Tufte TeX.jl Example with Sigmoid

Robert Moss MOSSR@CS.STANFORD.EDU
Stanford University, Stanford, CA 94305
2020-08-21

1 *Loss Function*

In mathematical optimization, statistics, decision theory and machine learning, a *loss function* or *cost function* is a function that maps an event or values of one or more variables onto a real number intuitively representing some ''cost'' associated with the event.[1] An optimization problem seeks to minimize a loss function. An objective function is either a loss function or its negative (sometimes called a *reward function* or a *utility function*), in which case it is to be maximized.

[1] https://en.wikipedia.org/wiki/Loss_function

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left[-y^{(i)}\log(h_\theta(x^{(i)})) - (1-y^{(i)})\log(1-h_\theta(x^{(i)}))\right] \tag{1}$$

```
function loss_function(theta, X, y)
    m = length(y) # number of training examples
    grad = zeros(size(theta))
    h = sigmoid(X * theta)
    J = 1/m*sum(-y'*log(h)-(1 .- y)'*log(1 .- h))
    grad = 1/m*(X'*(h-y))
    return (J, grad)
end
```

2 *Sigmoid*

A sigmoid function is a mathematical function having an "S" shape (sigmoid curve). Often, sigmoid function refers to the special case of the logistic function shown in the first figure and defined by the formula.[2]

[2] https://en.wikipedia.org/wiki/Sigmoid_function

$$\sigma(t) = \frac{1}{1+e^{-t}} \tag{2}$$

```
# return is of size/dimensions: zeros(size(t))
sigmoid(t) = 1 / (1 + exp(-t))
```