MAJORIZACIJA VEKTORJEV V $\mathbb{R}^N$

ANA JULIJA PREŠEREN

Fakulteta za matematiko in fiziko
Univerza v Ljubljani

V članku je predstavljena majorizacija vektorjev, njena geometrijska interpretacija in uporaba v ekonomiji. Predstavljena je tudi povezava med majorizacijo in dvojno stohastičnimi matrikami, v zadnjem poglavju pa so zbrane in dokazane tri izjave, ki so majorizaciji ekvivalentne.

VECTOR MAJORIZATION IN $\mathbb{R}^N$

The article presents vector majorization, its geometric interpretation and application in economy. It also presents the connection between vector majorization and doubly stochastic matrices. Three statements, equivalent to vector majorization, are formulated and proven in the last chapter.

1. Uvod

Ko med seboj primerjamo dva vektorja iz $\mathbb{R}^n$, ki imata enako vsoto koordinat, opazimo, da so včasih koordinate enega izmed vektorjev bližje skupaj kot koordinate drugega vektorja. Majorizacija vektorjev se ukvarja z enakomernostjo porazdelitve koordinat vektorjev: če ima vektor x koordinate bližje skupaj kot vektor y (pri čemer imata vektorja enako vsoto koordinat), pravimo, da je x majoriziran z y.
Eden najbolj znanih problemov, ki ga lahko karakteriziramo z majorizacijo, je enakomernost porazdelitve premoženja. Če želimo ugotoviti, katera izmed dveh porazdelitev premoženja je enakomernejša, lahko to storimo z uporabo majorizacije. Majorizacija vektorjev se uporablja tudi v povezavi z majorizacijo matrik, Schurovimi konveksnimi in konkavnimi funkcijami, Hadamardovo neenakostjo (angl. *Hadamard's inequality*), verjetnostjo prenosa bolezni, statistično mehaniko in še na mnogih drugih področjih, vendar pa ta niso zajeta v članku. Bralec lahko več primerov uporabe majorizacije najde v članku [1].

2. Osnovne definicije

Uvedimo dve oznaki, ki bosta nastopali v definiciji majorizacije in kasnejših izrekih.

Naj bo $x = (x_1, \dots, x_n)$ vektor iz $\mathbb{R}^n$. Naj bosta $x^\downarrow$ in $x^\uparrow$ vektorja, dobljena s preureditvijo koordinat vektorja x v padajoči oziroma naraščajoči vrstni red. Torej, če je $x^\downarrow = (x_1^\downarrow, \dots, x_n^\downarrow)$, potem je $x_1^\downarrow \geq \dots \geq x_n^\downarrow$. Podobno, če je $x^\uparrow = (x_1^\uparrow, \dots, x_n^\uparrow)$, potem je $x_1^\uparrow \leq \dots \leq x_n^\uparrow$.

Opazimo, da velja:
$$x_j^\uparrow = x_{n-j+1}^\downarrow, \quad 1 \leq j \leq n. \tag{1}$$

Definicija 1. Naj bosta $x, y \in \mathbb{R}^n$. Pravimo, da je x ***majoriziran*** z y, kar označimo $x \prec y$, če velja:
$$\sum_{j=1}^{k} x_j^\downarrow \leq \sum_{j=1}^{k} y_j^\downarrow, \quad 1 \leq k \leq n \tag{2}$$

in
$$\sum_{j=1}^{n} x_j^\downarrow = \sum_{j=1}^{n} y_j^\downarrow. \tag{3}$$

Primer 1. Če so $x_i \geq 0$ in $\sum_{i=1}^{n} x_i = 1$, potem velja:

$$\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \prec (x_1, \ldots, x_n) \prec (1, 0, \ldots, 0).$$

Primer 2. Vrstni red koordinat v zapisu x in y ne vpliva na majorizacijo, saj koordinate vedno uredimo po velikosti. Tako je na primer $(1, 2) \prec (0, 3)$ ekvivalentno $(2, 1) \prec (3, 0)$.

Primer 3. Za vektorje z n komponentami velja:

$$\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \prec \left(\frac{1}{n-1}, \ldots, \frac{1}{n-1}, 0\right) \prec \cdots \prec \left(\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0\right) \prec (1, 0, \ldots, 0)\,.$$

Včasih je prikladnejša uporaba $x^{\uparrow}$ kot $x^{\downarrow}$, zato želimo pogoje iz definicije majorizacije opisati še z $x^{\uparrow}$. Z upoštevanjem (1) lahko zapišemo:

$$\sum_{j=1}^{k} x_j^{\uparrow} = \sum_{j=1}^{n} x_j - \sum_{j=1}^{n-k} x_j^{\downarrow}.$$

Torej je $x \prec y$, če in samo če velja:

$$\sum_{j=1}^{k} x_j^{\uparrow} \geq \sum_{j=1}^{k} y_j^{\uparrow}, \quad 1 \leq k \leq n \tag{4}$$

in

$$\sum_{j=1}^{n} x_j^{\uparrow} = \sum_{j=1}^{n} y_j^{\uparrow}. \tag{5}$$

Definicija 2. x je ***podmajoriziran*** z y, kar označimo $x \prec_w y$, če je izpolnjen pogoj (2).

