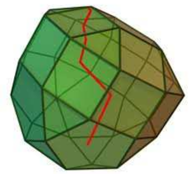
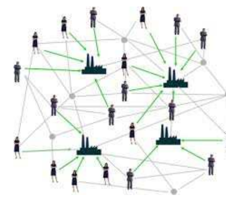
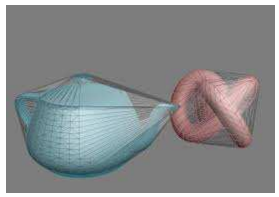
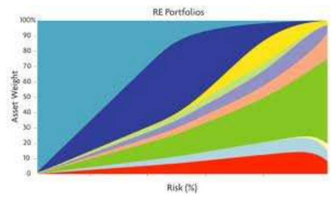
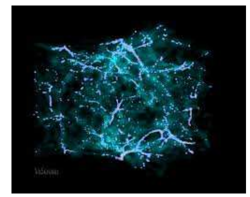
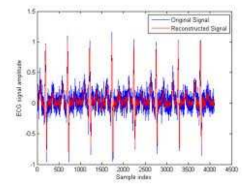
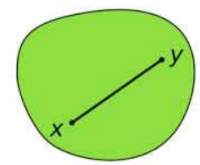


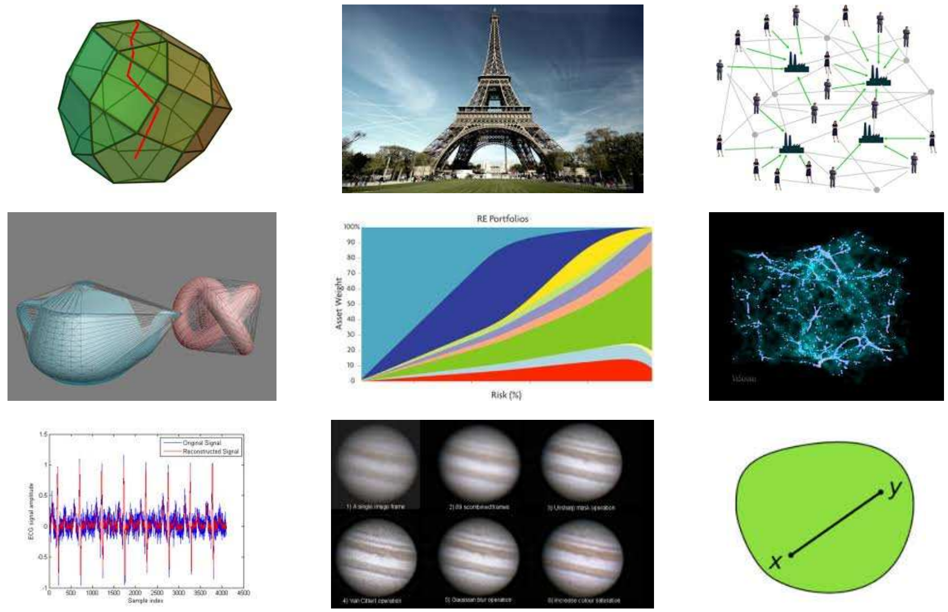






Konvexná optimalizácia - letný semester 2018

M. Trnovská, KAMŠ, FMFI UK



Téma 7: Aplikácie kónického lineárneho programovania

Optimalizácia portfólia s pravdepodobnostnými ohraničeniami

- premenná $x \in \mathbb{R}^n$; x_i je podiel aktíva i;

- vektor cien p, $p \sim \mathcal{N}(\bar{p}, \Sigma)$;

- výnos portfólia r, $r \sim \mathcal{N}(\bar{r}, \sigma^2)$, $\bar{r} = \bar{p}^T x$, $\sigma^2 = x^T \Sigma x$

Úloha maximalizácie priemerného výnosu s ohraničením v tvare pravdepodobnosti:

$$\begin{array}{ll} \max & \bar{p}^T x \\ & P(r \leq \alpha) \leq \beta, \\ & \Sigma_{i=1}^n x_i = 1, x \geq 0. \end{array} \tag{1}$$

Optimalizácia portfólia s pravdepodobnostnými ohraničeniami

Ohraničenie $P(r \leq \alpha) \leq \beta$ je ekvivalentné

$$P\left(\frac{r-\bar{r}}{\sigma} \leq \frac{\alpha-\bar{r}}{\sigma}\right) \leq \beta,$$

pričom $\frac{r-\bar{r}}{\sigma} \sim \mathcal{N}(0,1)$.

