times n} \mapsto \mathbb{R}_0^+$ произвољна матрична норма и $\rho(A)$ спектрални радијус матрице $A \in \mathcal{M}_{n\times n}$, тада је $\rho(A) \le \|A\|$.

ДОКАЗ: Нека је $\lambda \in \mathrm{Sp}(A)$ доминантна сопствена вредност матрице A и v одговарајући сопствени вектор. Тада је $\rho(A) = |\lambda|$ и

$$\|Av\| = \|\lambda v\| = |\lambda|\,\|v\| \le \|A\|\|v\| \quad \Longrightarrow \quad \|A\| \ge |\lambda| = \rho(A). \quad \square$$

Последица 12. Нека је $A \in \mathcal{M}_{n\times n}$ матрица која за неку матричну норму има својство $\|A\| < 1$. Тада је $\lim\limits_{k\to\infty} A^k = O$, тј. све компоненте матрице A^k теже нули када степен k тежи бесконачности.

ДОКАЗ: За $\|A\| < 1$ важи $\|A^2 - O\| = \|A^2\| \le \|A\|\,\|A\| = \|A\|^2$. Настављајући индуктивно закључујемо да је $\|A^k - O\| = \|A^k\| \le \|A\|^k \to 0,\ k \to \infty$. $\square$

Сагледавања норме Фробенијуса кроз норму вектора (3.47) доводи до нових начина дефинисања норми на скупу матрица. Посматрање матрица као скупова вектора доноси различите нове могућности за дефинисање матричне норме. Матрице као скупови вектора-врста или вектора-колона могу користити векторску норму за сваки од вектора скупа. На добијени вектор норми поново можемо применити неку векторску норму из дефиниције 3.1.

Пример 3.5.3. За матрицу $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{bmatrix}$ на сваку од колона можемо применити нпр. L_1 норму. Резултат је вектор-врста норми сваке од колона,

$$v = \begin{bmatrix} 1+|-1| & 2+0 & 3+2 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 5 \end{bmatrix}.$$

Сада на вектор v можемо применити, нпр. L_2 норму и то прогласимо нормом матрице A,

$$\|A\|_{1,2} = \sqrt{2^2 + 2^2 + 5^2} = \sqrt{33}.$$

Слично, можемо израчунати $L_{2,1}$ норму матрице A, да бисмо утврдили $L_{1,2} \ne L_{2,1}$.

$$\begin{aligned} u &= \begin{bmatrix} \sqrt{1^2 + (-1)^2} & \sqrt{2^2 + 0^2} & \sqrt{3^2 + 2^2} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 2 & \sqrt{13} \end{bmatrix}, \\ \|A\|_{2,1} &= \sqrt{2} + 2 + \sqrt{13}. \end{aligned}$$

Комбинација L_1 и L_∞ норме даје $L_{1,\infty}$ матричну норму која рачуна максималну суму модула компоненти вектора-колона матрице. У случају матрице A

$$\|A\|_{1,\infty} = \|v\|_\infty = 5.$$

Ако два пута применимо L_∞ норму (видети 3.1) добијамо одговарајућу члан по члан $L_{\infty,\infty}$ норму за матрице.

$$A = [a_{ij}] \in \mathcal{M}_{m\times n}, \quad \|A\|_{\infty,\infty} = \max_{i,j}\{|a_{ij}|\}.$$

За матрицу $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \end{bmatrix}$ је тада $\|A\|_{\infty,\infty} = 3$.

Норма матрица може да се дефинише на још неке веома корисне начине. Кључ рада са њима је неједнакост облика (3.45).

Дефинисање норме на простору матрица носи специфичне захтеве интерпретације самих матрица. За примену матрица веома је битна операција множења. Посебан нагласак је на операцији множења матрице и вектора. Због тога су направљени покушаји дефинисања норме матрица која би се слагала са овом операцијом, тј. да поседује особину мултипликативности $\|AB\| = \|A\|\,\|B\|$. Међутим, најбољи постигнут резултат јесте субмултипликативност норме матрица,

$$\|AB\| \leq \|A\|\,\|B\|.$$

Један од примера субмултипликативне норме је норма матрица индукована векторском нормом. Индукована матрична норма има особине

$$\|AB\| \leq \|A\|\ \|B\|,\ \text{а посебно}\ \|Av\|_V \leq \|A\|\ \|v\|_V, \tag{3.48}$$

где је $\|\cdot\|_V$ нека норма вектора. Индуковану матричну норму дефинисаћемо на скупу квадратних матрица[7].