Opomba 1. Ob odsotnosti pogoja (3) si (2) in (4) nista več ekvivalentna.

3. Lorenzova krivulja

Sedaj si poglejmo primer uporabe majorizacije v ekonomiji. Ogledali si bomo, kako lahko različne porazdelitve istega premoženja med seboj primerjamo z majorizacijo.

Lorenzova krivulja je grafična predstavitev porazdelitve premoženja, poimenovana po ameriškemu ekonomistu Maxu Lorenzu. Graf, ki prikazuje Lorenzovo krivuljo, na abscisni osi prikazuje odstotek populacije, na ordinatni osi pa odstotek skupnega premoženja dela populacije, glede na celotno premoženje. Denimo, da je pri vrednosti 0,45 na abscisni osi vrednost ordinate 0,14, kar pomeni, da ima 45 % prebivalstva 14 % vrednosti vsega premoženja.

Oglejmo si, kako lahko dve porazdelitvi premoženja med seboj primerjamo z majorizacijo. Naj ima populacija n posameznikov in z x_i označimo premoženje i-tega posameznika. Vrednosti premoženja posameznikov nato uredimo od najrevnejšega do najbogatejšega, kar označimo z $x_1^{\uparrow}, \ldots, x_n^{\uparrow}$. Sedaj narišemo graf iz točk $\left(\frac{k}{n}, \frac{S_k}{S_n}\right)$, $k = 0, \ldots, n$, kjer je $S_0 = 0$ in $S_k = \sum_{i=1}^{k} x_i^{\uparrow}$ vsota premoženja najrevnejših k posameznikov. Te točke sedaj povežemo z daljicami in dobimo Lorenzovo krivuljo, ki povezuje začetno točko (0,0) in končno točko (1,1).

Če je celotno premoženje enakomerno razporejeno med posameznike (vsak dobi $\frac{1}{n}$ premoženja), je graf kar premica od $(0, 0)$ do $(1, 1)$. Na sliki 1 to porazdelitev predstavlja krivulja A. V primeru neenakomerne porazdelitve se bo krivulja prav tako začela in končala v točkah $(0, 0)$ in $(1, 1)$, toda med točkama bo upognjena in bo ležala pod premico. Manjša upognjenost krivulje torej ponazarja enakomernejšo porazdelitev. Na sliki 1 krivulja B prikazuje enakomernejšo porazdelitev kot krivulja C.

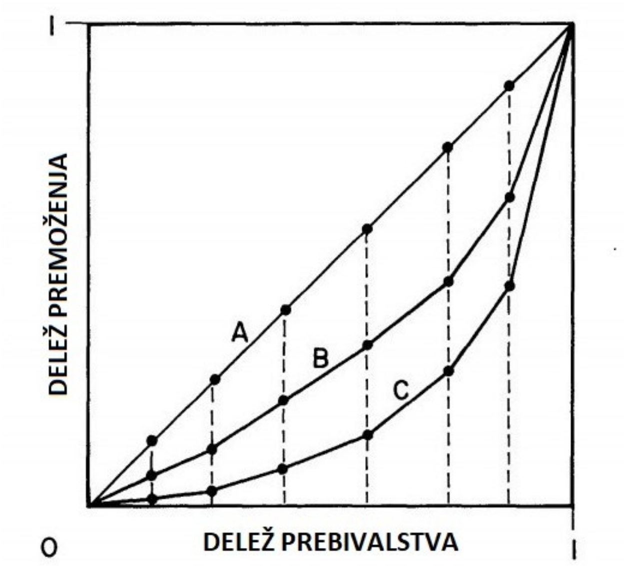
Kako lahko enakomernost dveh porazdelitev premoženja med seboj primerjamo z majorizacijo? Naj bo T skupno premoženje, $x_1, \ldots, x_n$ in $y_1, \ldots, y_n$ pa dve možni porazdelitvi premoženja T med n posameznikov. $(x_1, \ldots, x_n)$ predstavlja enakomernejšo porazdelitev premoženja kot $(y_1, \ldots, y_n)$, če in samo če velja:

$$\sum_{i=1}^{k} x_i^{\uparrow} \geq \sum_{i=1}^{k} y_i^{\uparrow}, \quad k = 1, \ldots, n$$

in seveda

$$\sum_{i=1}^{n} x_i^{\uparrow} = \sum_{i=1}^{n} y_i^{\uparrow} = T.$$

Kot lahko vidimo, sta to ravno enačbi (4) in (5), ki definirata, da je x majoriziran z y. Ker je mogoče relacijo majorizacije hitro izračunati, je ta metoda primerna za hitro primerjavo porazdelitev.



Slika 1. Lorenzova krivulja porazdelitve premoženja [2].