Vyjadrenie pomocou kumulatívnej distribučnej funckie $\phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$:

$$P\left(\frac{r-\bar{r}}{\sigma} \leq \frac{\alpha-\bar{r}}{\sigma}\right) = \phi\left(\frac{\alpha-\bar{r}}{\sigma}\right).$$

Optimalizácia portfólia s pravdepodobnostnými ohraničeniami

Úloha je ekvivalentná:

$$\begin{array}{ll} \max & \bar{p}^T x \\ & \alpha - \bar{p}^T x \leq \phi^{-1}(\beta)\sqrt{x^T \Sigma x}, \\ & \Sigma_{i=1}^n x_i = 1, x \geq 0. \end{array} \tag{2}$$

- $\beta < \frac{1}{2}$ - konvexná SOCP úloha

- $\beta = \frac{1}{2}$ - úloha LP

- $\beta > \frac{1}{2}$ - nekonvexná úloha

Optimalizácia portfólia s pravdepodobnostnými ohraničeniami

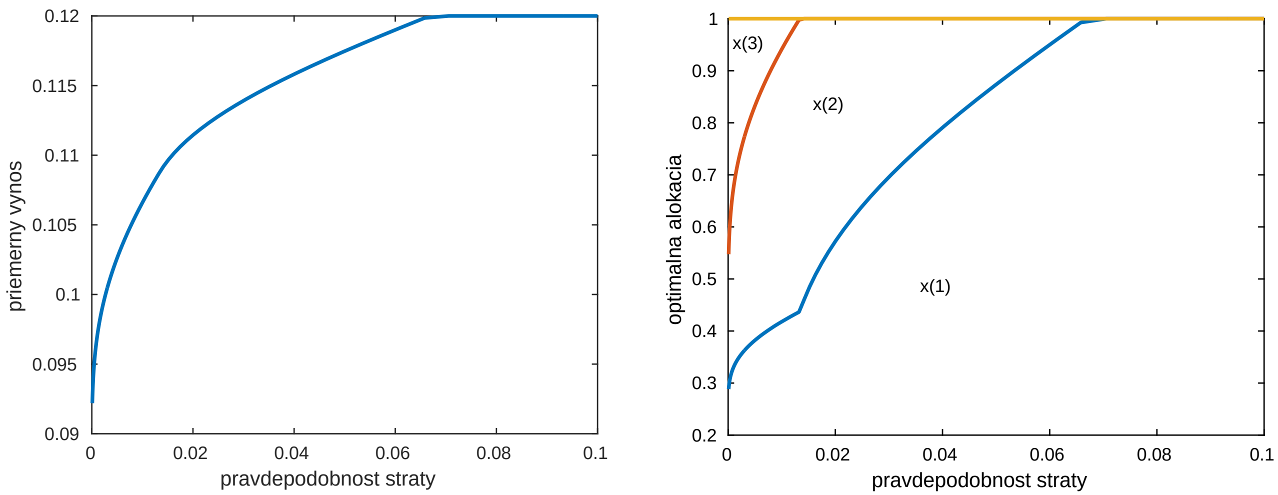
Príklad.