Дефиниција 44. Матрична норма $\|\cdot\|$ индукована векторском нормом $\|\cdot\|_V$ дефинисана је са

$$\|A\| = \sup_{v\neq\theta}\frac{\|Av\|_V}{\|v\|_V} = \sup_{v\neq\theta}\frac{\|A\frac{v}{\|v\|_V}\|_V}{\|\frac{v}{\|v\|_V}\|_V} = \max_{\|v\|_V=1}\|Av\|_V. \tag{3.49}$$

Другим речима, матрична норма $\|A\|$ је најмања вредност C за коју важи

$$\|Av\|_V \leq C\|v\|_V, \qquad \text{за сваки вектор } v\in\mathbb{R}^n.$$

Матрична норма описује максимално релативно истезање вектора под дејством матрице A, тј. максимално истезање вектора са јединичне сфере под дејством матрице A, мерено одговарајућом векторском нормом. Геометрија јединичне сфере векторског простора условљена је избором векторске норме $\|\cdot\|_V$, видети пример 3.1.3. Норма матрица дефинисана са (3.49) испуњава услове MH1-MH4 дефиниције 41. Кључ доказа ових особина је неједнакост

$$\forall v\in\mathbb{R}^n, \qquad \|Av\|_V \overset{v\neq\theta}{=} \frac{\|Av\|_V}{\|v\|_V}\|v\|_V \leq \left(\sup_{u\neq\theta}\frac{\|Au\|_V}{\|u\|_V}\right)\|v\|_V = \|A\|\,\|v\|_V \tag{3.50}$$

<u>H1 :</u> $\|A\| \geq 0$ очигледно важи, јер је израз $\|A\|$ дат као максимална вредност ненегативних вредности. Покажимо да је $\|A\| = 0 \iff A = O$.

$$\|A\| = 0 \iff \sup_{v\neq\theta}\frac{\|Av\|_V}{\|v\|_V} = 0 \iff \|Av\|_V = 0, \forall v\in V \iff A = O.$$

<u>H2 :</u> $\|\alpha A\| = |\alpha|\,\|A\|,\ \forall A$

$$\|\alpha A\| = \max_{\|v\|_V=1}\|\alpha Av\|_V = \max_{\|v\|_V=1}|\alpha|\|Av\|_V = \alpha\max_{\|v\|_V=1}\|Av\|_V = \alpha\|A\|.$$

[7] Норма матрица може да се дефинише и на скупу правоугаоних матрица.

Н3 : $\|A+B\| \leq \|A\| + \|B\|, \ \forall A, B$

$$\begin{aligned}\|A+B\| &= \max_{\|v\|_V=1} \|(A+B)v\|_V = \max_{\|v\|_V=1} \|Av+Bv\|_V \leq \max_{\|v\|_V=1} \{\|Av\|_V + \|Bv\|_V\} \\ &\leq \max_{\|v\|_V=1} \|Av\|_V + \max_{\|v\|_V=1} \|Bv\|_V = \|A\| + \|B\|.\end{aligned}$$

$\|AB\| \leq \|A\|\,\|B\|$:

$$\begin{aligned}\|AB\| &= \max_{\|v\|_V=1} \|ABv\|_V = \max_{\|v\|_V=1} \|A(Bv)\|_V \overset{(3.50)}{\leq} \max_{\|v\|_V=1} \|A\|\,\|(Bv)\|_V \\ &= \|A\| \max_{\|v\|_V=1} \|(Bv)\|_V = \|A\|\,\|B\|\end{aligned}$$

Напомена 6. Иста дефиниција индуковане норме важи за комплексне матрице у односу на норме вектора комплексног векторског простора $\mathbb{C}^n$. Такође, особине матричних норми које ћемо приказати у наставку важе како за реалне, тако и за комплексне матрице.

Пример 3.5.4. У свакој индукованој матричној норми је $\|I\| = 1$ и $\|O\| = 0$.

$$\|I\| = \sup_{v\neq\theta} \frac{\|Iv\|_V}{\|v\|_V} = \sup_{v\neq\theta} \frac{\|v\|_V}{\|v\|_V} = 1.$$

Теорема 25. Нека је A регуларна матрица. За индуковану матричну норму важи да је

$$\|A^{-1}\| \geq \frac{1}{\|A\|}.$$

ДОКАЗ: Нека је v ненула вектор.

$$\|v\|_V = \|A^{-1}Av\|_V \leq \|A^{-1}\|\ \|Av\|_V \leq \|A^{-1}\|\ \|A\|\ \|v\|_V \qquad \Big| : \|v\|_V$$

$$\implies 1 \leq \|A^{-1}\|\ \|A\| \implies \|A^{-1}\| \geq \frac{1}{\|A\|}. \quad \square$$

У наставку користићемо исту ознаку за норму вектора $\|\cdot\|_V$ и њом индуковану норму матрица. Из контекста моћиће да се протумачи о којој норми је реч.

