Coursera

Week 2

Koray Poyraz

Table of Contents

[Multivariate Linear Regression 2](#_Toc528242375)

[Multiple features 2](#_Toc528242376)

[Gradient Descent for Multiple Variables 4](#_Toc528242377)

[Gradient Descent in Practice I - Feature Scaling 6](#_Toc528242378)

[Feature scaling 6](#_Toc528242379)

[Mean normalization 7](#_Toc528242380)

[Gradient Descent in Practice II - Learning Rate 8](#_Toc528242381)

[Features and Polynomial Regression 11](#_Toc528242382)

[Polynomial regressie 11](#_Toc528242383)

[Computing Parameters Analytically 13](#_Toc528242384)

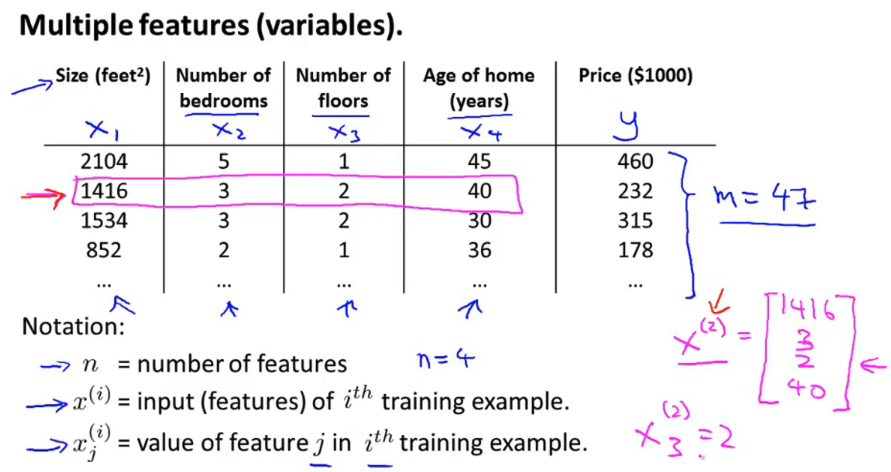
[Normal equation 13](#_Toc528242385)

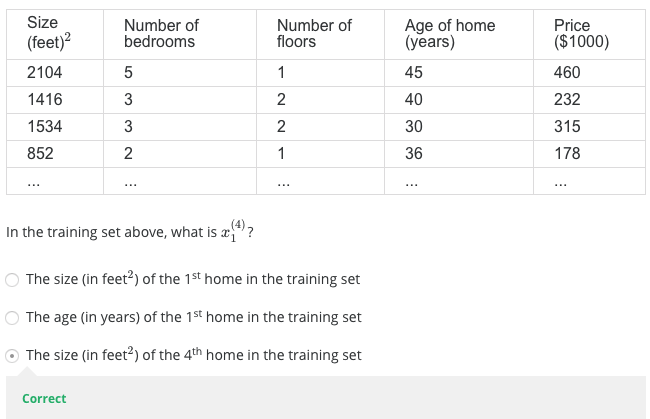
[Normal Equation Noninvertibility (Optioneel) 15](#_Toc528242386)

# Multivariate Linear Regression

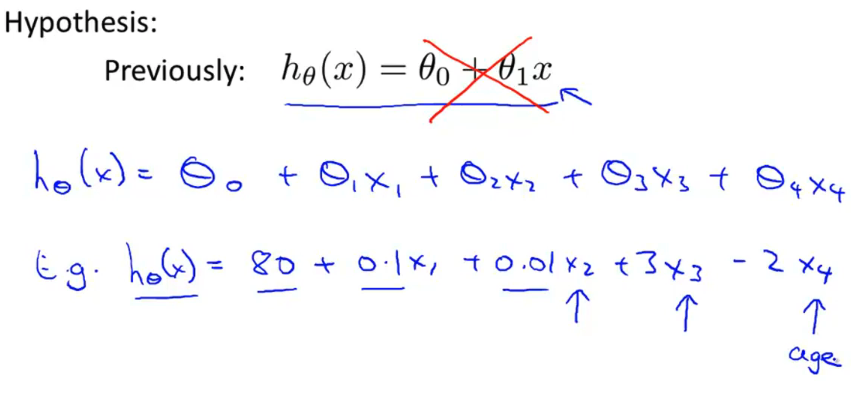
## Multiple features

Mutliple features is gewoon meerdere eigenschappen bijvoorbeeld eigenschappen van een klasse. Dus grootte en prijs van een huis zijn features.

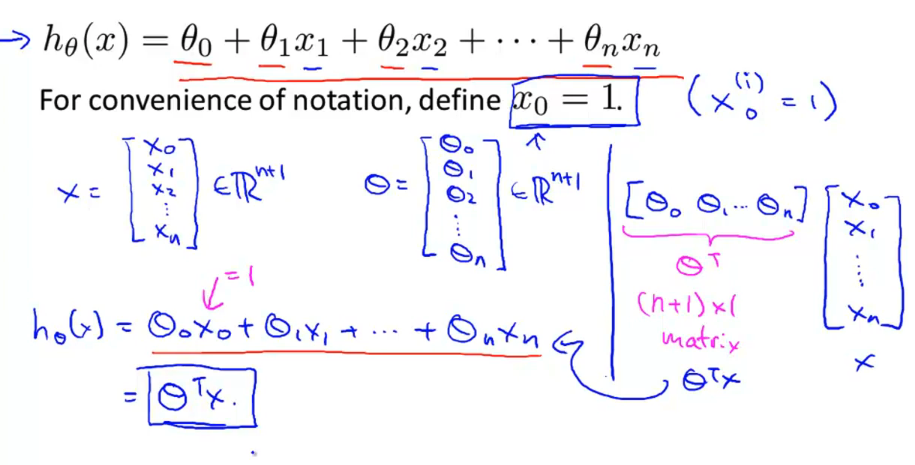


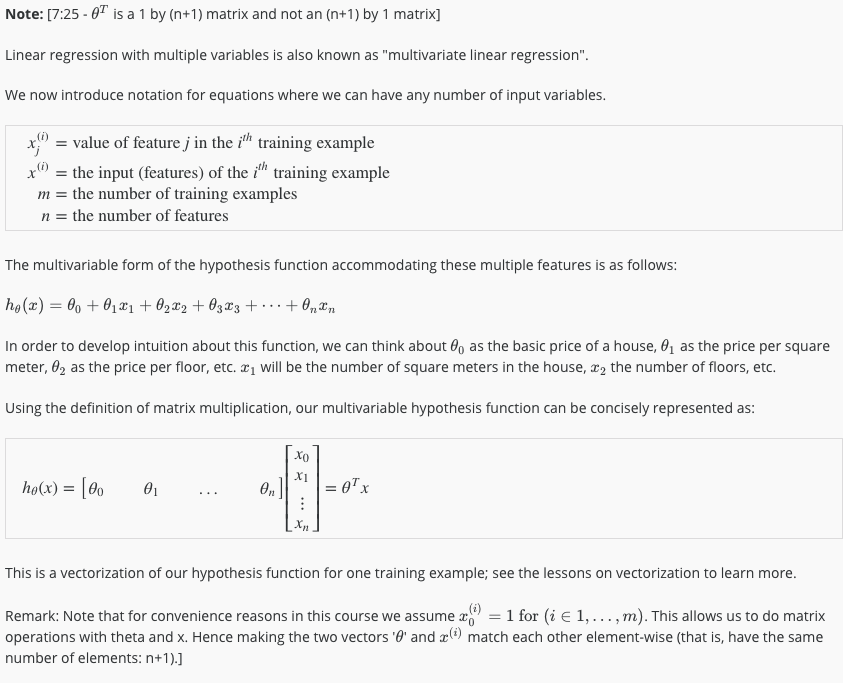


De hypothese bij multiple features ziet er dan als volgt uit:



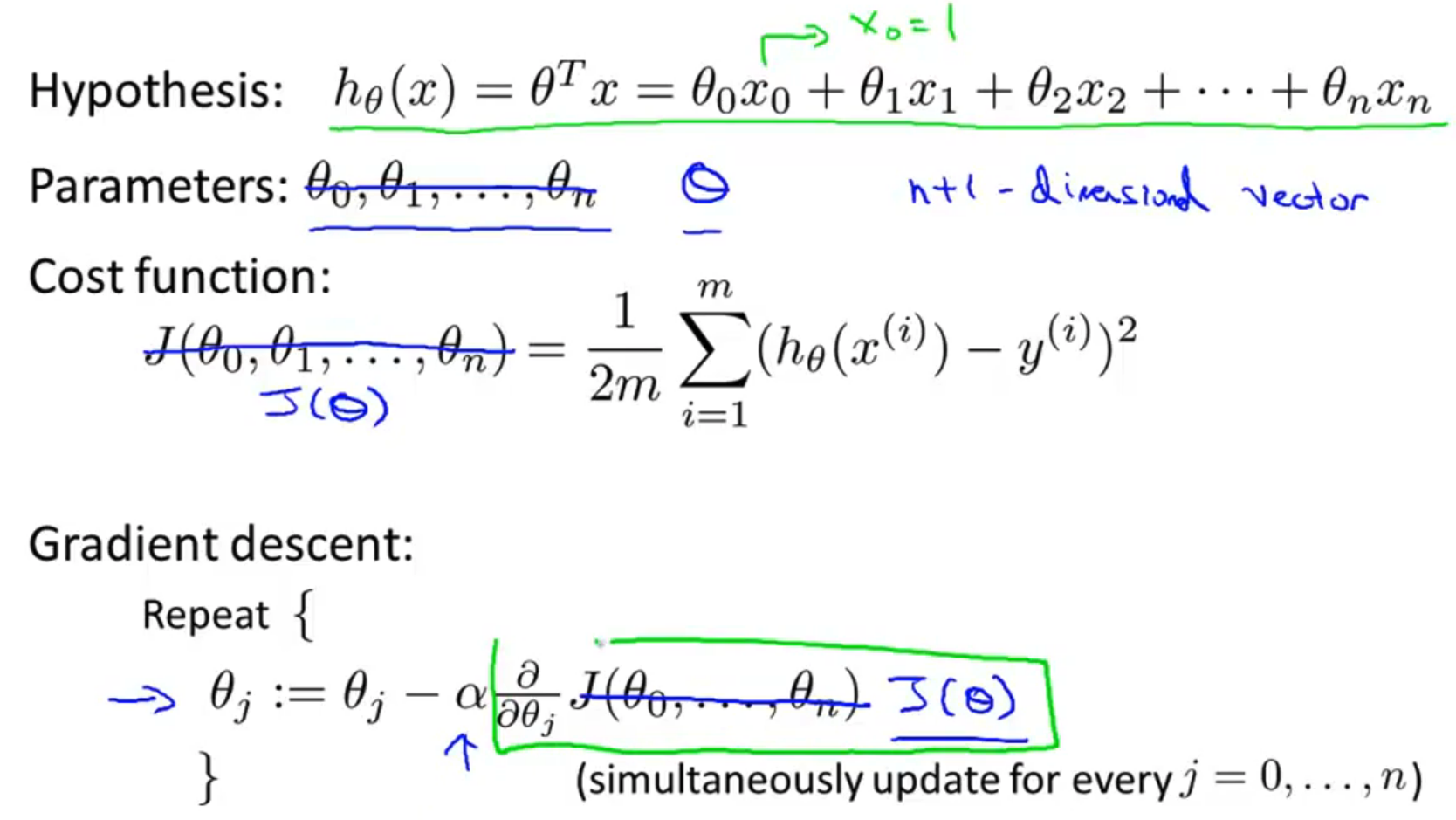
**Multiple features** wordt ook wel **multivariate linear regressie** genoemd.