Zvoľme dáta

$$\bar{p} = \begin{pmatrix} 0.12 \\ 0.10 \\ 0.07 \\ 0.03 \end{pmatrix}, \ \Sigma = 0.0001 \begin{pmatrix} 64 & 8 & -11 & 0 \\ 8 & 25 & 0 & 0 \\ -11 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ \alpha = 0.$$

Zaujíma nás optimálne portfólio, ak hodnota pravdepodobnosti straty je $\beta \in [10^{-4}, 10^{-1}]$

Optimalizácia portfólia s pravdepodobnostnými ohraničeniami



Vľavo: trade-off medzi pravdepodobnosťou straty a priemerným výnosom

Vpravo: optimálna alokácia portfólia

Optimalizácia portfólia s pravdepodobnostnými ohraničeniami

```
beta=logspace(-4, -1,100);
for i=1:100
    cvx_begin quiet
        variable x(4);
        maximize (p'*x)
        subject to
           p'*x+norminv(beta(i))*norm(Sigma^(1/2)*x,2)>=0;
           sum(x)==1;
           x>=0;
    cvx_end
    ER(i)=p'*x; X(:,i)=x;
    v1(i)=X(1,i); v2(i)=v1(i)+X(2,i); v3(i)=v2(i)+X(3,i);
end;
subplot(1,2,1)
plot(beta,ER,'LineWidth',2)
xlabel('pravdepodobnost straty'); ylabel('priemerny vynos');
subplot(1,2,2)
plot(beta,v1,beta,v2,beta,v3,'LineWidth',2)
xlabel('pravdepodobnost straty'); ylabel('optimalna alokacia');
```

Optimálny návrh experimentov

Lineárny regresný model

$$y(x) = f(x)^T \theta + \varepsilon,$$

- $y(x)$ je vektor pozorovaní,

- $x \in X$, $X = \{x_1, \ldots, x_n\}$ je množina reprezentujúca možné experimentálne podmienky (body merania)

- $f : X \to \mathbb{R}^m$ je vektor známych regresných funkcií,

- $\theta \in \mathbb{R}^m$ je vektor neznámych parametrov

- ε - vektor náhodných chýb, $\varepsilon \sim \mathcal{N}(0, I)$.

Optimálny návrh experimentov

- Chceme odhadnúť vektor θ z čiastkových experimentov alebo meraní v bodoch z X.

- Návrh experimentu veľkosti N je postupnosť N bodov z X, v ktorých budú vykonané merania.

- Označme N_i počet meraní v bode x_i, t. j. $N_1 + \cdots + N_n = N$. V štatistike sa zvyčajne definuje **exaktný dizajn** ako diskrétna miera ξ, kde $\xi(x_i) = \frac{N_i}{N}$, $x_i \in X$.

Optimálny návrh experimentov

- Základná otázka v optimálnom navrhovaní experimentov: Ako optimálne rozmiestniť N meraní do potenciálnych bodov z X, teda ako vybrať hodnoty $N_1, \ldots, N_n$?

- Takáto úloha je vo všeobecnosti tažký kombinatorický problém. Preto sa diskrétne premenné N_i nahradia reálnymi hodnotami w_i spĺňajúcimi

$$w_i \geq 0, \sum_{i=1}^{n} w_i = 1, \; Nw_i \in \mathbb{Z},$$

a posledná požiadavka sa ignoruje.

- Táto idea viedla k pojmu tzv. **asymptotického dizajnu**, ktorý sa v štatistike zvyčajne definuje ako diskrétna pravdepodobnostná miera ξ na X.

Optimálny návrh experimentov

- V prípade konečnej množiny X je asymptotický dizajn určený hodnotami $w_i = \xi(x_i)$. BUNV budeme pod asymptotickým dizajnom rozumieť vektor w zo simplexu

$$\mathbb{S}^n = \left\{ w \in \mathbb{R}^n_+ \;\middle|\; \sum_{i=1}^{n} w_i = 1 \right\}.$$

- Kvalita dizajnu závisí od tzv. informačnej matice

$$M(w) = \sum_{i=1}^{n} w_i f(x_i) f(x_i)^T,$$

ktorá je kladne definitná, pokiaľ vektory $f(x_1), \ldots, f(x_n)$ generujú $\mathbb{R}^m$ a $w > 0$.

Optimálny návrh experimentov

- Úloha v optimálnom navrhovaní experimentov:

$$\begin{array}{ll} \max_w & M(w) \\ & w \in \mathbb{S}^n. \end{array}$$

- V štatistike existuje niekoľko krítérií optimality, t.j. funkcií

$$\Phi : \mathcal{S}^n_{++} \to \mathbb{R}$$

ktoré slúžia ako miera veľkosti matice.