Пример 3.5.5. Нека је A регуларна матрица и A^{-1} њена инверзна матрица. Тада важи

$$\frac{1}{\|A^{-1}\|} = \frac{1}{\sup\limits_{v\neq\theta} \dfrac{\|A^{-1}v\|}{\|v\|}} = \inf_{v\neq\theta} \frac{\|v\|}{\|A^{-1}v\|} = \inf_{v\neq\theta} \frac{\|AA^{-1}v\|}{\|A^{-1}v\|}. \tag{3.51}$$

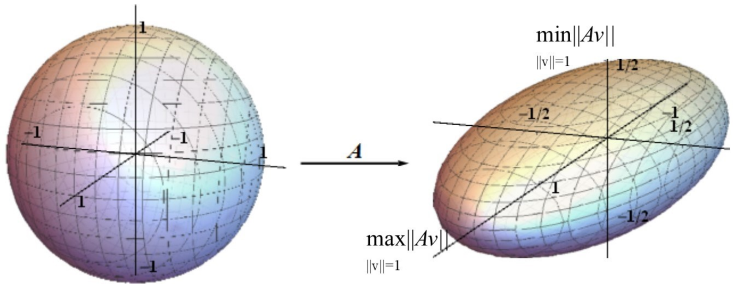
Уведимо ознаку $u = A^{-1}v$. С обзиром да је A регуларна матрица то је

$$v \neq \theta \iff A^{-1}v \neq \theta \iff u \neq \theta.$$

Једнакост (3.51) тада постаје

$$\frac{1}{\|A^{-1}\|} = \inf_{u\neq\theta} \frac{\|Au\|}{\|u\|} = \min_{\|u\|=1} \|Au\|. \tag{3.52}$$

Другим речима, вредност $\dfrac{1}{\|A^{-1}\|}$ показује нам у којој мери регуларна матрица A највише контрахује векторе са јединичне сфере.



Слика 3.25: Норма матрица и слика јединичне сфере

Најзначајније индуковане матричне норме су у тесној вези са најчешће коришћеним векторским L_p нормама.

Пример 3.5.6. Матрична норма индукована векторском нормом L_1 носи исту ознаку:

$$L_1: \ \|A\|_1 = \sup_{v \neq \theta} \frac{\|Av\|_1}{\|v\|_1} = \max_{\|v\|_1=1} \|Av\|_1 .$$

Да бисмо извели израз за рачунање матричне норме L_1, уведимо ознаке за колоне матрице $A \in \mathcal{M}_{n\times n}$ и за компоненте јединичног вектора $v \in \mathbb{R}^n$.

$$A = \left[\begin{array}{c|c|c|c} a_1 & a_2 & \dots & a_n \end{array}\right], \qquad v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Тада, како је $\|v\|_1 = 1 \iff \sum_{i=1}^{n} |x_i| = 1$, то је

$$\|Av\|_1 = \left\| \sum_{i=1}^{n} x_i a_i \right\|_1 \overset{\text{H3}}{\leq} \sum_{i=1}^{n} \|x_i a_i\|_1 \overset{\text{H2}}{=} \sum_{i=1}^{n} |x_i| \, \|a_i\|_1 \leq \max_i \|a_i\|_1 \sum_{i=1}^{n} |x_i| = \max_i \|a_i\|_1.$$

Због тога за матрицу $A = [a_{ij}]$ важи

$$\|A\|_1 = \max_{\|v\|_1=1} \|Av\|_1 = \max_i \|a_i\|_1 = \max_i \sum_{j=1}^{n} |a_{ji}|. \tag{3.53}$$

L_1 матрична норма представља максималну суму модула компоненти вектора-колона матрице A, тј. еквивалентна је $L_{1,\infty}$ норми.