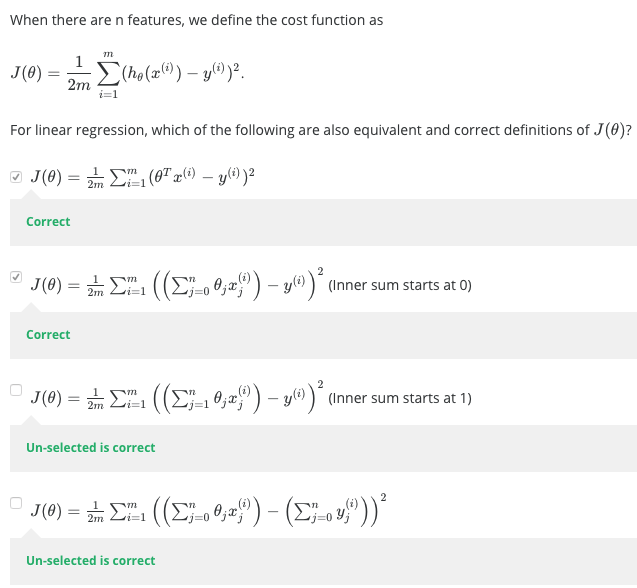


## Gradient Descent for Multiple Variables

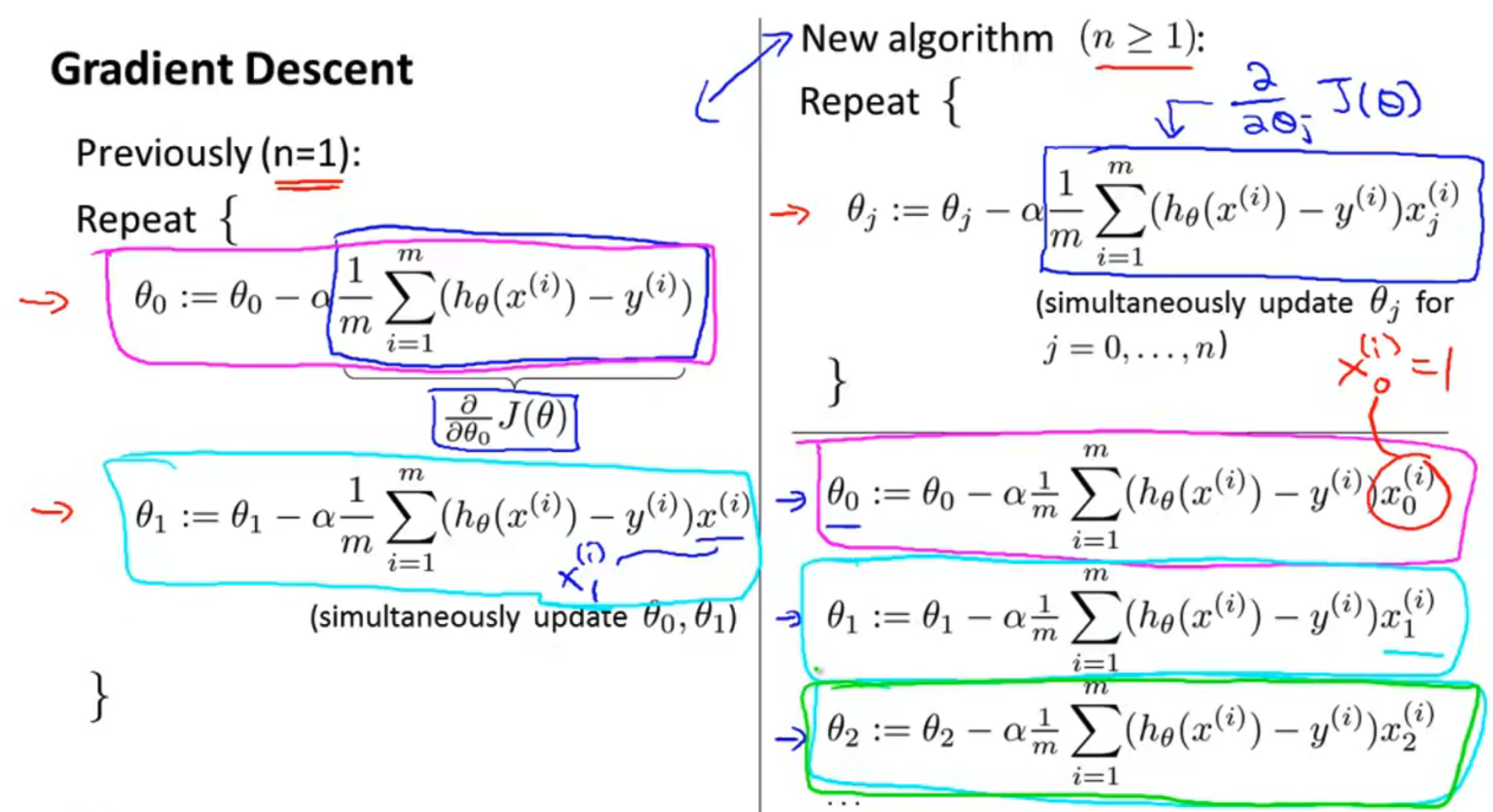
Ziet er als volgt uit:

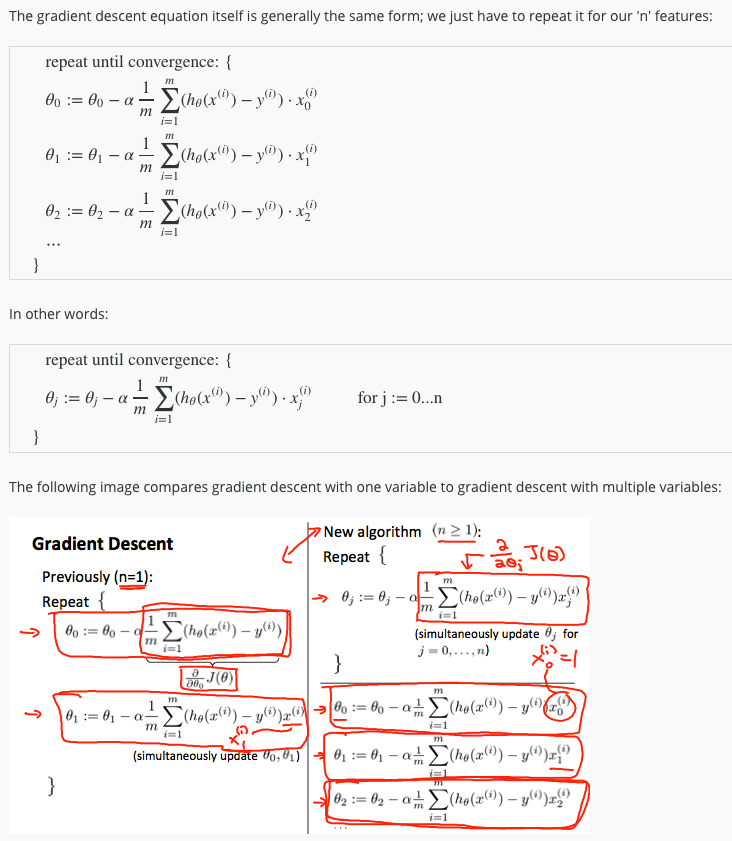


Zoals je kunt zien wordt voor simplificatie J(Ø) gebruikt welke dus een Vector is van thetas.



Gradient descent voorheen (univariate) en nu (multivariate):



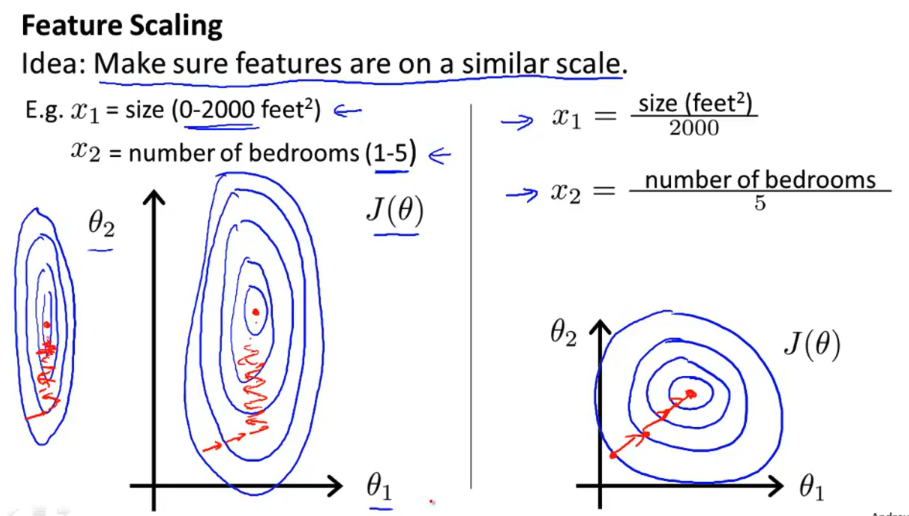


## Gradient Descent in Practice I - Feature Scaling

### Feature scaling

Als de features hetzelfde zijn geschaald dan converge de gradient descent veel sneller.

Linker is langzamer, rechter is sneller (hier zie je ook voorbeeld formule van scaling):

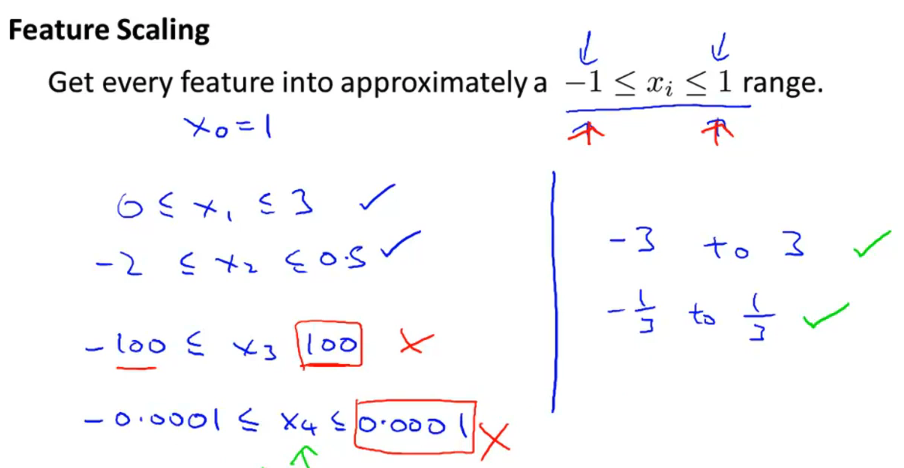


**Formule voor scaling:**

X1 = feature (size) gedeeld door max size

X2 = aantal slaapkamers gedeeld door max slaapkamers

Het is handig om de feature scaling waardes van de features te houden tussen -1 <= Xi <= 1 range.