4. Geometrična interpretacija

Vemo že, kdaj za vektor x velja, da je majoriziran z y, ne vemo pa še, kako najti vse vektorje x, ki jih y majorizira. Izkaže se, da lahko take vektorje geometrično zelo lepo opišemo z uporabo naslednje definicije:

Definicija 3. ***Konveksna ogrinjača množice točk*** X je najmanjša konveksna množica, ki vsebuje X kot podmnožico.

Kako najti konveksno ogrinjačo množice X? Problem si predstavimo s primerom iz vsakdanjega življenja. Imamo nekaj žebljičkov, zapičenih v desko; to so točke iz X. Okoli njih razpremo elastiko, zajamemo vse žebljičke, nato pa jo spustimo, da se napne okoli njih. Oblika elastike predstavlja rob konveksne ogrinjače množice žebljičkov.

Za še boljše razumevanje pojma konveksne ogrinjače vpeljimo pojem konveksne kombinacije, ki se bo pojavil tudi kasneje v članku.

Definicija 4. Vsota vektorjev $\sum_{i=1}^{n} \lambda_i x_i$ se imenuje ***konveksna kombinacija vektorjev*** $x_1, \ldots, x_n$, če so vsi koeficienti λ_i nenegativni, njihova vsota pa je 1.

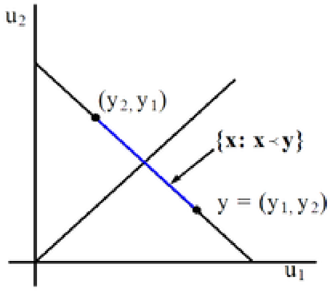
Sedaj si oglejmo, kako lahko konveksno ogrinjačo množice X opišemo s konveksnimi kombinacijami elementov iz množice X. To nam pove spodnji izrek. Bralec lahko njegov dokaz in več informacij o konveksni ogrinjači in konveksnih kombinacijah najde v [3].

Izrek 1. *Naj bo $X \subset \mathbb{R}^n$. Konveksna ogrinjača množice X je sestavljena iz vseh konveksnih kombinacij elementov iz X.*

Sedaj povežimo pojem majorizacije in konveksne ogrinjače. Sledeči izrek nam pove, kdaj za $x, y \in \mathbb{R}^n$ velja $x \prec y$. Izrek bo dokazan na koncu članka.

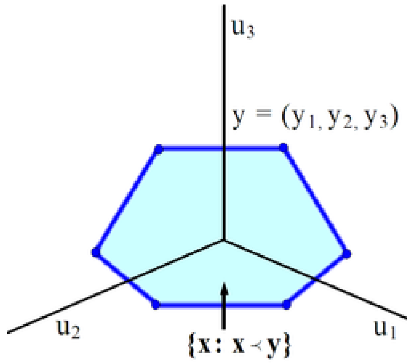
Izrek 2. *Za $x, y \in \mathbb{R}^n$ velja $x \prec y \Leftrightarrow x$ je vsebovan v konveksni ogrinjači vseh vektorjev, ki jih dobimo s permutacijo koordinat vektorja y.*

Primer 4. Slika 2 prikazuje konveksno ogrinjačo za $y = (3, 1)$ v 2D prostoru. Edini drugi vektor, ki ga pridobimo s permutacijo koordinat, je $y' = (1, 3)$. Konveksna ogrinjača teh dveh vektorjev je kar daljica med njima, torej velja $x \prec y$ natanko tedaj, ko x leži na tej daljici. Jasno, z daljico so označene točke, kjer je vsota koordinat konstanta in s tem izpolnjen pogoj (5), krajišča daljice pa so omejena tako, da je izpolnjen tudi pogoj (4). V večrazsežnem prostoru je popolnoma enako, le grafe predstavimo nekoliko težje.



Slika 2. Konveksna ogrinjača vektorjev (y_1, y_2) in (y_2, y_1), $y_1 \neq y_2$, v 2D [4].

Primer 5. Slika 3 prikazuje konveksno ogrinjačo vektorja $y = (y_1, y_2, y_3)$. Če so koordinate med seboj paroma različne, z njihovo permutacijo dobimo šest različnih vektorjev, konveksna ogrinjača teh vektorjev pa je šestkotnik med njimi. Velja: $x \prec y \Leftrightarrow x$ leži na šestkotniku. Podobno kot prej je ravnina šestkotnika ploskev, na kateri je vsota koordinat točk konstantna, robovi šestkotnika pa skrbijo za izpolnitev pogoja (4). V primeru, da sta dve komponenti vektorja y enaki, s permutacijo koordinat dobimo tri različne vektorje, njihova konveksna ogrinjača pa je trikotnik. Če so vse tri koordinate vektorja enake, s permutacijo teh seveda dobimo le en vektor, njegova konveksna ogrinjača pa je točka.



Slika 3. Konveksna ogrinjača vseh permutacij vektorja (y_1, y_2, y_3) v 3D [4].

5. Osnovne trditve

V tem poglavju bomo formulirali in dokazali nekaj osnovnih trditev o majorizaciji vektorjev. Bralec lahko še več trditev in njihove dokaze najde v članku [5].