- kritérium E-optimality: $\Phi_E(M) = \lambda_{min}(M)$;
- kritérium A-optimality: $\Phi_A(M) = [\mathrm{tr}(\mathrm{M}^{-1})]^{-1}$;
- kritérium D-optimality: $\Phi_D(M) = \log\det(M)$;

Optimálny návrh experimentov

Pri hľadaní E-optimálneho dizajnu rieši optimalizačná úloha

$$\begin{array}{ll} \max & \lambda_{min}(M(w)) \\ & w \in \mathbb{S}^n. \end{array}$$

Keďže

$$t \leq \lambda_{min}(M(w)) \;\Leftrightarrow\; M(w) \succeq tI,$$

predchádzajúca úloha sa dá naformulovať ako úloha SDP

$$\begin{array}{ll} \max & t \\ & M(w) - tI \succeq 0, \\ & w \in \mathbb{S}^n. \end{array} \tag{3}$$

Optimálny návrh experimentov

Duálna úloha k úlohe E-optimálneho dizajnu

Úlohu (3) si preformulujeme do nasledovného minimalizačného tvaru:

$$\begin{array}{ll} \min & \frac{1}{t} \\ & \Sigma_{i=1}^{n} w_i f_i f_i^T - tI \succeq 0, \\ & w \geq 0, \Sigma_{i=1}^{n} w_i = 1. \end{array}$$

Lagrangeova funkcia k tejto úlohe je $L : (\mathbb{R}^n \times \mathbb{R}_{++}) \times (\mathcal{S}_+^n \times \mathbb{R}_+^n \times \mathbb{R}) \to \mathbb{R}$ definovaná ako

$$L(w,t;Z,y,\alpha) = \frac{1}{t} - \mathsf{tr}\left(Z\left(\sum_{i=1}^{n} w_i f_i f_i^T - tI \succeq 0\right)\right) - y^T w + \alpha\left(\sum_{i=1}^{n} w_i - 1\right)$$

$$= \frac{1}{t} + t\mathsf{tr}(Z) - \sum_{i=1}^{n} w_i \left(f_i^T Z f_i + y_i - \alpha\right) - \alpha.$$

Optimálny návrh experimentov

Lagrangeova duálna funkcia je definovaná ako

$$g(Z, y, \alpha) = \inf_{w,t} L(w, t; Z, y, \alpha).$$

Odvodíme

$$g(Z, y, \alpha) = \begin{cases} 2\sqrt{\mathsf{tr}(Z)} - \alpha & f_i^T Z f_i + y_i - \alpha = 0, \\ -\infty & \text{inak.} \end{cases}$$

Optimálny návrh experimentov

Lagrangeovu duálnu úlohu môžeme preto naformulovať ako

$$\begin{array}{ll} \max & 2\sqrt{\mathsf{tr}(Z)} - \alpha \\ & f_i^T Z f_i + y_i - \alpha = 0, \;\; i = 1, \ldots, n \\ & y \geq 0, Z \succeq 0, \end{array}$$

Ekvivalentná formulácia (po zjednodušení):

$$\begin{array}{ll} \max & \mathsf{tr}(W) \\ & f_i^T W f_i \leq 1, \;\; i = 1, \ldots, n \\ & W \succeq 0, \end{array} \tag{4}$$