### Mean normalization

Bij mean normalization pas je de volgende formule toe:

Feature – gemiddelde waarde van de feature in de trainingset

------------------ gedeeld door ---------------

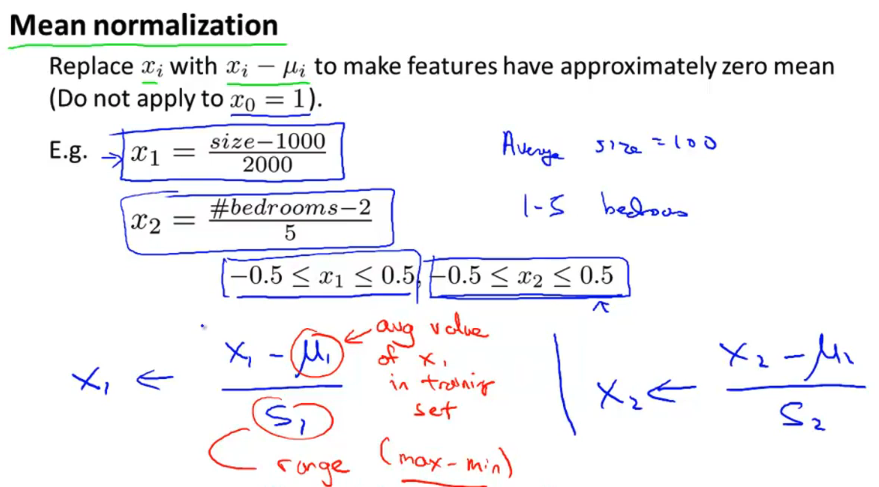
Range (max waarde – min waarde)

