

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Matematika – 1. stopnja

Gal Kuhar

Problem Maksimalne Korelacije

Osnutek dela diplomskega seminarja

Mentor: Bor Plestenjak

Ljubljana, 2022

1. UVOD

Projektna naloga spada na področje numeričnih metod. Problem maksimalne korelacije je problem, ki naravno nastane iz statistične analize več spremenljivk, ko iščemo linearno kombinacijo slučajnih spremenljivk, ki ima maksimalno korelacijo z drugo linearno kombinacijo slučajnih spremenljivk.

Cilj tega osnutka je predstaviti problem maksimalne korelacije, njegovo statistično ozadje in izpeljavo problema iz statističnih podatkov.

2. PREDSTAVITEV PROBLEMA MAKSIMALNE KORELACIJE

Definicija 2.1. Problem maksimalne korelacije

Dana je simetrična pozitivno definitna matrika $A \in \mathbb{R}^{n \times n}$, zapisana v bločni obliki

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix},$$

kjer blok $A_{ij} \in \mathbb{R}^{n_i \times n_j}$ in $\sum_{i=1}^m n_i = n$

Pri problemu maksimalne korelacije (MCP) iščemo maksimum $x^T A x$ po vseh vektorjih oblike $x = [x_1^T \ \dots \ x_m^T]^T$ kjer je $x_i \in \mathbb{R}^{n_i}$ in $\|x_i\|_2 = 1$.

Preko metode Lagrangevih večkratnikov lahko ugotovimo, da je zadosten pogoj za lokalni ekstrem obstoj skalarjev $\lambda_1, \dots, \lambda_m$, da velja:

$$\begin{aligned} A_{11}x_1 + \cdots + A_{1m}x_m &= \lambda_1 x_1 \\ &\vdots \\ A_{m1}x_1 + \cdots + A_{mm}x_m &= \lambda_m x_m \end{aligned}$$

To pravzaprav posplošitev standardnega problema lastnih vrednosti

3. IZPELJAVA

V statistiki pogosto naletimo na situacije, kjer hočemo ugotoviti, kako so med sabo povezane slučajne spremenljivke. Mero medsebojne povezanosti imenujemo korelacija.

Pri delu s slučajnimi vektorji nas pogosto kako posamezne spremenljivke nekega vektorja vplivajo na spremenljivke drugih. To matematično modeliramo kot linearno kombinacijo spremenljivk vektorja, ki ima maksimalno korelacijo z linearno kombinacijo spremenljivk drugega vektorja.

V tem razdelku bom naredil izpeljavo problema maksimalne korelacije za dva slučajna vektorja.

Imamo slučajna vektorja:

$$\bar{X}_1 = [X_{11} \ \dots \ X_{1n_1}] \text{ in } \bar{X}_2 = [X_{21} \ \dots \ X_{2n_2}]$$

ter pripadajoča vzorca k poskusov:

$$\bar{V}_1 = [V_1 \ \dots \ V_{n_1}] = \begin{bmatrix} v_{11}^1 & \cdots & v_{n_1 1}^1 \\ \vdots & \ddots & \vdots \\ v_{1k}^1 & \cdots & v_{n_1 k}^1 \end{bmatrix} \text{ in } \bar{V}_2 = [V_2 \ \dots \ V_{n_2}] = \begin{bmatrix} v_{11}^2 & \cdots & v_{n_2 1}^2 \\ \vdots & \ddots & \vdots \\ v_{1k}^2 & \cdots & v_{n_2 k}^2 \end{bmatrix}$$

Poskušamo najti taka vektorja $b_1 \in \mathbb{R}^{n_1}$ ter $b_2 \in \mathbb{R}^{n_2}$, da maksimiziramo korelacijski koeficient:

$$r = \text{corr}(X_1 b_1, X_2 b_2)$$

Brez škode za splošnost lahko predpostavimo, da je povprečna vrednost vsakega stolpca v vzorcu enaka 0.

Potem lahko definiramo kovariančno matriko

$$\Delta = [\bar{V}_1 \quad \bar{V}_2]^T [\bar{V}_1 \quad \bar{V}_2] = \begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix}$$

Δ je očitno simetrična in pozitivno definitna matrika.