Za začetek vpeljimo nekaj oznak, ki jih bomo potrebovali v nadaljevanju.
Za x, $y \in \mathbb{R}$ definiramo:

$$x \vee y = \max(x, y),$$
$$x^+ = x \vee 0.$$

Izrek 3. *Naj bosta x, $y \in \mathbb{R}^n$. Potem veljata naslednji ekvivalenci:*

$$x \prec_w y \iff \textit{za vsak}\, t \in \mathbb{R}\ \textit{velja}: \ \sum_{j=1}^{n} (x_j - t)^+ \leq \sum_{j=1}^{n} (y_j - t)^+ \tag{6}$$

$$x \prec y \iff \textit{za vsak}\, t \in \mathbb{R}\ \textit{velja}: \ \sum_{j=1}^{n} |(x_j - t)| \leq \sum_{j=1}^{n} |(y_j - t)| \tag{7}$$

Dokaz. Najprej bomo dokazali prvo ekvivalenco izreka.
Naj bo $x \prec_w y$. Če je $t \geq x_1^{\downarrow}$, potem je $(x_j - t)^+ = 0$ za vsak j. Torej ekvivalenca (6) velja za vsak $t \geq x_1^{\downarrow}$.
Naj bo $x_{k+1}^{\downarrow} \leq t \leq x_k^{\downarrow}$ za nek $1 \leq k \leq n$, kjer si $x_{n+1}^{\downarrow}$ predstavljamo kot $-\infty$. Potem velja:

$$\begin{aligned} \sum_{j=1}^{n} (x_j - t)^+ &= \sum_{j=1}^{k} (x_j^{\downarrow} - t) = \sum_{j=1}^{k} x_j^{\downarrow} - kt \\ &\leq \sum_{j=1}^{k} y_j^{\downarrow} - kt = \sum_{j=1}^{k} (y_j^{\downarrow} - t) \leq \sum_{j=1}^{n} (y_j^{\downarrow} - t)^+. \end{aligned}$$

Torej velja implikacija v desno, dokazati pa je potrebno še implikacijo v levo.
Predpostavimo torej, da velja:

$$\sum_{j=1}^{n} (x_j - t)^+ \leq \sum_{j=1}^{n} (y_j - t)^+ \text{ za vsak } t \in \mathbb{R}. \tag{8}$$

Ker predpostavka velja za za vsak $t \in \mathbb{R}$, mora veljati tudi v primeru $t = y_k^{\downarrow}$. Dobimo:

$$\sum_{j=1}^{n} (y_j - t)^+ = \sum_{j=1}^{k} (y_j^{\downarrow} - t) = \sum_{j=1}^{k} y_j^{\downarrow} - kt.$$

Velja pa tudi:

$$\sum_{j=1}^{k} x_j^{\downarrow} - kt = \sum_{j=1}^{k}(x_j^{\downarrow} - t) \leq \sum_{j=1}^{k}(x_j^{\downarrow} - t)^{+}$$
$$\leq \sum_{j=1}^{n}(x_j^{\downarrow} - t)^{+} = \sum_{j=1}^{n}(x_j - t)^{+}.$$

Iz predpostavke (8) torej sledi:

$$\sum_{j=1}^{k} x_j^{\downarrow} - kt \leq \sum_{j=1}^{n}(x_j - t)^{+} \leq \sum_{j=1}^{n}(y_j - t)^{+} = \sum_{j=1}^{k} y_j^{\downarrow} - kt. \tag{9}$$

Če si nato pogledamo začetni in končni izračun enačbe (9), dobimo:

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow},$$

kar je ravno definicija $x \prec_w y$.

Dokažimo še drugo ekvivalenco iz izreka.
Predpostavimo, da velja $x \prec y$. Če je $t \leq x_n^{\downarrow}$, potem je $(x_j - t) \geq 0$ za vsak j.

$$\sum_{j=1}^{n} |x_j - t| = \sum_{j=1}^{n}(x_j - t) = \sum_{j=1}^{n} x_j - nt = \sum_{j=1}^{n} y_j - nt$$
$$= \sum_{j=1}^{n}(y_j - t) \leq \sum_{j=1}^{n} |y_j - t|$$

Če je $t \geq x_1^{\downarrow}$, potem je $(x_j - t) \leq 0$ za vsak j.

$$\sum_{j=1}^{n} |x_j - t| = \sum_{j=1}^{n}(t - x_j) = nt - \sum_{j=1}^{n} x_j = nt - \sum_{j=1}^{n} y_j$$
$$= \sum_{j=1}^{n}(t - y_j) \leq \sum_{j=1}^{n} |t - y_j|$$
$$= \sum_{j=1}^{n} |y_j - t|$$

Preostane le še možnost, da je $x_{k+1}^{\downarrow} \leq t \leq x_k^{\downarrow}$ za nek k, $1 \leq k \leq n-1$.