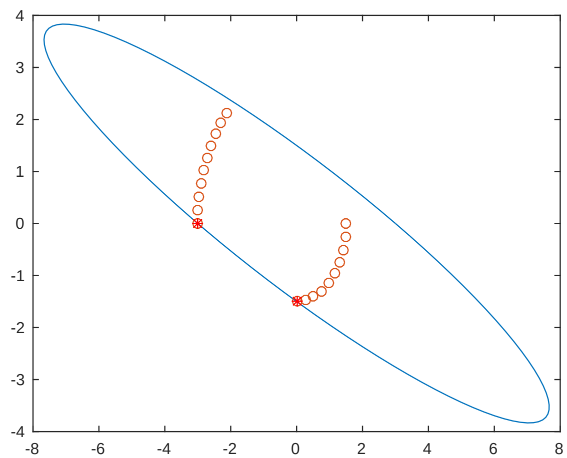
Optimálny návrh experimentov

Dané dú dáta $m = 2, n = 20$ a hodnoty regresných funkcií uložené v $n \times n$ matici F. Dátové body (riadky matice F) sú vykreslené na obrázku.
Zdroj: `http://cvxopt.org/examples/book/expdesign.html`

Pomocou CVX spočítame hodnotu optimálneho dizajnu:

$$w_i^* = \begin{cases} 0.2 & i = 10; \\ 0.8 & i = 20; \\ 0 & \text{inak.} \end{cases}$$

Optimálny návrh experimentov



Elipsa na obrázku zodpovedá optimálnej matici W^* z duálnej úlohy. Všimnime si, že body zodpovedajúce kladným hodnotám w_i^* ležia na hranici elipsy, čo je dôsledok podmienok komplementarity.

Optimálny návrh experimentov

CVX kód:

```
n=size(F,1);
m=size(F,2);
cvx_begin sdp
   variable w(n,1)
   variable t
   dual variable Z
   dual variable alfa
   maximize t
   subject to
      Z:F'*diag(w)*F-t*eye(m)>=0;
      w>=0;
      alfa:sum(w)==1;
cvx_end
```

Úloha minimálnej plochy

Uvažujme diferencovateľnú funkciu $f : C \subset \mathbb{R}^2 \to \mathbb{R}$. Plocha daná grafom funkcie f je vyjadrená ako

$$A(f) = \int_C \sqrt{1 + \|\nabla f(x)\|_2^2} dx = \int_C \left\| \begin{pmatrix} \nabla f(x) \\ 1 \end{pmatrix} \right\|_2 dx.$$

Úloha minimalizácie plochy je o nájdení funkcie f, pre ktorú je za daných ohraničení plocha A minimálna.

- $C = [0,1] \times [0,1]$ - diskretizuje sa mriežkou s bodmi $(ih, jh), 0 \leq i, j \leq K$, kde K je celé číslo a $h = 1/K$ je (rovnomerné) delenie

- **Premenné:** hodnoty f_{ij} funkcie f v bodoch (ih, jh), čo možno reprezentovať maticou $F \in \mathbb{R}^{(K+1)\times(K+1)}$, $F_{ij} = f(ih, jh)$

Úloha minimálnej plochy

Aproximácia parciálnych derivácií funkcie f pomocou pomerných differencií

$$\frac{\partial f(x)}{\partial x_1} = \frac{f(x_1 + h, x_2) - f(x_1, x_2)}{h}, \quad \frac{\partial f(x)}{\partial x_2} = \frac{f(x_1, x_2 + h) - f(x_1, x_2)}{h}.$$

Aproximácia gradientu funkcie f v bode $x = (ih, jh)$:

$$\nabla f(x) \approx K \begin{bmatrix} f_{i+1,j} - f_{i,j} \\ f_{i,j+1} - f_{i,j} \end{bmatrix}$$

Aproximácia plochy danej grafom funkcie f:

$$A(f) \approx A_{disk}(f) = \frac{1}{K^2} \sum_{i,j=0}^{K-1} \left\| \begin{pmatrix} K(f_{i+1,j} - f_{i,j}) \\ K(f_{i,j+1} - f_{i,j}) \\ 1 \end{pmatrix} \right\|_2 .$$

Úloha minimálnej plochy

Možno uvažovať rôzne ohraničenia na f_{ij}, napr. ohraničenia na hraničné hodnoty C.

$$\begin{array}{ll} \min & A_{disk} \\ & f_{0,j} = l_j, \ j = 0, 1, \ldots, K, \\ & f_{K,j} = r_j, \ j = 0, 1, \ldots, K, \end{array} \tag{5}$$

kde l, r sú vektory hraničných hodnôt (ktoré možno získať diskretizáciou daných jednorozmerných funkcií).

