

1 Úvod

V našej práci sa budeme venovať implementácii lineárnej regresie ako úlohy lineárneho programovania. Lineárna regresia je spôsob odhadovania závislej premennej $y \in \mathbb{R}^n$ ako lineárnej kombinácie nezávislých premenných $x_1, \dots, x_k \in \mathbb{R}^n$ s pridaným skalárnym členom: $\hat{y} = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$. Takýto problém môžeme interpretovať ako n pozorovaní, kde pre každé pozorovanie sledujeme k atribútov, čiže vektor x_i pre $i = 1, \dots, k$ predstavuje dáta atribútu i pre všetkých n pozorovaní. Pomocou lineárnej funkcie týchto premenných sa budeme snažiť čo najlepšie predikovať atribút y .

Na meranie vzdialenosti medzi vektorom y a odhadovaným vektorom \hat{y} budeme používať L^1 a L^∞ normy, keďže práve pre tie sa dá tento problém naformulovať ako úloha lineárneho programovania. V kapitole ?? sa venujeme matematickej formulácii LP úlohy a dokazovaniu jej optimality. V kapitole ?? vizualizujeme funkčnosť modelu na arbitrárnych 2D dátach `A04plotregres.npz`. Následne, v kapitole ?? sa venujeme predikovaniu ceny vína podľa dátového súboru `A04wine.csv`. Pre tieto predikcie následne spočítame koeficient determinácie v ?? . Na záver, sekcia ?? popisuje našu implementáciu L^1 a L^∞ lineárnej regresie pre ľubovoľné dáta v programovacom jazyku Python. Tiež sa tam venujeme porovnávaniu správania takýchto regresii a formulácii a implementácii minimalizovania váženej sumy týchto noriem.