Пример 3.5.7. Слично претходном примеру, показаћемо да је индукована L_∞ норма матрице A еквивалентна $\|A^T\|_{1,\infty}$. Нека је $v \in \mathbb{C}^n$ јединични вектор у односу на L_∞ норму,

$$\|v\|_\infty = \max_{1\leq i\leq n} |x_i| = 1.$$

Посматрамо поделу матрице A на векторе-врста, $A = \begin{bmatrix} a_1^T \\ \hline a_2^T \\ \hline \vdots \\ \hline a_n^T \end{bmatrix}$. Тада је

$$\|Av\|_\infty = \left\| \begin{bmatrix} a_1^T v \\ a_2^T v \\ \vdots \\ a_n^T v \end{bmatrix} \right\|_\infty = \max_{1\le i\le n} |a_i^T v| = \max_{1\le i\le n} |\sum_{j=1}^n a_{ij}x_j| \le \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}x_j|$$
$$\overset{|x_j|\le 1}{\le} \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|.$$

L_∞ матрична норма матрице A представља максималну суму модула компоненти вектора-врста матрице A.

Пример 3.5.8. За извођење израза L_2 матричне норме користићемо SVD матрице A, и својства сингуларних вредности.

$$\|A\|_2 = \max_{\|v\|_2=1} \|Av\|_2 = \sigma_1.$$

Дакле L_2 матрична норма представља максималну сингуларну вредност матрице A и назива се још и спектрална норма. Уколико је матрица A регуларна, и σ_n њена најмања сингуларна вредност, тада је

$$A = U\Sigma V^T \quad \Longrightarrow \quad A^{-1} = V\Sigma^{-1}U^T,$$

па је $\dfrac{1}{\sigma_n}$ максимална сингуларна вредност матрице A^{-1}. Закључујемо да је $\|A^{-1}\|_2 = \dfrac{1}{\sigma_n}$.

За $A = [a_{ij}] \in \mathcal{M}_{n\times n}$ најзначајније матричне норме индуковане векторском нормом су:

L_1: $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \qquad \dfrac{1}{\|A^{-1}\|_1} = \min_j \sum_{i=1}^n |a_{ij}|,$

L_2: $\|A\|_2 = \sigma_1 = \sqrt{\rho(A^H A)} = \max\{\sqrt{|\lambda|} \mid \lambda \in Sp(A^H A)\},$

$\dfrac{1}{\|A^{-1}\|_2} = \sigma_n = \min\{\sqrt{|\lambda|} \mid \lambda \in Sp(A^H A)\},$

L_∞: $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|, \qquad \dfrac{1}{\|A^{-1}\|_\infty} = \min_i \sum_{j=1}^n |a_{ij}|.$

Сличност између L_2 и Фробенијусове матричне норме (3.44) лежи у њиховој повезаности са Еуклидовом нормом вектора. Ту сличност потврђујемо и кроз наредно тврђење.

Теорема 26. Ако је Q унитарна матрица реда n и A произвољна квадратна матрица реда n тада је
$$\|A\|_2 = \|QA\|_2.$$

ДОКАЗ: $\|QA\|_2 = \max_{\|v\|_2=1} \|QAv\|_2 \overset{\|Q(Av)\|_2=\|Av\|_2}{=} \max_{\|v\|_2=1} \|Av\|_2 = \|A\|_2$. $\square$

Пример 3.5.9. Нека је T регуларна матрица и $\| \ \|$ произвољна матрична норма. Показаћемо да је још једна матрична норма тада дата са

$$(|A|) = \|T^{-1}AT\|.$$

Н1: $(|A|) = \|T^{-1}AT\| \geq 0$ јер је $\| \ \|$ норма. Из истог разлога је

$$0 = (|A|) = \|T^{-1}AT\| \iff T^{-1}AT = O \iff A = O.$$

Н2: $(|\lambda A|) = \|T^{-1}(\lambda A)T\| = \|\lambda(T^{-1}AT)\| = \lambda\|T^{-1}AT\| = \lambda(|A|)$.

Н3: $(|A + B|) = \|T^{-1}(A + B)T\| = \|T^{-1}AT + T^{-1}BT\| \leq \|T^{-1}AT\| + \|T^{-1}BT\| = (|A|) + (|B|)$.

Норма матрица игра велику улогу у испитивању стабилности израчунавања. У те сврхе, користићемо основне појмове из теорије грешака раније уведене.