$$\sum_{j=1}^{n} |x_j - t| = \sum_{j=1}^{k} |x_j^{\downarrow} - t| + \sum_{j=k+1}^{n} |x_j^{\downarrow} - t| = \sum_{j=1}^{k}(x_j^{\downarrow} - t) + \sum_{j=k+1}^{n}(t - x_j^{\downarrow})$$
$$= \sum_{j=1}^{k} x_j^{\downarrow} - kt - \sum_{j=k+1}^{n} x_j^{\downarrow} + (n-k)t$$
$$= \sum_{j=1}^{k} x_j^{\downarrow} - kt + (\sum_{j=1}^{k} x_j^{\downarrow} - \sum_{j=1}^{n} x_j^{\downarrow}) + (n-k)t$$

Sedaj upoštevamo, da velja $x \prec y$, torej je $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$.

$$\begin{aligned}
&\leq \left(\sum_{j=1}^{k} y_j^{\downarrow} - kt\right) + \left(\sum_{j=1}^{k} y_j^{\downarrow} - \sum_{j=1}^{n} y_j^{\downarrow} + (n-k)t\right) \\
&= \sum_{j=1}^{k} (y_j^{\downarrow} - t) + \left(- \sum_{j=k+1}^{n} y_j^{\downarrow} + (n-k)t\right) \\
&= \sum_{j=1}^{k} (y_j^{\downarrow} - t) + \sum_{j=k+1}^{n} (t - y_j^{\downarrow}) \leq \sum_{j=1}^{k} |y_j^{\downarrow} - t| + \sum_{j=k+1}^{n} |t - y_j^{\downarrow}| \\
&= \sum_{j=1}^{n} |y_j^{\downarrow} - t| = \sum_{j=1}^{n} |y_j - t|
\end{aligned}$$

Dokažimo še implikacijo v levo.
Ker predpostavka velja za vsak $t \in \mathbb{R}$, si najprej izberemo t, ki je večji od vseh koordinat vektorjev x in y. Potem velja:

$$\begin{aligned}
\sum_{j=1}^{n} |x_j - t| \leq \sum_{j=1}^{n} |y_j - t| &\Longleftrightarrow \sum_{j=1}^{n} (t - x_j) \leq \sum_{j=1}^{n} (t - y_j) \\
&\Longleftrightarrow \sum_{j=1}^{n} -x_j \leq \sum_{j=1}^{n} -y_j \Longleftrightarrow \sum_{j=1}^{n} x_j \geq \sum_{j=1}^{n} y_j.
\end{aligned}$$

Sedaj izberemo t, ki je manjši od vseh koordinat vektorjev x in y.

$$\sum_{j=1}^{n} |x_j - t| \leq \sum_{j=1}^{n} |y_j - t| \Longleftrightarrow \sum_{j=1}^{n} (x_j - t) \leq \sum_{j=1}^{n} (y_j - t) \Longleftrightarrow \sum_{j=1}^{n} x_j \leq \sum_{j=1}^{n} y_j$$

Iz teh dveh neenakosti sledi $\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j$.

Dokazati je potrebno še $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$ za vsak k. To je ravno definicija $x \prec_w y$.
Zgoraj smo že dokazali, da velja ekvivalenca (6), torej je dovolj dokazati, da za vsak $t \in \mathbb{R}$ velja:

$$\sum_{j=1}^{n} (x_j - t)^{+} \leq \sum_{j=1}^{n} (y_j - t)^{+}.$$

Opazimo, da za vsak vektor z velja: $|z| + z = 2z^{+}$. Torej lahko vektor z^{+} zapišemo kot $z^{+} = \frac{1}{2}|z| + \frac{1}{2}z$. Velja:

$$\begin{aligned}
\sum_{j=1}^{n} (x_j - t)^{+} &= \frac{1}{2} \sum_{j=1}^{n} |x_j - t| + \frac{1}{2} \sum_{j=1}^{n} (x_j - t) = \frac{1}{2} \sum_{j=1}^{n} |x_j - t| + \frac{1}{2} \sum_{j=1}^{n} x_j - \frac{1}{2} nt \\
&= \frac{1}{2} \sum_{j=1}^{n} |x_j - t| + \frac{1}{2} \sum_{j=1}^{n} (y_j - t) \leq \frac{1}{2} \sum_{j=1}^{n} |y_j - t| + \frac{1}{2} \sum_{j=1}^{n} (y_j - t) \\
&= \sum_{j=1}^{n} (y_j - t)^{+}.
\end{aligned}$$

$\square$

Posledica 4. *Če $x \prec y$ v $\mathbb{R}^n$ in $u \prec w$ v $\mathbb{R}^m$, potem velja $(x, u) \prec (y, w)$ v $\mathbb{R}^{n+m}$. Natančneje, $x \prec y$ velja natanko tedaj ko $(x, u) \prec (y, u)$ za vsak u.*

Dokaz. Želimo dokazati, da iz $x \prec y$ v $\mathbb{R}^n$ in $u \prec w$ v $\mathbb{R}^m$ sledi $(x,u) \prec (y,w)$ v $\mathbb{R}^{n+m}$. Označimo $x' = (x,u)$ in $y' = (y,w)$. Prejšnji izrek nam pove, da sta trditvi

- $x' \prec y'$ in
- $\sum_{j=1}^{n+m} |(x'_j - t)| \leq \sum_{j=1}^{n+m} |(y'_j - t)|$ za vsak $t \in \mathbb{R}$

ekvivalentni. Dokažimo zadnjo:

$$\begin{aligned}\sum_{j=1}^{n+m} |(x'_j - t)| &= \sum_{i=1}^{n} |(x_i - t)| + \sum_{j=1}^{m} |(u_j - t)| \\ &\leq \sum_{i=1}^{n} |(y_i - t)| + \sum_{j=1}^{m} |(w_j - t)| = \sum_{j=1}^{n+m} |(y'_j - t)|,\end{aligned}$$

kjer smo pri neenakosti upoštevali $x \prec y$ in $u \prec w$ ter ekvivalenco (7).