Úloha minimálnej plochy

SOCP reformulácia:

$$\min \quad \Sigma_{i,j=0}^{K-1} t_{ij}$$

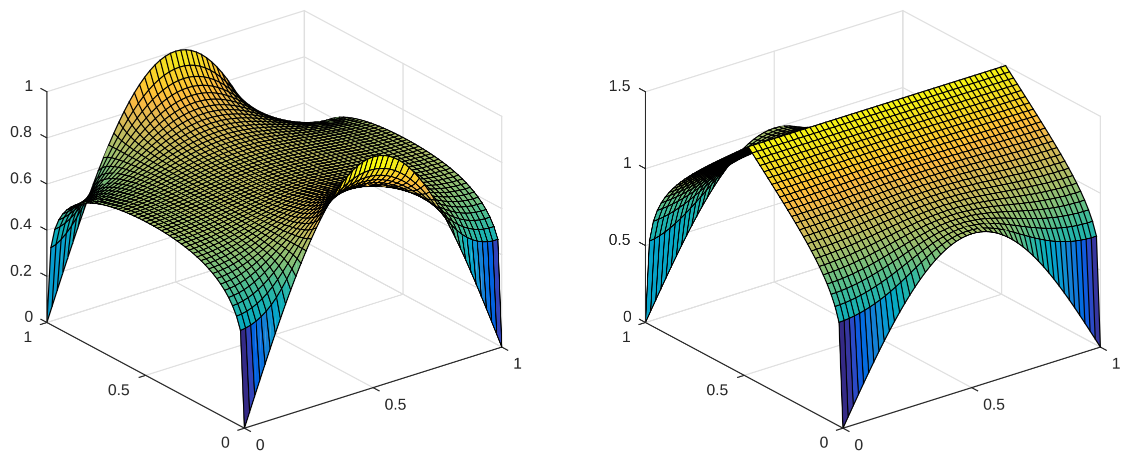
$$\left\| \begin{pmatrix} K(f_{i+1,j} - f_{i,j}) \\ K(f_{i,j+1} - f_{i,j}) \\ 1 \end{pmatrix} \right\|_2 \leq t_{ij}, \; i,j = 0,1,\dots,K-1 \qquad (6)$$

$$f_{0,j} = l_j, \; j = 0,1,\dots,K,$$

$$f_{K,j} = r_j, \; j = 0,1,\dots,K,$$

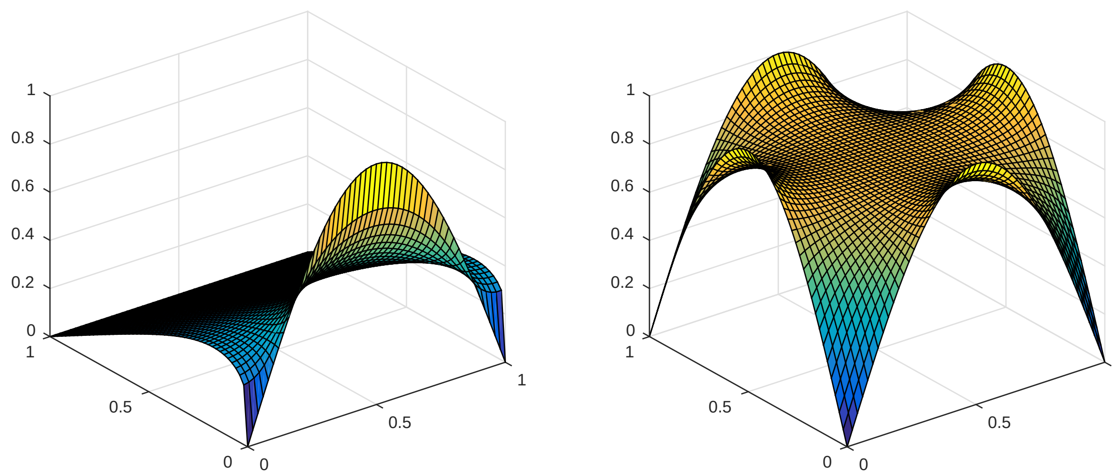
Úloha minimálnej plochy

Grafický výstup (CVX/Matlab):



