Features and Polynomial Regression

We can improve our features and the form of our hypothesis function in a couple different ways.

We can **combine** multiple features into one. For example, we can combine $\mathbf{x_1}$ and $\mathbf{x_2}$ into a new feature $\mathbf{x_3}$ by taking $\mathbf{x_1} \cdot \mathbf{x_2}$.

Polynomial Regression

Our hypothesis function need not be linear (a straight line) if that does not fit the data well.

We can **change the behavior or curve** of our hypothesis function by making it a quadratic, cubic or square root function (or any other form).

For example, if our hypothesis function is $\mathbf{h_\theta}$ **(x)=** $\boldsymbol{\theta_0+\theta_1 x_1}$ then we can create additional features based on $\mathbf{x_1}$, to get the quadratic function $\mathbf{h_\theta}$ **(x)=** $\boldsymbol{\theta_0+\theta_1 x_1+\theta_2 x_1^2}$ or the cubic function $\mathbf{h_\theta}$ **(x)=** $\boldsymbol{\theta_0+\theta_1 x_1+\theta_1 x_1^2+\theta_1 x_1^3}$
In the cubic version, we have created new features $\mathbf{x_2}$ and $\mathbf{x_3}$ where $\mathbf{x_2=x_1^2}$ and $\mathbf{x_3=x_1^3}$**.**

To make it a square root function, we could do: $\mathbf{h_\theta}$ **(x)=** $\boldsymbol{\theta_0+\theta_1 x_1+\theta_2\sqrt{x_1}}$
One important thing to keep in mind is, if you choose your features this way then feature scaling becomes very important.

eg. if $\mathbf{x_1}$ has range 1 - 1000 then range of $\mathbf{x_1^2}$ becomes 1 - 1000000 and that of $\mathbf{x_1^3}$ becomes 1 - 1000000000