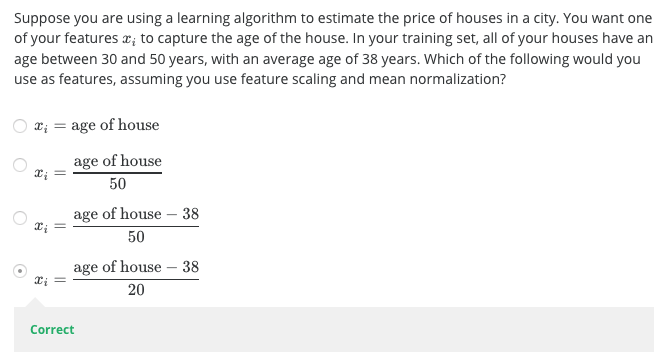
Dus,,

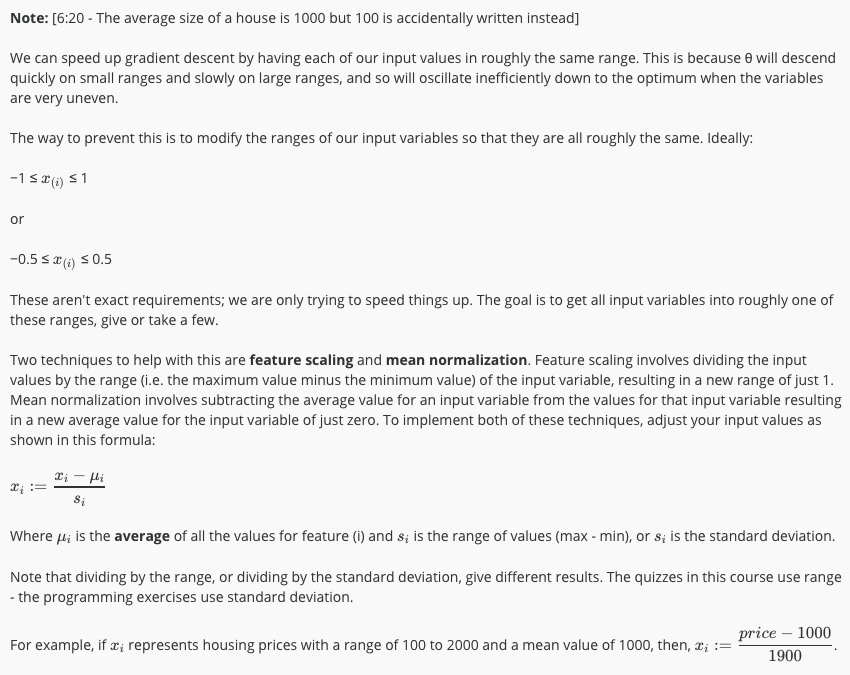
X1 – U1

------------------ gedeeld door ---------------

S1

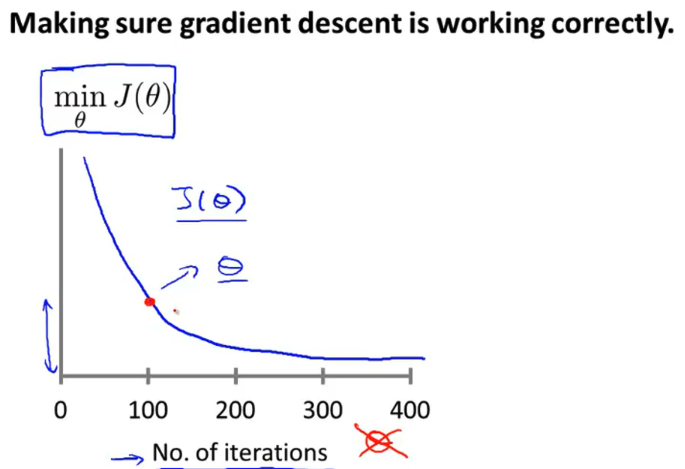


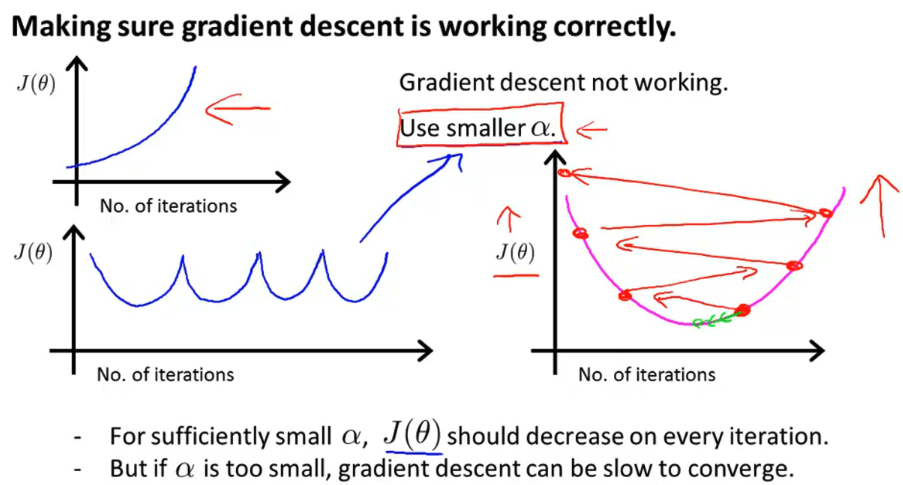


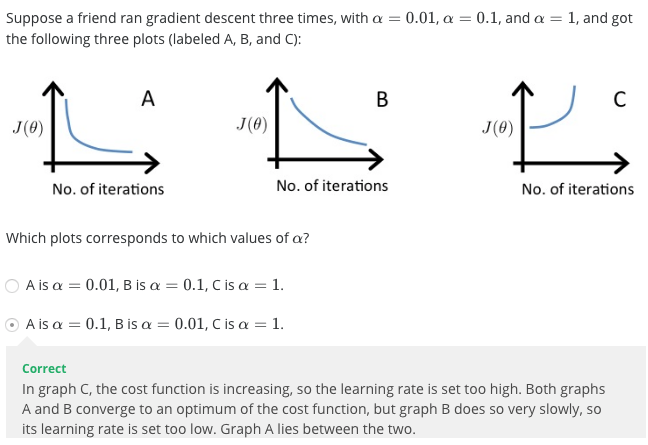


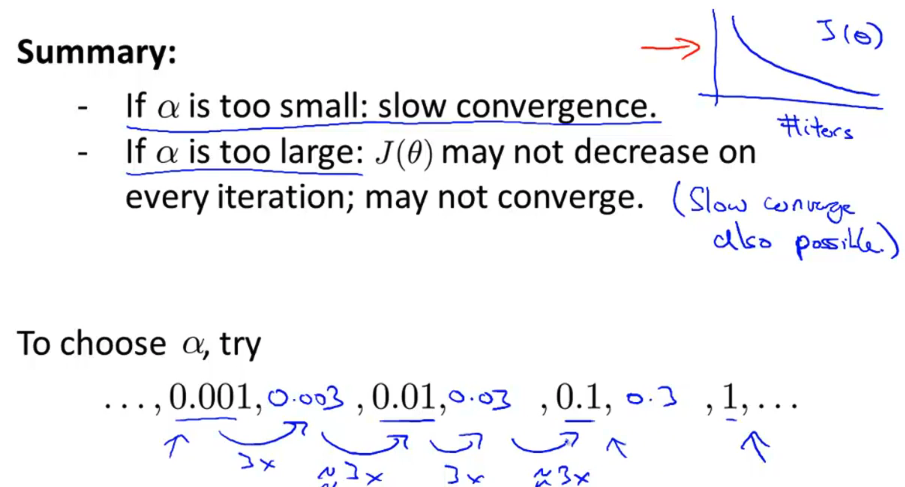
## Gradient Descent in Practice II - Learning Rate

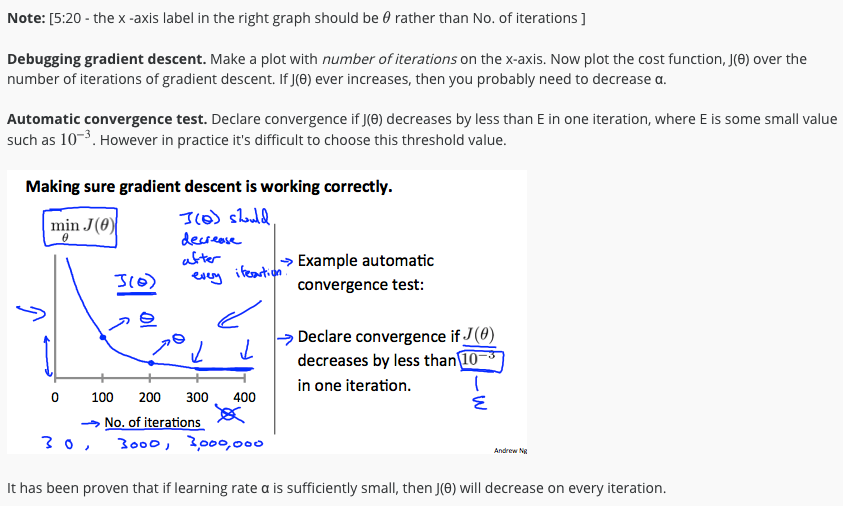
Een manier om de werking van gradient descent te kunnen debuggen kun je gebruik maken van plotten. Door de cost funtie te plotten kun je in de iteraties zien of J(Ø) daalt na iedere iteratie.

