

ÚLOHA 2

Lineární regrese jedné proměnné

Zadáno na cvičení: 3
Mezní termín: 24.10. 2017
Maximální počet bodů: 10
Nepovinná úloha

Zadání

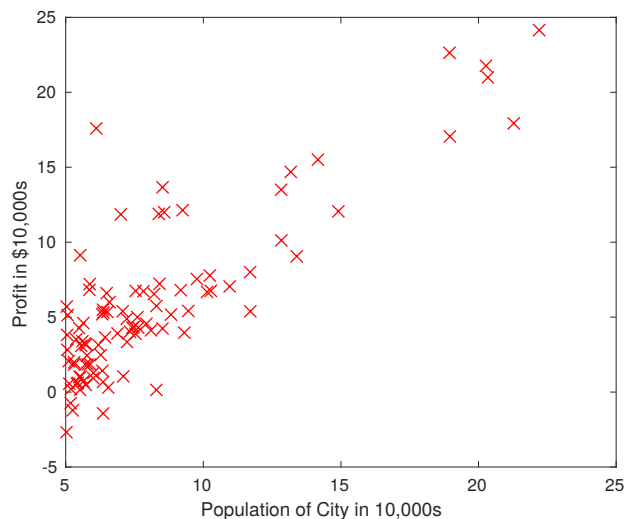
Stáhněte si archiv *linRegSimple.zip* ze stránky 2 *Lineární regrese*. Archiv obsahuje tyto soubory:

- *computeCost.m* ★ – výpočet cenové funkce
- *data.txt* – vstupní data
- *linRegSimple.m* ★ – hlavní spouštěcí skript celé úlohy
- *plotData.m* – vizualizace dat
- *predict.m* ★ – hypotéza lineární regrese
- *train.m* ★ – Vypočítá parametry regresní přímky ze vstupních dat (pomocí gradientního sestupu).

Soubory označené ★ budete doplňovat¹.

Vstupní data

Firma chce predikovat předpokládaný zisk v případě založení pobočky v novém městě podle počtu obyvatel.



Obrázek 1: Vizualizace dat.

¹Což ale neznamená, že nemůžete v rámci lepší dekompozice vytvářet další soubory.

Úkoly

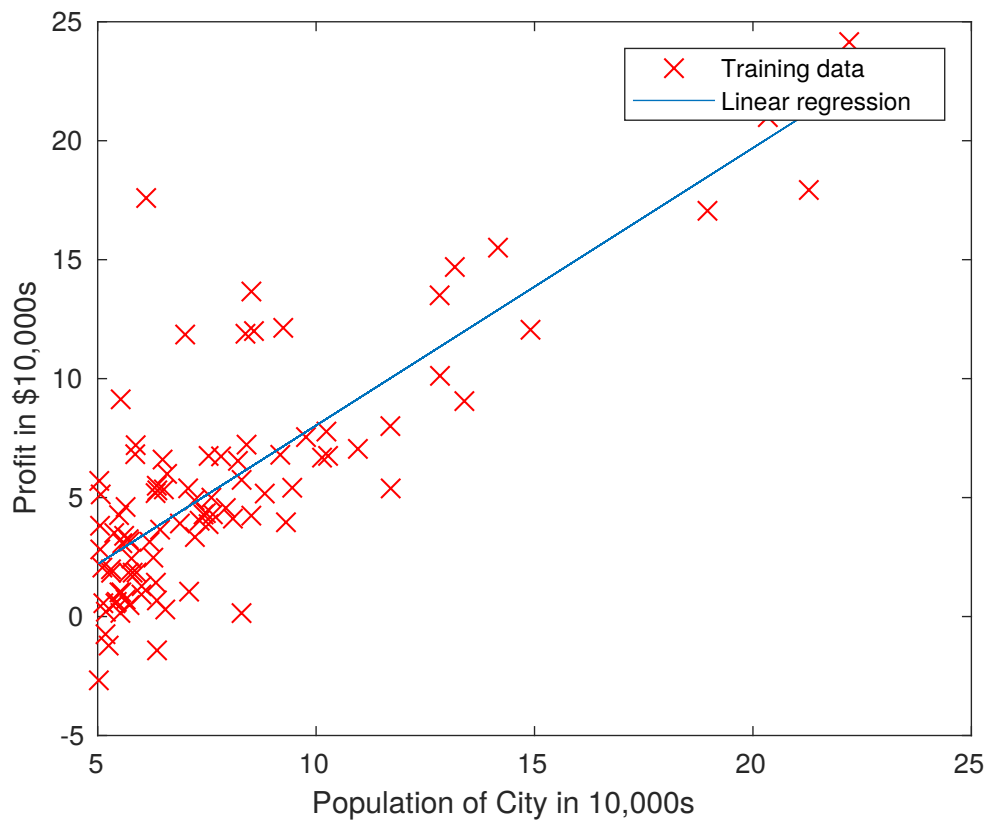
1. Naprogramujte lineární regresi gradientním sestupem.