Úloha minimálnej plochy

Grafický výstup (CVX/Matlab):



Klasifikácia dát

Dané sú dne množiny bodov

$$X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^n \text{ a } Y = \{y_1, \ldots, y_M\} \subset \mathbb{R}^n.$$

Úlohou je nájsť funkciu $f : \mathbb{R}^n \to \mathbb{R}$ (z danej riedy funkcií) takú, že

$$f(x_i) < 0,\ i = 1, \ldots, N,\ f(y_i) > 0,\ i = 1, \ldots, M.$$

Pokiaľ sú tieto nerovnosti splnené, tak hovoríme, že množina

$$\{x \mid f(x) = 0\}$$

separuje množiny bodov X, Y.

Klasifikácia dát

Lineárna diskriminácia

Ak zvolíme pre klasifikáciu triedu afínnych funkcií, t. j.

$$f(x) = a^T x + b,$$

hovoríme a lineárnej diskriminácii.

Úloha: nájsť nadrovinu $\{x \mid a^T x - b = 0\}$ ktorá separuje množiny bodov X a Y, t.j. platí

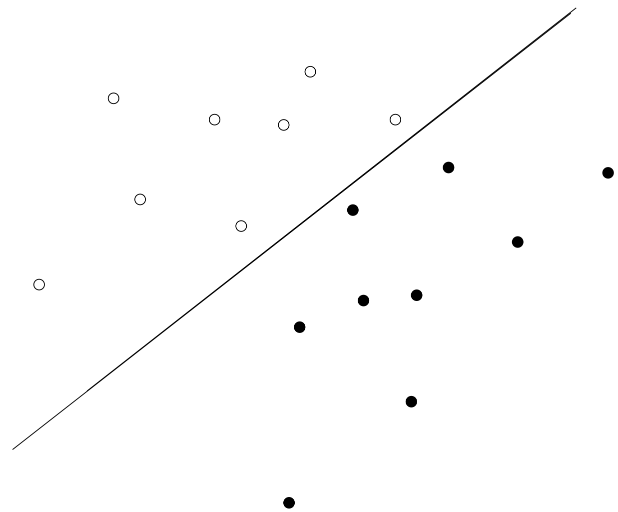
$$a^T x_i - b > 0,\ i = 1, \dots, N,\ a^T y_i - b < 0,\ i = 1, \dots, M.$$

Homognita nerovností v premenných (a, b) - ekvivalentné nerovnosti

$$a^T x_i - b \geq 1,\ i = 1, \dots, N,\ a^T y_i - b \leq -1,\ i = 1, \dots, M.$$

Takáto úloha sa dá riešiť ako **lineárna úloha prípustnosti**.

Klasifikácia dát



Ak sú dátové množiny v $\mathbb{R}^2$ separovateľné, lineárnou úlohou prípustnosti možno nájsť separujúcu priamku.

Klasifikácia dát

Prípad neseparovateľných dátových množín (SVM)

Zavedú sa nové **nezáporné** premenné $u_1, \ldots, u_N$ a $v_1, \ldots, v_M$.

Lineárne ohraničenia sa relaxujú:

$$a^T x_i - b \geq 1 - u_i,\ i = 1, \ldots, N,\ a^T y_i - b \leq -1 + v_i,\ i = 1, \ldots, M.$$

LP formulácia:

$$\begin{array}{ll} \min_{a,b,u,v} & \mathbf{e}^T u + \mathbf{e}^T v \\ & a^T x_i - b \geq -1 - u_i,\ i = 1, \ldots, N, \\ & a^T y_i - b \leq 1 + v_i,\ i = 1, \ldots, M, \\ & u, v \geq 0. \end{array}$$

