Constrained optimization notes

James Graham

February 23, 2017

1 Introduction

This note discusses options for solving constrained optimization problems. First via Newton methods, and then by others.

2 Single variable constrained optimization

2.1 Change of variables: logistic transform

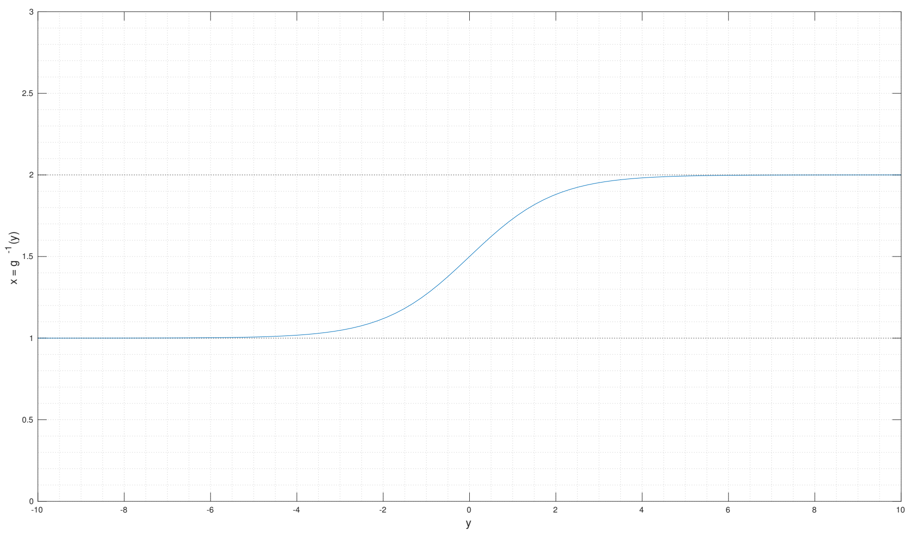
Consider the constrained optimization problem:

$$\max_{x \in [\underline{x}, \overline{x}]} f(x)$$

where the function f is twice differentiable, and the choice variable x is subject to a "box constraint". One way to solve this problem numerically is to transofrm the problem into an unconstrained problem using a logistic transformation. Let $y = g(x) \equiv \log\left(\frac{x - \underline{x}}{\overline{x} - x}\right)$. We can also re-express x as a function of y via the inverse transform: $x = \frac{\overline{x}\exp(y) + \underline{x}}{1 + \exp(y)}$. Figure 1 shows the inverse transformation subject to $\underline{x} = 1$ and $\overline{x} = 2$. Observe that $x = g^{-1}(y)$ is defined for y on the whole teal line, however x always remains within the box constraint. We rewrite the constrained optimization problem as an unconstrained problem:

$$\max_{y \in \mathbb{R}} f(g^{-1}(y))$$

Figure 1: Inverse logit transformation



Now we can solve this problem using any unconstrained optimization method. However, note that the inverse logit transform $g^{-1}(y)$ is infinitely differentiable. This means both its Jacobian and Hessian are well-defined. These are, respectively:

$$\bigtriangledown f(g^{-1}(y)) = f'(g^{-1}(y))\frac{\exp(y)(\overline{x}-\underline{x})}{(1+\exp(y))^2}$$

$$Hf(g^{-1}(y)) = f''(g^{-1}(y))\frac{\exp(y)(\overline{x}-\underline{x})}{(1+\exp(y))^2} + f'(g^{-1}(y))\frac{\exp(y)(1-\exp(y))(\overline{x}-\underline{x})}{(1+\exp(y))^3}$$

Now make an initial guess for the unconstrained variable, y_0, and we can update the guess for y via:

$$y_{k+1} = y_k - \left(Hf(g^{-1}(y))\right)^{-1}\left(\bigtriangledown f(g^{-1}(y))\right)$$

We can check for convergence of the algorithm by looking at the distance of the Jacobian from zero (at an unconstrained optimum, the Jacobian equals zero).

2.2 Change of variables: smooth function of m

We can transform the constrained optimization problem to an unconstrained problem in other ways.[1] Suppose, again, we have the problem:

$$\max_{x \in [\underline{x}, \overline{x}]} f(x)$$

The Lagrangian for this problem is:

$$\mathcal{L}(x) = f(x) + \lambda_1(x - \underline{x}) + \lambda_2(\bar(x) - x)$$

where λ_1 and λ_2 are the Lagrange multipliers on the lower and upper constraints for x. The Kuhn Tucker conditions for this problem are:

$$\begin{aligned} f'(x) + \lambda_1 - \lambda_2 = 0 \\ x \geq \underline{x} \\ x \leq \overline{x} \\ \lambda_1, \lambda_2 \geq 0 \end{aligned}$$

Which can be written as:

$$\begin{aligned} f'(x) + \lambda_1 - \lambda_2 = 0 \\ \lambda_1(x - \underline{x}) = 0 \\ \lambda_2(\overline{x} - x) = 0 \end{aligned}$$