Dokažimo še, da velja $x \prec y \Leftrightarrow (x,u) \prec (y,u)$ za vsak u. Ker vedno velja $u \prec u$, iz $x \prec y$ sledi $(x,u) \prec (y,u)$ po zgornjem dokazu.

Dokažimo še, da iz $(x,u) \prec (y,u)$ za vsak u sledi $x \prec y$.
Izberemo si u, ki ima vse koordinate manjše od najmanjše koordinate x in y. Če označimo $a = (x,u)$ in $b = (y,u)$, nastopajo v $a^{\downarrow}$ in $b^{\downarrow}$ koordinate u povsem na koncu. Zato za $k \leq n$ očitno velja:

$$\sum_{j=1}^{k} a_j^{\downarrow} \leq \sum_{j=1}^{k} b_j^{\downarrow} \Longleftrightarrow \sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}.$$

Očitno velja tudi $\sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{n} y_j^{\downarrow}$.

□

6. Povezava z matrikami

Izkaže se, da je majorizacija vektorjev tesno povezana z dvojno stohastičnimi matrikami, katerih lastnosti so opisane v sledeči definiciji.

Definicija 5. Matriko $A = (a_{ij})$ velikosti $n \times n$ imenujemo ***dvojno stohastična***, kadar velja:

$$a_{ij} \geq 0 \quad \text{za vsaka} \quad i,j \tag{10}$$

$$\sum_{i=1}^{n} a_{ij} = 1 \quad \text{za vsak} \quad j \tag{11}$$

$$\sum_{j=1}^{n} a_{ij} = 1 \quad \text{za vsak} \quad i \tag{12}$$

Opomba 2. Produkt dvojno stohastičnih matrik je dvojno stohastičen.

Izrek 5. *Matrika A je dvojno stohastična natanko tedaj, ko je $Ax \prec x$ za vsak vektor x.*

Dokaz. Najprej dokažimo implikacijo v levo. Naj torej velja $Ax \prec x$ za vsak vektor x.

Najprej za vektor x izberemo $e = (1, 1, \ldots, 1)$. Če za nek vektor y velja $y \prec e = (1, \ldots, 1)$, potem je gotovo $y = e$, saj mora biti skupna vsota n koordinat enaka n in velja $y_i \leq 1$ za vsak i, $1 \leq i \leq n$. Edina možna rešitev je, da za vsak i velja $y_i = 1$. Torej je $Ae = e$ in velja (12).

Nato obravnavamo vektor $x = e_j$ za katerega velja $x_j = 1, x_i = 0$ za $i \neq j$. Vemo, da velja $Ae_j = (a_{1j}, a_{2j}, \ldots, a_{nj}) \prec e_j$. Torej je $\sum_{i=1}^n a_{ij} = 1$ za vsak j in velja (11).
Vidimo, da so vsi členi $a_{ij} \geq 0$, saj relacija $a \prec b$ pove, da je $\min_i a_i \geq \min_i b_i$. V našem primeru to pomeni $\min_i a_i \geq 0$ za vsak j oziroma $a_{ij} \geq 0$ za vsaka i, j.
Torej je matrika A dvojno stohastična.

Dokažimo še implikacijo v desno.
Naj bo A dvojno stohastična matrika in naj bo $y = Ax$. Pri dokazu $y \prec x$ lahko brez izgube splošnosti predpostavimo, da so kooridinate vektorjev x in y urejene v padajočem vrstnem redu. Opazimo, da za vsak k, $1 \leq k \leq n$ velja:

$$\sum_{j=1}^{k} y_j = \sum_{j=1}^{k} \sum_{i=1}^{n} a_{ji} x_i.$$

Vpeljimo oznako $t_i = \sum_{j=1}^k a_{ji}$. Potem velja $0 \leq t_i \leq 1$ in $\sum_{i=1}^n t_i = k$, saj so a_{ji} koeficienti matike A, A pa je dvojno stohastična.
Dokažimo, da velja $\sum_{j=1}^k y_j \leq \sum_{j=1}^k x_j$ za $1 \leq k \leq n$, v primeru $k = n$ pa mora veljati enačaj.

$$\begin{aligned}\sum_{j=1}^{k} y_j - \sum_{j=1}^{k} x_j &= \sum_{i=1}^{n} t_i x_i - \sum_{i=1}^{k} x_i = \sum_{i=1}^{n} t_i x_i - \sum_{i=1}^{k} x_i + (k - \sum_{i=1}^{n} t_i) x_k \\ &= \sum_{i=1}^{k} (t_i - 1)(x_i - x_k) + \sum_{i=k+1}^{n} t_i (x_i - x_k) \leq 0\end{aligned}$$

Pri neenakosti smo upoštevali:

- $t_i \leq 1$,
- $x_k \leq x_i$ za $i \leq k$,
- $t_i \geq 0$,
- $x_k \geq x_i$ za $i > k$.