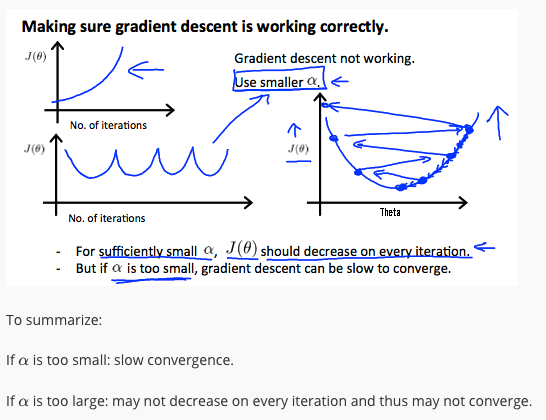






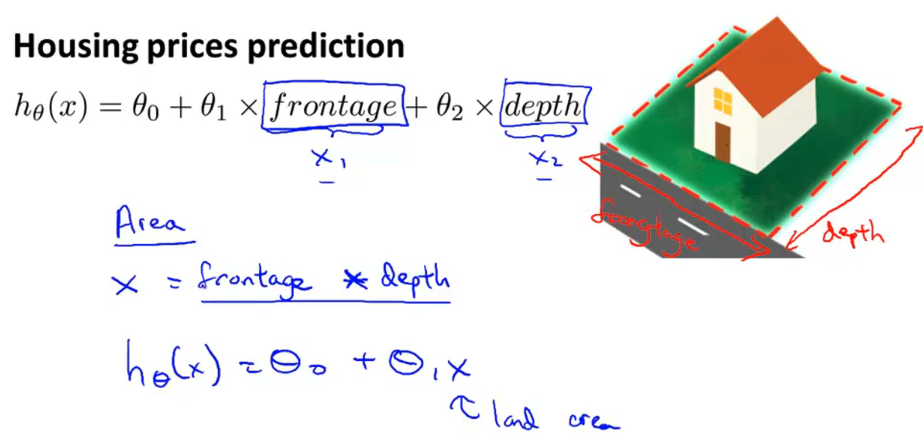




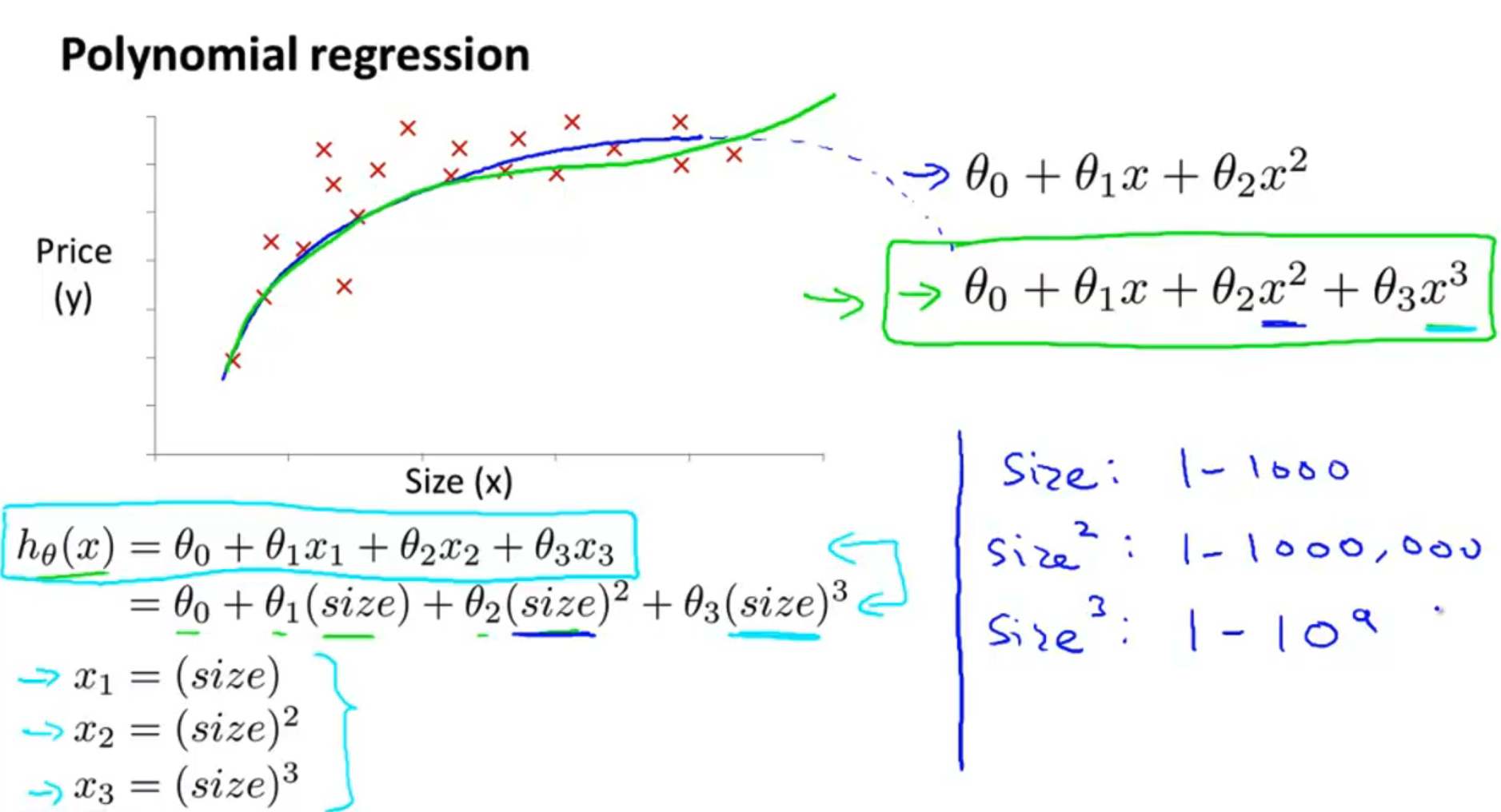


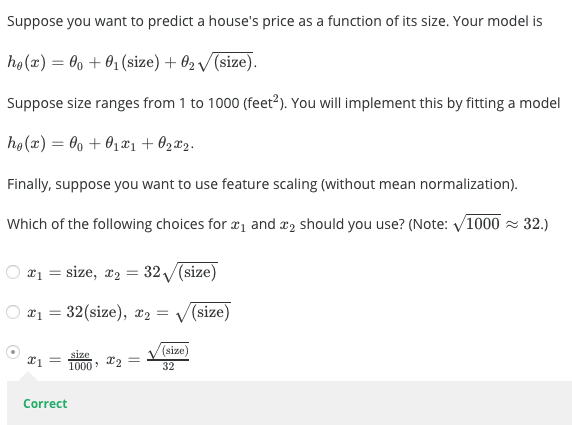
## Features and Polynomial Regression

Vaak kun je features met elkaar vermenigvuldigen tot 1 feature. Bijvoorbeeld bij features zoals breedte of diepte van een huis die features vermenigvuldigen tot 1 feature.

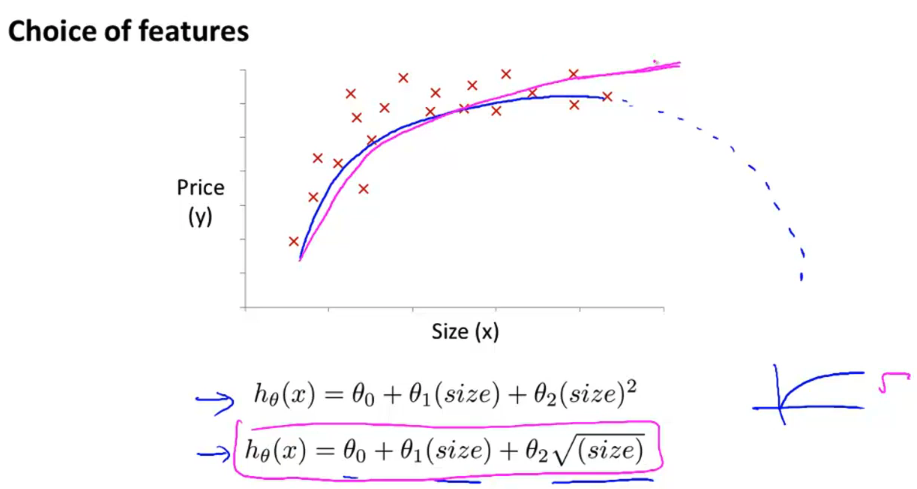


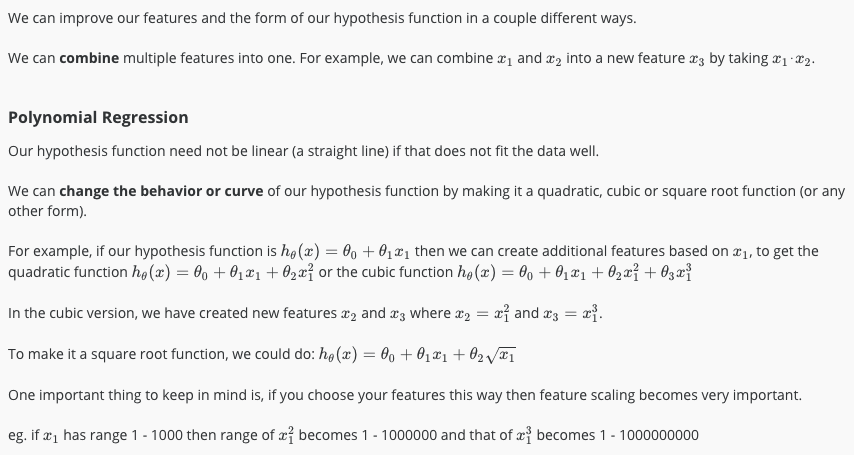
### Polynomial regressie





Je kunt ook gebruik maken van de WORTEL om een mooie polynomial lijn te kunnen creëren.

