

[placeholder]6.867 Project 1[placeholder]

September 29, 2015

1

First, we discuss variations on gradient descent, including analytic gradient descent, finite difference gradient descent, and `Matlab`'s `fminunc`, which includes some variations such as adaptive step size, using a quadratic approximation instead of linear approximation, and using a 2-dimensional subspace to reduce computational complexity. First, we'll examine the analytic gradient descent, which relies on having an analytic gradient expression at all points in space.

We'll refer to three example functions: the n -dimensional quadratic "bowl", Q_n , the n -dimensional inverted Gaussian (centered at $\mu = 0$, with $\Sigma = \mathbf{I}_n$), N_n , and the n -dimensional sum of sin's, S_n . These are defined as follows (leaving off the normalization constant on the inverted Gaussian for simplicity):

$$Q_n = \|\mathbf{x}\|^2$$

$$N_n = -\exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right)$$

$$S_n = \sum_{i=1}^n \sin(x_i)$$

Next, we benchmark the gradient descent variations on these two functions (seeding with an initial guess of $(1, 1, \dots)$):

Table 1: Iterations (step = .1, threshold = .001)									
n	Analytic Gradient			Finite Differences			fminunc		
	Q_n	N_n	S_n	Q_n	N_n	S_n	Q_n	N_n	S_n
2	16	33		16	33				
5	18	60		18	60				
20	21	1		21	1				
100	25	1		25	1				

One thing we notice is that the finite differences and analytic methods yield the exact same number of iterations for Q_n . This is easily explainable by noticing $\nabla Q_n = 2\mathbf{x}$. Since the gradient is linear, the finite differences method is still an exact solution to the gradient.

We also notice that the finite differences and analytic methods yield the exact same number of iterations for N_n . This is explained by the fact that the finite difference approximation (with $\delta = .05$) differs from the gradient by just .03% between $(0, 1)$. This small difference is purely due to the gradient being very well behaved, with small values for the diagonal of the Hessian matrix for the inverted Gaussian.