Definirajmo še matriko

$$B = \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$$

Sedaj definiramo novo slučajno spremenljivko kot linearno kombinacijo slučajnih spremenljivk $\bar{X}_i b_i$ in njeno pripadajočo vzorčno matriko

$$Z_i = \bar{V}_i b_i \text{ in } Z = [Z_1 \quad Z_2]$$

Potem dobimo novo kovariančno matriko

$$\Omega = Z^T Z = \begin{bmatrix} Z_1^T Z_1 & Z_1^T Z_2 \\ Z_2^T Z_1 & Z_2^T Z_2 \end{bmatrix} = B^T \Delta B = \begin{bmatrix} b_1^T \Delta_{11} b_1 & b_1^T \Delta_{12} b_2 \\ b_2^T \Delta_{21} b_1 & b_2^T \Delta_{22} b_2 \end{bmatrix}$$

Iz tega lahko sedaj razberemo korelacijski koeficient

$$r = \frac{Z_1^T Z_2}{\sqrt{Z_1^T Z_1} \sqrt{Z_2^T Z_2}} = \frac{b_1^T \Delta_{12} b_2}{\sqrt{b_1^T \Delta_{11} b_1} \sqrt{b_2^T \Delta_{22} b_2}}$$

Tega poskušamo maksimizirati, torej iščemo

$$\text{Maksimum } b^T \Delta b$$

$$\text{Pri pogojih } b_1^T \Delta_{11} b_1 = 1$$

$$b_2^T \Delta_{22} b_2 = 1$$

$$\text{kjer } b = [b_1^T \quad b_2^T]^T$$

Problem pri reševanju tega sta pogoja $b_i^T \Delta_{ii} b_i = 1$, saj tega ne moremo hitro rešiti.

Ta pogoja bi radi spremenili v $x_i^T x_i = 1$, saj je ta računsko mnogo lažji. To naredimo tako, da Δ_{ii} razcepimo.

Ker je vsak Δ_{ii} simetrična pozitivno definitna matrika, zanjo obstaja razcep Choleskega:

$$\Delta_{ii} = C_i^T C_i$$

, ki razdeli Δ_{ii} na dve zgornjetrikotni matriki.

Potem lahko definiramo

$$x_i = C_i b_i \text{ in } A = (C^{-1})^T \Delta C^T$$

Ker sta matriki C_i zgornjetrikotni, nam to omogoča da iz x_i in C_i hitro izračunamo b_i , saj je rečunanje rešitev sistema $x_i = C_i b_i$ trivialno.

S tem dobimo korelacijski koeficient

$$r = \frac{x^T A x}{\sqrt{x_1^T x_1} \sqrt{x_2^T x_2}}$$

kjer $x = \begin{bmatrix} x_1^T & x_2^T \end{bmatrix}^T$

To maksimiziramo in dobimo problem maksimalne korelacije.

Maksimum $x^T A x$

Pri pogojih $x_1^T x_1 = 1$

$x_2^T x_2 = 1$

4. NUMERIČNO REŠEVANJE

Na problem maksimalne korelacije lahko gledamo kot posplišitev potenčne metode za iskanje lastnih vrednosti matrik.

$$\begin{array}{llll} \text{Izberemo nek začetni vektor} & x_0 & = & \begin{bmatrix} x_{01} & \dots & x_{0m} \end{bmatrix} \\ \text{za } k & = & & 1 : N \\ & \text{za } i & = & 1 : m \\ & & y_{ki} & = Ax_{(k-1)i} \\ & & x_{ki} & = \frac{y_{ki}}{\|y_{ki}\|} \end{array}$$

5. PROBLEMI Z METODO IN IZBOLJŠAVE

Metodo je prvi predstavil Horst leta 1961. Kljub uspešni ekperimentalni uporabi, konvergenčnost metode dolgo časa ni bila matematično dokazana. Šele leta 1993 sta Chu in Watterson pokazala da je metoda res lokalno konvergentna ter nato leta 2015 Liu in Wang dokazala, da je metoda ob pravilni izbiri začetne točke globalno konvergentna.

Potenčna metoda je le ena izmed metod za iskanje lastnih vrednosti matrik. Obstajajo posplošitve drugih metod, naprimer Gauss-Seidelova in P-SOR metoda, ki konvergirajo hitreje kot Horstova.

6. DODATNO

V okviru diplomske naloge, bodo omenjeni algoritmi implementirani v Matlabu in preizkušeni na numeričnih primerih.

LITERATURA

- [1] M. T. Chu in J. L. Watterson, *On multivariate eigenvalue problem: I. Algebraic theory and power method.*, SIAM J. Sci. Comput. **14** (2011) 1089–1006.