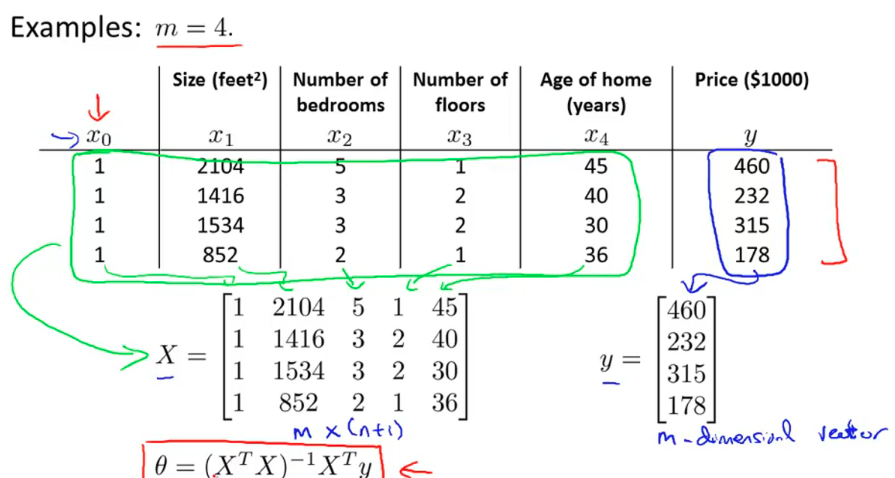


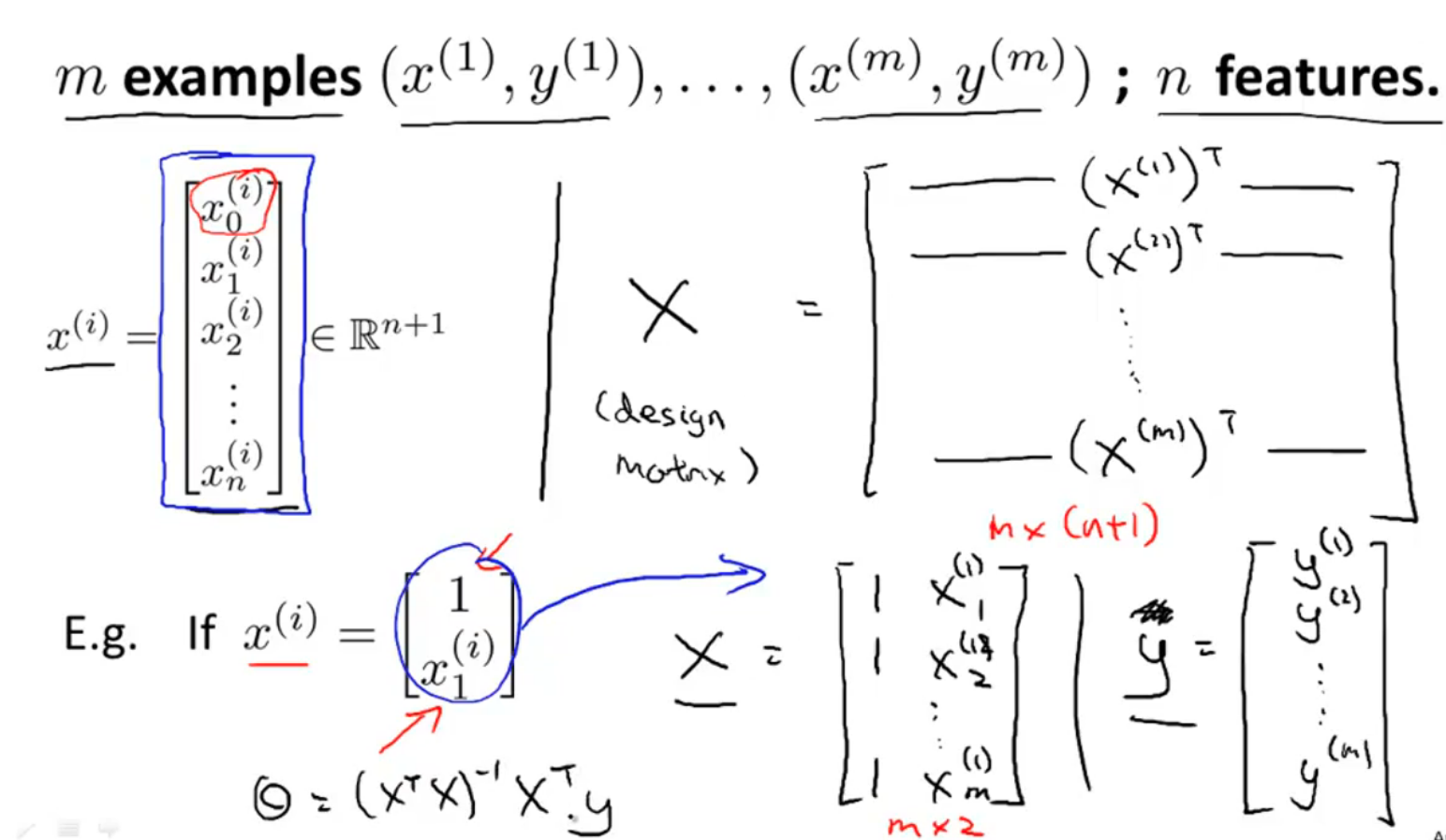


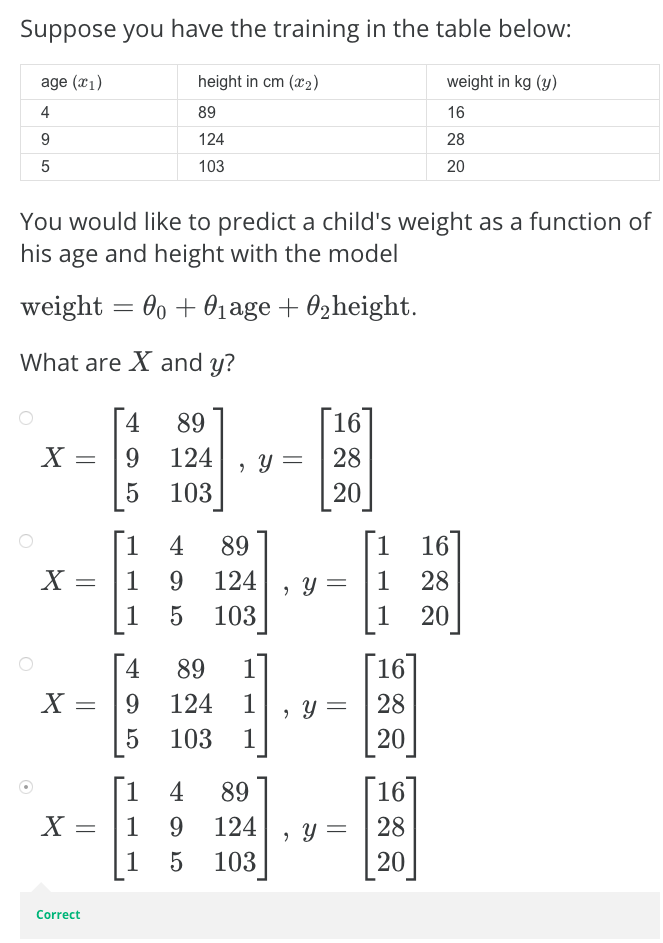
# Computing Parameters Analytically

## Normal equation

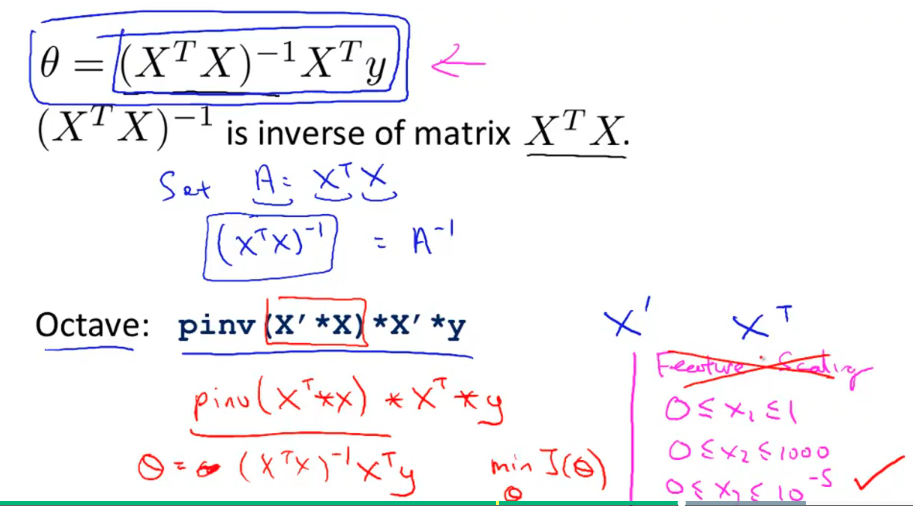
Normal equation is een methode om de theta analytisch op te lossen.