Пример 3.5.10. Претпоставимо да је потребно израчунати Av, али нам v није тачно дато. То може бити последица грешака заокруживања, мерења или претходних нумеричких израчунавања. Уместо са тачном вредношћу v располажемо са $v+\varepsilon$. Због тога је граница апсолутне грешке улазног податка дата са $\|\varepsilon\|$. Желимо да проценимо утицај грешке ε на израчунавање Av. Граница апсолутне грешке израчунавања је дата изразом

$$\Delta(Av) = \|A(v + \varepsilon) - Av\| = \|A\varepsilon\| \leq \|A\| \ \|\varepsilon\| = \|A\|\Delta v.$$

Норма матрице представља грубу процену коефицијента пропорционалности између граница апсолутне грешке израчунавања Av и апсолутне грешке улазног податка v.

За процену релативне грешке $\delta(Av) = \dfrac{\|A(v + \varepsilon) - Av\|}{\|Av\|}$ израчунавања Av, претпоставимо регуларност матрице A. Тада је

$$\delta(Av) = \frac{\|A(v + \varepsilon) - Av\|}{\|Av\|} \leq \frac{\|A\| \ \|\varepsilon\|}{\|Av\|} = \frac{\|A\| \ \|A^{-1}\| \ \|\varepsilon\|}{\|A^{-1}\| \ \|Av\|}.$$

$$\|A^{-1}\| \ \|Av\| \geq \|A^{-1}Av\| = \|v\| \Longrightarrow \frac{1}{\|A^{-1}\| \ \|Av\|} \leq \frac{1}{\|v\|}.$$

$$\Longrightarrow \frac{\|A(v + \varepsilon) - Av\|}{\|Av\|} \leq \|A\| \ \|A^{-1}\| \frac{\|\varepsilon\|}{\|v\|} = \|A\| \ \|A^{-1}\| \ \delta v$$

Граница релативне грешке резултата $\delta(Av) = \dfrac{\|A(v + \varepsilon) - Av\|}{\|Av\|}$ је пропорционална граници релативне грешке улазних података $\delta v = \dfrac{\|\varepsilon\|}{\|v\|}$. Горња граница коефицијента пропорционалности дата је изразом

$$\kappa(A) = \|A\| \ \|A^{-1}\|$$

и назива се кондициони број матрице A. Представља грубу процену коефицијента увећања релативне грешке у израчунавањима.

Уколико је L_2 норма матрица у односу на коју се рачуна кондициони број регуларне матрице $A \in \mathcal{M}_{n\times n}$, тада је

$$\kappa(A) = \|A\|\;\|A^{-1}\| = \frac{\sigma_1}{\sigma_n}\,,$$

где су σ_1 и σ_n највећа и најмања сингуларна вредност матрице A, редом.

Регуларна матрица A је лоше условљена уколико мала релативна промена унутар вредности матрице A производи велике релативне промене међу вредностима матрице A^{-1}. Степен лоше условљености исказује се кондиционим бројем матрице $\kappa(A) = \|A\|\;\|A^{-1}\|$.

Системи линеарних једначина $Ax = b$ који се појављују у пракси готово увек долазе са мање или више грешака у улазним подацима. Веома је битно проценити утицај тих грешака на грешку излазног резултата, тј. на решење система $x = A^{-1}b$.

Пример 3.5.11. У овом примеру размотрићемо случај утицај грешке улазног податка b на грешку у добијеном решењу x. Претпоставимо да је $\widetilde{b} = b - \varepsilon$, и да су решења два система дата са x и $\widetilde{x}$:

$$Ax = b, \qquad A\widetilde{x} = \widetilde{b}.$$

Проценићемо коефицијент пропорционалности релативних грешака

$$\delta x = \frac{\|x - \widetilde{x}\|}{\|x\|} \qquad \text{и} \qquad \delta b = \frac{\|b - \widetilde{b}\|}{\|b\|} = \frac{\|\varepsilon\|}{\|b\|}\,.$$

Приметимо да је $\|b\| = \|Ax\| \le \|A\|\;\|x\|$, па је $\dfrac{1}{\|A\|\;\|x\|} \le \dfrac{1}{\|b\|}\,.$

$$\begin{aligned}\delta x &= \frac{\|x - \widetilde{x}\|}{\|x\|} = \frac{\|A^{-1}b - A^{-1}\widetilde{b}\|}{\|x\|} = \frac{\|A^{-1}(b - \widetilde{b})\|}{\|x\|} = \frac{\|A^{-1}\varepsilon\|}{\|x\|}\\ &\le \frac{\|A^{-1}\|\;\|\varepsilon\|}{\|x\|} = \frac{\|A^{-1}\|\|A\|\;\|\varepsilon\|}{\|A\|\|x\|} \le \|A^{-1}\|\|A\|\frac{\|\varepsilon\|}{\|b\|} = \kappa(A)\;\delta b.\end{aligned}$$

