```
function [theta, J_history]
    = gradientDescent(X, y, theta, alpha, num_iters)

%GRADIENTDESCENT gradient descent to learn theta
%   updates theta by taking num_iters gradient
%   steps with learning rate alpha.

m = length(y); % number of training examples
J_history = zeros(num_iters, 1);

for iter = 1:num_iters
    theta = theta-alpha*1/m*(X'*(X*theta-y));
    J_history(iter) = computeCost(X, y, theta);
end
```