Klasifikácia dát

Kvadratická diskriminácia

Ak zvolíme pre klasifikáciu triedu konvexných kvadratických funkcií

$$f(x) = x^T Px + q^T x + r,$$

hovoríme a kvadratickej (elipsoidnej) diskriminácii. Úlohou je nájsť $P \in \mathcal{S}^n_{++}, q \in \mathbb{R}^n, r \in \mathbb{R}$ platí

$$x_i^T Px_i + q^T x_i + r < 0,\ i = 1, \ldots, N,\ y_i^T Py_i + q^T y_i + r > 0,\ i = 1, \ldots, M.$$

Homogenita premenných (P, q, r) - ekvivalentné nerovnosti

$$x_i^T Px_i + q^T x_i + r \le -1,\ i = 1, \ldots, N,\ y_i^T Py_i + q^T y_i + r \ge 1,\ i = 1, \ldots, M.$$

Klasifikácia dát

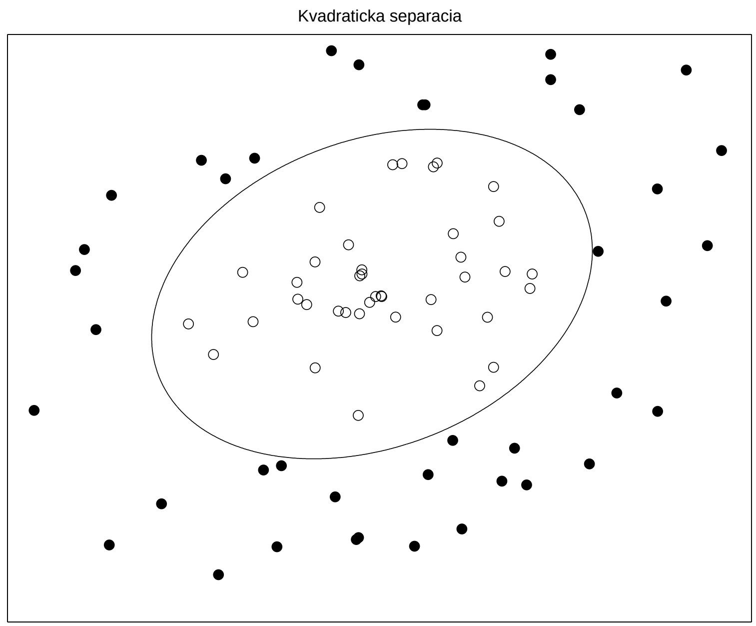
Kvadratická (elipsoidná) diskriminácia

Semidefinitná úloha prípustnosti: nájsť P, q, r také, že

$$
\begin{array}{l}
x_i^T P x_i + q^T x_i + r \leq -1, \ i = 1, \ldots, N, \\
y_i^T P y_i + q^T y_i + r \geq 1, \ i = 1, \ldots, M, \\
P \succeq I,
\end{array}
$$

kde podmienku $P \in \mathcal{S}^n_{++}$ môžeme kvôli homogenite premenných nahradiť LMI ohraničením $P \succeq I$.

Klasifikácia dát



Ak sú dátové množiny v $\mathbb{R}^2$ separovateľné, semidefinitnou úlohou prípustnosti možno nájsť separujúcu elispu.

Klasifikácia dát

Prípad neseparovateľných dátových množín (SVM)

Zavedú sa nové **nezáporné** premenné $u_1, \ldots, u_N$ a $v_1, \ldots, v_M$.

Kvadratické ohraničenia sa relaxujú:

$$
\begin{array}{l}
x_i^T P x_i + q^T x_i + r \leq -1 + u_i,\ i = 1, \ldots, N, \\
y_i^T P y_i + q^T y_i + r \geq 1 - v_i,\ i = 1, \ldots, M,
\end{array}
$$

SDP formulácia:

$$
\begin{array}{ll}
\min_{a,b,u,v} & \mathbf{e}^T u + \mathbf{e}^T v \\
& x_i^T P x_i + q^T x_i + r \leq -1 + u_i,\ i = 1, \ldots, N, \\
& y_i^T P y_i + q^T y_i + r \geq 1 - v_i,\ i = 1, \ldots, M, \\
& u, v \geq 0, P \succeq I.
\end{array}
$$