V primeru $k = n$ mora veljati enačaj:

$$\sum_{j=1}^{n} y_j = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} x_i = \sum_{i=1}^{n} x_i,$$

saj je A dvojno stohastična, torej je vsota posameznega stolpca enaka 1.

□

Pred formulacijo glavnega izreka, ki nam poda ekvivalenco med majorizacijo in dvojnostohastičnimi matrikami, moramo uvesti še pojem T-transformacije.

Definicija 6. Linearna preslikava T se imenuje ***T-transformacija***, če obstaja t, $0 \leq t \leq 1$, in indeksa j, k, da velja:

$$Ty = (y_1, \ldots, y_{j-1}, ty_j + (1-t)y_k, y_{j+1}, \ldots, (1-t)y_j + ty_k, y_{k+1}, \ldots, y_n)$$

Potem velja $Ty \prec y$ za vsak y.

Matrika T-transformacije ima obliko: $T = \lambda I + (1-\lambda)Q,\ \ 0 \leq \lambda \leq 1$, kjer je Q permutacijska matrika, ki le zamenja dve koordinati.

Lema 6. *Če $x \prec y$, potem lahko x dobimo iz y s končnim številom T-transformacij.*

Dokaz. Permutacijska matrika Q, ki le zamenja dve koordinati vektorja, je seveda T-transformacija. Vsaka permutacijska matrika je produkt matrik, ki zamenjajo po dve koordinati. Permutacijska matrika je torej produkt T-transformacij, zato lema velja, če imata x in y enake koordinate, morda le v drugem vrstnem redu.

Sedaj predpostavimo, da x in y nimata vseh koordinat enakih in da BŠS[1] velja $x_1 \geq x_2 \geq \ldots \geq x_n$ in $y_1 \geq y_2 \geq \ldots \geq y_n$.

Naj bo j največji indeks, za katerega velja $x_j < y_j$, in k najmanjši indeks, večji od j, za katerega velja $x_k > y_k$. Tak par j, k mora obstajati, saj mora največji indeks i, za katerega velja $x_i \neq y_i$, zadostiti pogoju $x_i > y_i$. Za izbrana j, k velja:

$$y_j > x_j \geq x_k > y_k. \tag{13}$$

Naj bodo:

$$\begin{aligned} \delta &= \min(y_j - x_j, x_k - y_k), \\ 1 - \lambda &= \frac{\delta}{(y_j - y_k)}, \\ y^* &= (y_1, \ldots, y_{j-1}, y_j - \delta, y_{j+1}, \ldots, y_{k-1}, y_k + \delta, y_{k+1}, \ldots, y_n). \end{aligned}$$

Iz (13) sledi $0 < \lambda < 1$, y^* pa lahko zapišemo v obliki:

$$y^* = \lambda y + (1-\lambda)(y_1, \ldots, y_{j-1}, y_k, y_{j+1}, \ldots, y_{k-1}, y_j, y_{k+1}, \ldots, y_n).$$

Torej je $y^* = Ty$ za $T = \lambda I + (1-\lambda)Q$, kjer Q zamenja j-te in k-te koordinate. Vidimo tudi, da velja $y^* \prec y$, saj je T dvojno stohastična.

Dokažimo še, da velja $x \prec y^*$. Upoštevamo, da je prvih $j-1$ koordinat vektorjev y in $y*$ enakih, ter da y majorizira x. Dobimo spodnjo formulo:

$$\sum_{i=1}^{\nu} y_i^* = \sum_{i=1}^{\nu} y_i \geq \sum_{i=1}^{\nu} x_i, \quad \nu = 1, \ldots, j-1.$$

Nato preverimo, da velja $y_j^* \geq x_j$. Če velja $\delta = y_j - x_j$, dobimo $y_j^* = y_j - \delta = x_j$. Če pa velja $\delta = x_k - y_k$, iz definicije δ sledi, da je $y_j - x_j \geq x_k - y_k$. Torej je $y_j^* \geq x_j$, saj od y_j odštejemo kvečjemu manj kot v prvem primeru.

Iz predpostavk na indeksa j in k sledi, da je $y_i = x_i$ za $i = j+1, \ldots, k-1$. Opazimo tudi, da

[1] Brez škode za splošnost

za $i = j+1, \ldots, k-1$ velja $y_i^* = y_i$. Torej velja $y_i^* = x_i$ za $i = j+1, \ldots, k-1$.