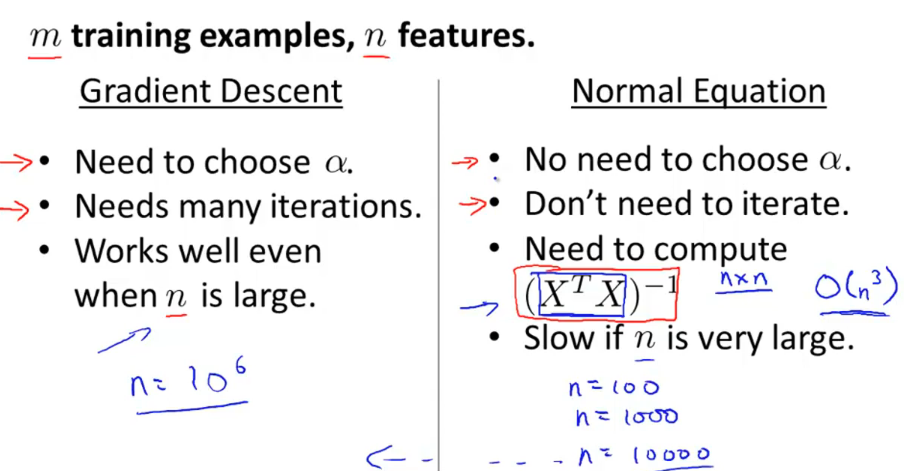


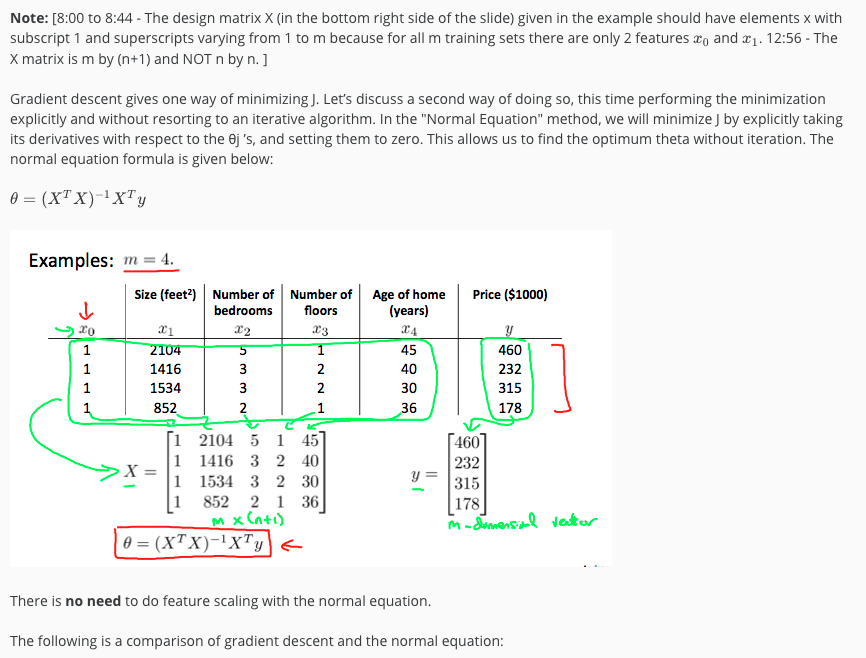


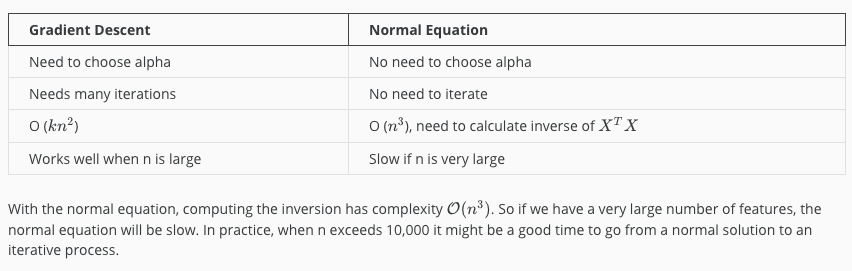


Als je Normal Equation gebruikt dan hoef je Feature Scaling niet te gebruiken.



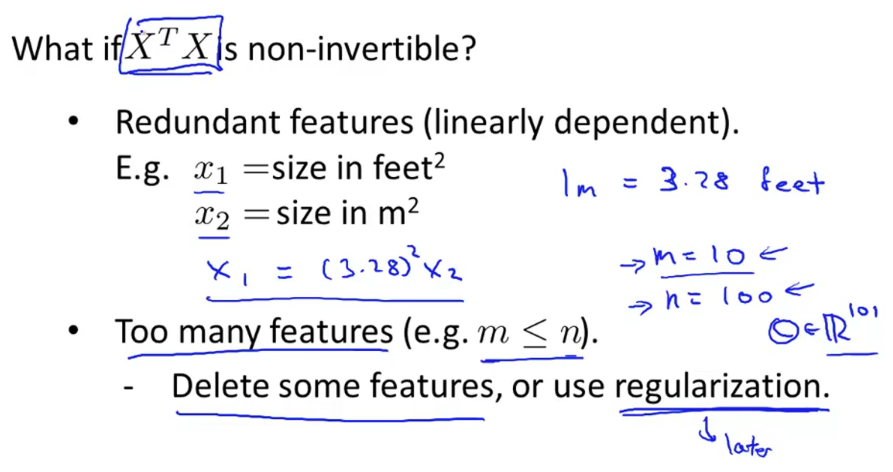


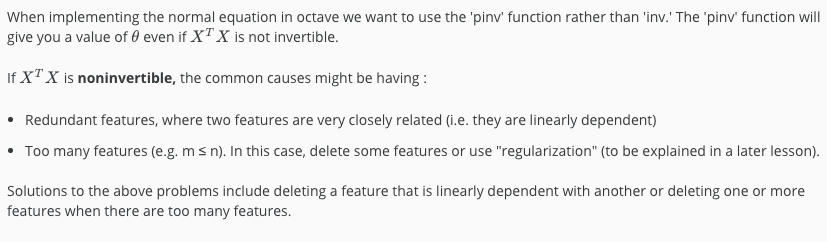




## Normal Equation Noninvertibility (Optioneel)

Bij te veel features: aantal features verwijderen of gebruik maken van regularization.





<https://github.com/mGalarnyk/datasciencecoursera/tree/master/Stanford_Machine_Learning/Week2>