Domáca úloha č. 1

2-INF-150: Strojové učenie, Zima 2017

Termín: 6.11.2017 9:00, M163 (pod dvere)
domácu úlohu odovzdávajte písomne (nie elektronicky)

1. Maticový počet.

a) Povedzte za akých podmienok platí a dokážte:

$$\nabla_A \mathrm{tr}(A^T BAC) = B^T AC^T + BAC$$

b) Uvažujme problém lineárnej regresie, pričom optimalizujeme normu L_2. Nech X je dizajnová matica, y je cieľový vektor trénovacej množiny a θ je vektor parametrov lineárnej funkcie, ktorú v procese trénovania hľadáme. Na prednáške sme odvodili, že pre vektor θ musí platiť:

$$X^T X\theta = X^T y.$$

Ak matica $(X^T X)$ je regulárna, môžeme túto rovnicu prenásobiť zľava inverznou maticou $(X^T X)^{-1}$ a ďalším odvodením dostaneme:

$$\theta = (X^T X)^{-1} X^T y = X^{-1} X^{T^{-1}} X^T y = X^{-1} y$$

$$\theta = X^{-1} y$$

Kde je v tomto odvodení problém?

2. Teória strojového učenia. Uvažujme problém regresie nad množinou hypotéz $H = \{h_b : x \to 2x + b\}$.

a) Popíšte algoritmus, ktorý spočíta pre danú trénovaciu množinu $(x_1, y_1), \ldots, (x_t, y_t)$ hypotézu, ktorá minimalizuje chybu charakterizovanú chybovou funkciou $J(b) = \frac{1}{t} \sum_{i=1}^{t} (h_b(x_i) - y_i)^2$.

b) Uvažujme pravdepodobnostnú distribúciu $P_{x,y}$ definovanú nasledujúcim spôsobom:

- rozdelenie x-ov je rovnomerné na intervale $[0, 100]$
- pre dané x je $\Pr(y = 2x+3 \mid x) = 0.3$ a $\Pr(y = 2x-2 \mid x) = 0.7$ (iné hodnoty y sa v kombinácii s daným x nevyskytujú).

Aká je výchylka pre množinu hypotéz H ak predpokladáme, že dáta sú nezávislými vzorkami z tejto distribúcie?

c) Pre pravdepodobnostnú distribúciu dát $P_{x,y}$ z časti b) a $t = 1$ spočítajte očakávanú trénovaciu a očakávanú testovaciu chybu.

d) Pre pravdepodobnostnú distribúciu dát $P_{x,y}$ z časti b) a všeobecné t spočítajte očakávanú trénovaciu a testovaciu chybu. V akom vzťahu sú tieto chyby ku výchylke keď $t \to \infty$? Vedeli by ste na základe vašich výsledkov zvoliť vhodnú veľkosť trénovacej množiny?