We want to specific a variable m, and functions $x(m), \lambda_1(m), \lambda_2(m)$, that have a continuous first derivative, and satisfy the Kuhn Tucker conditions. A suggested set of such functions are:

$$\begin{aligned} x(m) &= \begin{cases} \underline{x} & \text{, if } m < 0 \\ \underline{x} + am^2 & \text{, if } m \in [0, \frac{1}{2}) \\ \overline{x} - a(1-m)^2 & \text{, if } m \in [\frac{1}{2}, 1] \\ \overline{x} & \text{, if } m > 1 \end{cases} \\ \lambda_1(m) &= \begin{cases} m^2 & \text{, if } m < 0 \\ 0 & \text{, if } m \geq 0 \end{cases} \\ \lambda_2(m) &= \begin{cases} 0 & \text{, if } m \leq 1 \\ (m-1)^2 & \text{, if } m > 1 \end{cases} \end{aligned}$$

[1]This is based on a note by Simon Mongey (2016).

where $a = 2(\overline{x} - \underline{x})$. Now we want to minimize the Lagrangian function $\mathcal{L}(m)$ after this change of variable. Note that the Jacobian and Hessian of the Lagrangian function are:

$$\begin{aligned}
\bigtriangledown\mathcal{L}(m) &= f'(x(m)) + \lambda_1(m) - \lambda_2(m) \\
H\mathcal{L}(m) &= f''(x(m))x'(m) + \lambda_1'(m) - \lambda_2'(m)
\end{aligned}$$

where the derivatives are given by:

$$\begin{aligned}
x'(m) &= \begin{cases} 0 & \text{, if } m < 0 \\ 2am & \text{, if } m \in [0, \frac{1}{2}) \\ 2a(1-m) & \text{, if } m \in [\frac{1}{2}, 1] \\ 0 & \text{, if } m > 1 \end{cases} \\
\lambda_1'(m) &= \begin{cases} 2m & \text{, if } m < 0 \\ 0 & \text{, if } m \geq 0 \end{cases} \\
\lambda_2'(m) &= \begin{cases} 0 & \text{, if } m \leq 1 \\ 2(m-1) & \text{, if } m > 1 \end{cases}
\end{aligned}$$

Finally, we compute the Newton steps by updating m via:

$$m_{k+1} = m_k - (H\mathcal{L}(m))^{-1} (\bigtriangledown\mathcal{L}(m))$$

We can check for convergence of the algorithm by looking at the distance of the Jacobian from zero (at an unconstrained optimum, the Jacobian of the Lagrangian function equals zero).

3 Multivariate constrained optimization

Let's use the logit transform again. Consider the two variable case. Our constrained optimization problem is:

$$\max_{x_1 \in [\underline{x}_1, \overline{x}_1], x_2 \in [\underline{x}_2, \overline{x}_2]} f(x_1, x_2)$$

Using the logit transform, let

$$\begin{aligned}
y_1 &= g_1(x_1) \equiv \log\left(\frac{x_1 - \underline{x}_1}{\overline{x}_1 - x_1}\right) \\
y_2 &= g_2(x_2) \equiv \log\left(\frac{x_2 - \underline{x}_2}{\overline{x}_2 - x_2}\right)
\end{aligned}$$

notice that we use functions that are separate across the two choice variables. We express the inverse transforms as we did above:

$$
\begin{aligned}
x_1 &= \frac{\overline{x_1}\exp(y_1) + \underline{x_1}}{1+\exp(y_1)} \\
x_2 &= \frac{\overline{x_2}\exp(y_2) + \underline{x_2}}{1+\exp(y_2)}
\end{aligned}
$$

Now our unconstrained problem is:

$$
\max_{(y_1,y_2)\in\mathbb{R}^2} f(g_1^{-1}(y_1), g_2^{-1}(y_2))
$$

We compute the Jacobian and Hessian functions as:

$$
\begin{aligned}
\bigtriangledown f &= \left[\frac{\partial f}{\partial y_1}\frac{\exp(y_1)(\overline{x}_1-\underline{x}_1)}{(1+\exp(y_1))^2}\ , \quad \frac{\partial f}{\partial y_2}\frac{\exp(y_2)(\overline{x}_2-\underline{x}_2)}{(1+\exp(y_2))^2}\right] \\
Hf &= \begin{bmatrix} \frac{\partial^2 f}{\partial y_1^2}\frac{\exp(y_1)(\overline{x}_1-\underline{x}_1)}{(1+\exp(y_1))^2} + \frac{\partial f}{\partial y_1}\frac{\exp(y_1)(1-\exp(y_1))(\overline{x}_1-\underline{x}_1)}{(1+\exp(y_1))^3}\ , & 0 \\ 0\ , & \frac{\partial^2 f}{\partial y_2^2}\frac{\exp(y_2)(\overline{x}_2-\underline{x}_2)}{(1+\exp(y_2))^2} + \frac{\partial f}{\partial y_2}\frac{\exp(y_2)(1-\exp(y_2))(\overline{x}_2-\underline{x}_2)}{(1+\exp(y_2))^3} \end{bmatrix}
\end{aligned}
$$

Notice that the Hessian is diagonal. This follows from the fact that the logit transforms of x_1 and x_2 are independent, so that the cross-partial derivatives are zero.

Finally, we can update the unconstrained solution via Newton's method with:

$$
\begin{bmatrix} y_{1,k+1} \\ y_{2,k+1} \end{bmatrix} = \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix} - (Hf)^{-1}(\bigtriangledown f)^T
$$

We can check for convergence of the algorithm by looking at the distance of the Jacobian from zero by taking the norm of the Jacobian (at an unconstrained optimum, the Jacobian equals zero).

Appendix A