V této části je potřeba doplnit následující soubory:

- *computeCost.m* – výpočet cenové funkce
- *predict.m* – hypotéza lineární regrese
- *train.m* – Vypočítá parametry regresní přímky

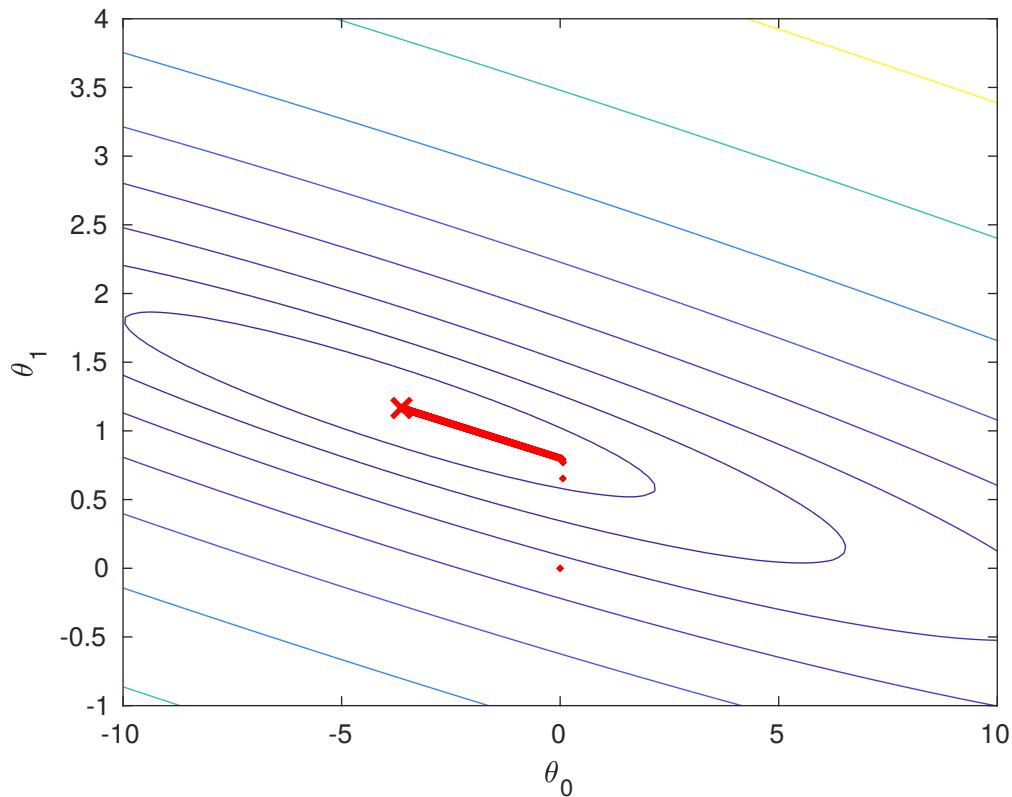
Gradientní sestup bude mít dva parametry. Míru učení α a počet iterací.

