Introduction to Optimization
Exercise 1

Marc Toussaint
Machine Learning & Robotics lab, U Stuttgart
Universitsttsstrae 38, 70569 Stuttgart, Germany

April 25, 2013

1 Equality Constraint Penalties and augmented Lagrangian

(We don't need to know what the Langangian is (yet) to solving this exercise.)
In the lecture we discussed the squared penalty method for inequality constraints. There is a straight-forward version for equality constraints: Instead of

$$\min_x \ f(x) \quad \text{s.t.} \quad h(x) = 0 \tag{1}$$

we address

$$\min_x \ f(x) + \mu \sum_{i=1}^m h_i(x)^2 \tag{2}$$

such that the squared penalty pulls the solution onto the constraint $h(x) = 0$. Assume that if we minimize (2) we end up at a solution x_1 for which each $h_i(x_1)$ is reasonable small, but not exactly zero.
We also mentioned the idea that we could add an additional term which counteracts the violation of the constraint. This can be realized by minimizing

$$\min_x \ f(x) + \mu \sum_{i=1}^m h_i(x)^2 + \sum_{i=1}^m \lambda_i h_i(x) \tag{3}$$

for a "good choice" of each λ_i. It turns we can infer this "good choice" from the solution x_1 of (2):
Proof that setting $\lambda_i = 2\mu h_i(x_1)$ will, if we assume that the gradients $\nabla f(x)$ and $\nabla h(x)$ are (locally) constant, ensure that the minimum of (3) fulfils exactly the constraints $h(x) = 0$.
Tip: Think intuitive. Think about how the gradient that arises from the penalty in (2) is now generated via the λ_i.

2 Squared Panalties & Log Barriers (worth 2 points)

In the last exercise we defined the "hole function" $f^c_{\text{hole}}(x)$, where we now assume a conditioning $c = 4$.
Consider the optimization problem

$$\min_x f^c_{\text{hole}}(x) \quad \text{s.t.} \quad g(x) \le 0 \tag{4}$$

$$g(x) = \begin{pmatrix} x^\top x - 1 \\ x_n + 1/c \end{pmatrix} \tag{5}$$

a) First, assume $n = 2$ ($x \in \mathbb{R}^2$ is 2-dimensional), $c = 4$, and draw on paper what the problem looks like and where you expect the optimum.

b) Implement the Squared Penalty Method. Choose as a start point $x = (\frac{1}{2}, \frac{1}{2})$. Plot its optimization path and report on the number of total function/gradient evaluations needed.

c) Test the scaling of the method for $n = 10$ dimensions.

d) Implement the Log Barrier Method and test as in b) and c). Compare the function/gradient evaluations needed.