Такође, на основу (3.52) важи да је

$$\begin{aligned}\delta x &= \frac{\|A^{-1}\varepsilon\|}{\|x\|} \ge \frac{\frac{1}{\|A\|}\|\varepsilon\|}{\|x\|} = \frac{\|\varepsilon\|}{\|A\|\|x\|} = \frac{\|\varepsilon\|}{\|A\|\|A^{-1}b\|}\\ &\ge \frac{\|\varepsilon\|}{\|A\|\|A^{-1}\|\;\|b\|} = \frac{1}{\kappa(A)}\;\delta b.\end{aligned}$$

Коначно,

$$\frac{1}{\kappa(A)}\frac{\|b - \widetilde{b}\|}{\|b\|} \le \frac{\|x - \widetilde{x}\|}{\|x\|} \le \kappa(A)\frac{\|b - \widetilde{b}\|}{\|b\|}. \tag{3.54}$$

Најприроднији начин да се провери квалитет добијеног резултата $\widetilde{x}$ за решење система $Ax = b$ јесте израчунавање резидуланог вектора

$$r = b - A\widetilde{x} = b - \widetilde{b}.$$

У случају $r = \theta$, јасно је да је $\widetilde{x}$ тачно решење система. Када је $r \neq \theta$, тада v није тачно решење и потребно је проверити број тачних цифара у вектору v. Користећи резултат (3.54) са $r = b - \widetilde{b}$, добијамо

$$\frac{1}{\kappa(A)}\frac{\|r\|}{\|b\|} \leq \frac{\|x - \widetilde{x}\|}{\|x\|} \leq \kappa(A)\frac{\|r\|}{\|b\|}.$$

Закључујемо да су резултати резидуала добра процена грешке апроксимације само у случају малих вредности кондиционог броја матрице. Због тога, за матрицу A кажемо да је добро условљена уколико је $\kappa(A)$ релативно мала вредност. За матрице изузетно велике вредности кондиционог броја кажемо да су слабо условљене.

3.6 Скаларни производ функција

Библиографија

[1] C. Rosant. *IPython Interactive Computing and Visualization Cookbook*. PACKT Publishing, Birmingham-Mumbai, 2014.

[2] C. Rosant. *Learning IPython for Interactive Computing and Data Visualization*. PACKT Publishing, Birmingham-Mumbai, 2015.

[3] P. Farrell. *Math Adventures with Python*. No Starch Press, San Francisco, 2019.

[4] I. Idris. *Learning NumPy Array*. PACKT Publishing, Birmingham-Mumbai, 2014.

[5] H.P. Langtangen. *A Primer on Scientific Programming with Python*. Springer Open, Springer Heidelberg Dordrecht London New York, 2011.

[6] S. Linge and H.P. Langtangen. *Programming for Computations - Python*. Springer Open, Springer Heidelberg Dordrecht London New York, 2020.

[7] R. Johansson. *Numerical Python*. Apress, Springer Science+Business Media New York, 2019.

[8] J. Unpingco. *Python for Probability, Statistics, and Machine Learning*. Springer, Springer International Publishing Switzerland, 2016.

[9] G. Strang. *Introduction to Linear Algebra*. Wellesley - Cambridge Press, Wellesley MA 02482 USA, 2009.

[10] G. Strang. *Linear Algebra and learning from Data*. Wellesley - Cambridge Press, Wellesley MA 02482 USA, 2019.

[11] M. Itskov. *Tensor Algebra and Tensor Analysis for Engineers*. Springer, Springer Berlin Heidelberg New York, 2007.

[12] Y. Liu, J. Liu, Z. Long, and C. Zhu. *Tensor Computation for Data Analysis*. Springer, Springer International Publishing Switzerland, 2022.

[13] C. Bocci and L. Chiantini. *An Introduction to Algebraic Statistics with Tensors*. Springer, Springer International Publishing Switzerland, 2019.

[14] Y. Panagakis, J. Kossaifi, G. Chrysos, J. Oldfield, M. Nicolaou, A. Anandkumar, and S. Zafeiriou. Tensor methods in computer vision and deep learning. *Proceedings of the IEEE*, 109(5):863–890, 2021.

[15] T. Kolda and B. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[16] P. Comon. Tensors : A brief introduction. *IEEE Signal Processing Magazine*, 31(3):44–53, 2014.

[17] D. Donoho. For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59(6):797–829, 2006.