Sedaj upoštevamo, da se pri seštevanju prvih k ali več koordinat vektorja y^* delti odštejeta. Zato je vsota prvih k ali več koordinat vektorja y^* enaka vsoti istega števila koordinat vektorja y. Upoštevamo še, da y majorizira x. Dobimo spodnjo formulo:

$$\sum_{i=1}^{\nu} y_i^* = \sum_{i=1}^{\nu} y_i \geq \sum_{i=1}^{\nu} x_i, \quad \nu = k, \ldots, n.$$

Vemo že, da velja $y^* \prec y$ in $x \prec y$, iz česar sledi:

$$\sum_{i=1}^{n} y_i^* = \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i.$$

Na podlagi zgornjih ugotovitev tako lahko zaključimo, da velja $x \prec y^*$.

Za katerakoli vektorja u, v, naj bo $d(u, v)$ število neničelnih razlik $u_i - v_i$.
Ker velja $y_j^* = x_j$, če je $\delta = y_j - x_j$, in $y_k^* = x_k$, če je $\delta = x_k - y_k$, sledi, da je $d(x, y^*) \leq d(x, y) - 1$. Torej lahko x dobimo iz y s končnim številom T-transformacij.

□

Opomba 3. Če $x \prec y$, potem lahko x dobimo iz y z uporabo največ $n-1$ T-transformacij.
To sledi iz zgornjega dokaza, saj je $d(u, v) \leq n$ in $d(u, v) \neq 1$ (v tem primeru bi namreč veljalo $\sum_{i=1}^{n} u_i \neq \sum_{i=1}^{n} v_i$).

Z uporabo leme lahko sedaj dokažemo osnovni izrek majorizacije.

Izrek 7. ***Hardy, Littlewood, Polya***
Za $x,\ y \in \mathbb{R}^n$ velja $x \prec y \iff x = Ay$ za neko dvojno stohastično matriko A.

Dokaz. Implikacija v levo je že dokazana pri Izreku 5. Dokažimo še implikacijo v desno. Predpostavimo, da $x \prec y$. T-transformacije so dvojno stohastične, njihov produkt pa je prav tako dvojno stohastičen. Iz leme vemo: $x \prec y \iff x = T_1 T_2 ... T_k y$, kjer so T_i T-transformacije, zgoraj pa smo razmislili, da je $T_1 T_2 ... T_k = A$ za neko dvojno stohastično matriko A, kar smo želeli dokazati.

□

Opomba 4. Matrika A iz izreka ni enolično določena.

Za konec si oglejmo še nekaj izjav, ki so ekvivalentne $x \prec y$.

Izrek 8. *Za $x, y \in \mathbb{R}^n$ so naslednje izjave ekvivalentne.*

1. *$x \prec y$*
2. *x lahko dobimo iz y s končnim številom T-transformacij.*
3. *x je vsebovan v konveksni ogrinjači vseh vektorjev, ki jih dobimo s permutacijo koordinat vektorja y.*
4. *$x = Ay$ za neko dvojno stohastično matriko A.*

Dokaz. Izrek bomo dokazali s serijo implikacij:

1. (1) $\Rightarrow$ (2):
   Že dokazano v Lemi 6.

2. (2) $\Rightarrow$ (3):
   T-transformacija je konveksna kombinacija identitete in neke permutacije. Kompozitum takih preslikav je torej konveksna kombinacija permutacij. Torej velja (3).

3. (3) $\Rightarrow$ (4):
   Če ima vektor le zamenjan vrstni red komponent vektorja y, ga lahko zapišemo kot Py, kjer je P permutacijska matrika. Konveksno kombinacijo takih vektorjev lahko potem izrazimo kot Ay, kjer je A konveksna kombinacija permutacijskih matrik, torej dvojno stohastična.

4. (4) $\Rightarrow$ (1):
   Že dokazano v Izreku 5.

$\square$

7. Zaključek

Kot smo videli, je majorizacija zelo preprost matematični pojem, prav ta preprostost pa omogoča njeno uporabo na mnogih področjih. Ker je mogoče relacijo majorizacije hitro izračunati, je ta metoda primerna za ocenjevanje optimalnih rešitev v problemih, ki vsebujejo diskretne porazdelitve, na primer za hitro primerjavo porazdelitev premoženja. Bralec lahko znanje o majorizaciji poglobi in razširi z branjem dela [6], ki je bilo tudi glavni vir za ta članek.

LITERATURA

[1] Barry C. Arnold. *Majorization: Here, there and everywhere*, Statist. Sci. 22 (2007), no. 3, 407–413. MR 2416816

[2] Albert W. Marshall and Ingram Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York, 1979.

[3] Marko Obid. *Konveksne funkcije* Pedagoška fakulteta, Univerza v Ljubljani, 2018.

[4] *https://en.wikipedia.org/wiki/Majorization*, dostopano 31. 3. 2021

[5] T. Ando. *Linear Algebra and its Applications: Majorization, Doubly Stochastic Matrices, and Comparison of Eigenvalues.* Elsevier Science Publishing Co., New York, 1989.

[6] Rajendra Bhatia. *Matrix Analysis.* S. Axler, Springer-Verlag, New York, 1997.