* 1. Linear regression

梯度下降

Tricks :

2.1.3 Feature scaling

2.1.4 Learning rate :

太小，拟合速度太慢； 太大，不会拟合，可以通过做x-Y （x轴：迭代次数和y轴：const fun）图看 .

学习率 如果太大，slow convergence also possible

2.1.5 多项式回归，可以变化成线性回归

J=aX1+bX1^2+cX1^3 =aY1+bY2+cY3

2.2.1 最小二乘法 公式直接求参数 ，

如果训练数据很大，求矩阵的逆太慢大约

2.2.2 公式法,如果 不可逆 情况：

1. Redundant features (linearly dependent)

E.g. =size in

=size in n

2. Too many features (E.g. mn )

Matlab 数组1为开始索引

2.3.6 Vectorization computation ： matlab 内部向量计算高度优化

3.1.1 logistic regression : classification

3.1.2 logistic regression : Hypothesis representation

g z=

h()=

3.1.3 logistic regression : decision boundary

3.2.1 logistic regression : cost fuction

cost(h(),y)= = -y-(1-y)

3.2.2 logistic regression : Gradient descent

, 其中 g

发现logistic regression 和linear regression 参数迭代公式一样，只是h()不同，最后cost function 也不同

3.2.3 logistic regression : advanced optimization (matlab内置很多最小值优化函数)

Gradient descent

Conjugate gradient

BFGS

L-BFGS

3.3.1 Logistic regression : Multiclass Classification (one vs all) 找出多少个分类的假设函数，哪个函数概率大属于哪一种

3.4.1 Regularization – the problem of overfit

3.4.2 Regularizatio – cost function

+

3.4.3 规范因子过大 则underfit，过小则overfit

4.1.1 Neural Network : 每层节点和上一层相当于logistic regression

5.1.2 Neural Networks - Backpropagation Algorithm

5.2.1 Backprogapation --unrolling parameters

5.2.2 Gradient checking 微积分定义检查得到参数是否相似 ,

5.23 Random initialization (打破对称，消除冗余计算)

NN　non-convex （非凸优化，只能找到局部最小值）

6.1.2 Evaluate a hypothesis :将数据集分成70% train data 和30% test data 防止过度拟合

6.1.3 Model selection

6.2.1 High Bias (underfitting) Hcv和Ht 都很大

High Variance(overfitting) Hcv>>Ht

6.2.2 Regulariation parameter selection

6.2.3 Learning curve