# Multivariate Linear Regression

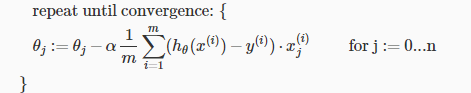
Vectorization

*hθ*(*x*)=*θ*0+*θ*1*x*1+*θ*2*x*2+*θ*3*x*3+⋯+*θnxn*

TO



Gradient Descent for Multiple Variables



Gradient Descent in Practice I - Feature Scaling



Where *μi* is the **average** of all the values for feature (i) and *si* is the range of values (max - min), or *si* is the standard deviation.

Gradient Descent in Practice II - Learning Rate

To summarize:

If α is too small: slow convergence.

If α is too large: ￼may not decrease on every iteration and thus may not converge.

Features and Polynomial Regression

*hθ*(*x*)=*θ*0+*θ*1*x*1+*θ*2*x*21 quadratic function

*hθ*(*x*)=*θ*0+*θ*1*x*1+*θ*2*x*21+*θ*3*x*31  cubic function

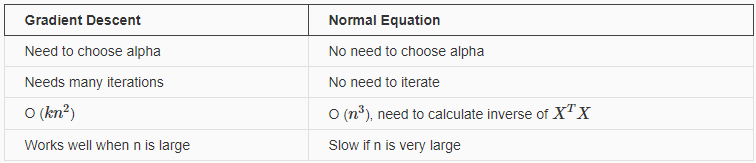
 new features *x*2 and *x*3 where *x*2=*x*21 and *x*3=*x*31. *hθ*(*x*)=*θ*0+*θ*1*x*1+*θ*2√*x*1  root function

 if you choose your features this way then feature scaling becomes very important.

# Computing Parameters Analytically

Normal Equation





Normal Equation Noninvertibility

If X’X is noninvertible, the common causes might be having :

Redundant features, where two features are very closely related (i.e. they are linearly dependent)

Too many features (e.g. m ≤ n). In this case, delete some features or use "regularization" (to be explained in a later lesson).

# Assignment EX1

Tips: Submit Function

For [406648511@qq.com](mailto:406648511@qq.com) CODE IN EX1: HybsTbcr4Ny4vE5Y

背景，预测食物卡车的利润。早有人口与利润的关系。选择下一个开店的 城市。

城市人口（万）与利润（万美元）的关系