- CS229 notes (15-17)

$$y^{(i)} = \theta^T x^{(i)} + \epsilon^{(i)},$$

- epsilon (i) is an error term that captures either unmodeled effects or random noise.
- **Assumptions**
  - epsilon (i) is IID according to a Gaussian distribution with zero mean and some variance
    - **Note:** This is not the only assumption which can be used to justify that least-squares is the rational procedure for linear regression. The final choice of theta did not depend on variance.