Očekávaný výsledek je na obrázku 2.



Obrázek 2: Regrese dat.

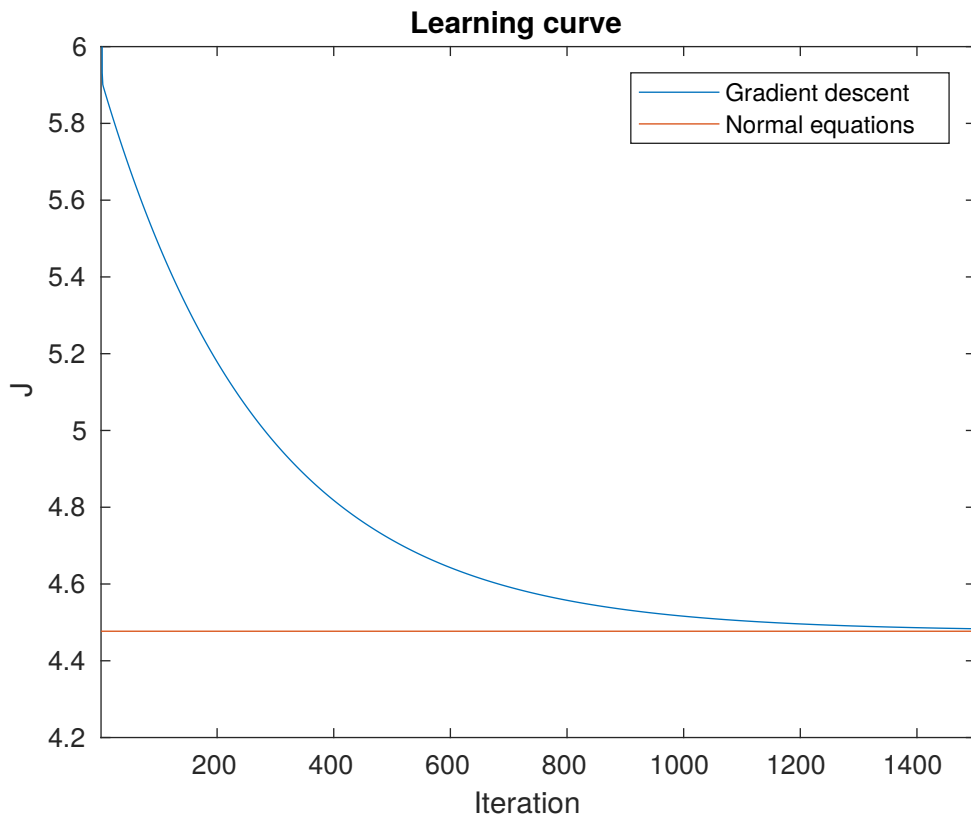
2. V souboru *linRegSimple.m* do obrázku 3 (vizualizace chyby v závislosti na parametrech θ) doplňte hodnoty v jednotlivých iteracích gradientního sestupu (`theta_history`).
3. Vyladte parametry α a počet iterací gradientního sestupu tak, abyste dosáhli co nejmenší chyby. Na grafu sledujte, co změna parametrů dělá s průběhem chyby. Očekávaný výsledek úkolů 2 a 3 je zobrazen na obrázku 3.



Obrázek 3: Vývoj cenové funkce v závislosti na parametrech θ .

4. **Vypočítejte parametry θ pomocí normálních rovnic a srovnejte výslednou chybu s gradientním sestupem.** Výsledky vykreslete do grafu. Vykreslete graf pro různé hodnoty α .

Očekávaný výsledek můžete vidět na obrázku 4.



Obrázek 4: Učící křivka.