Lecture X
Finding the Minimum Using Newton's Method

I. Finding the Univariate Minimum (Algorithm1.ma)

A. A Sufficiently Complex Function

1. It is obvious that finding the minimum of a simple univariate quadratic function is trivial given the rules we discussed in the preceding section. For example

$$U(x) = 5x^2 - 4x + 2$$

has a minimum determined by its first derivative

$$\frac{\partial U(x)}{\partial x} = 10x - 4 = 0$$
$$x = \frac{2}{5}.$$

In addition, straightforward transformations such as

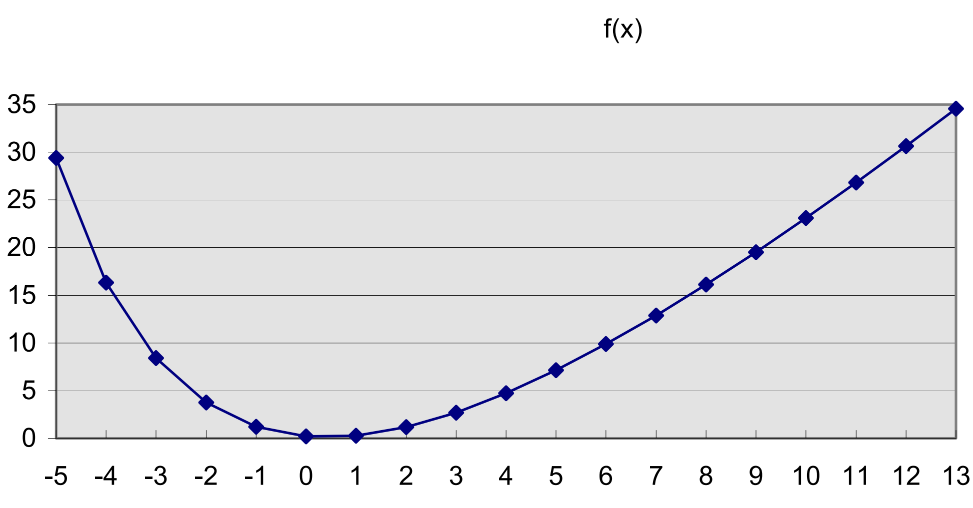
$$e^{5x^2 - 4x + 2}$$

offer little additional complexity

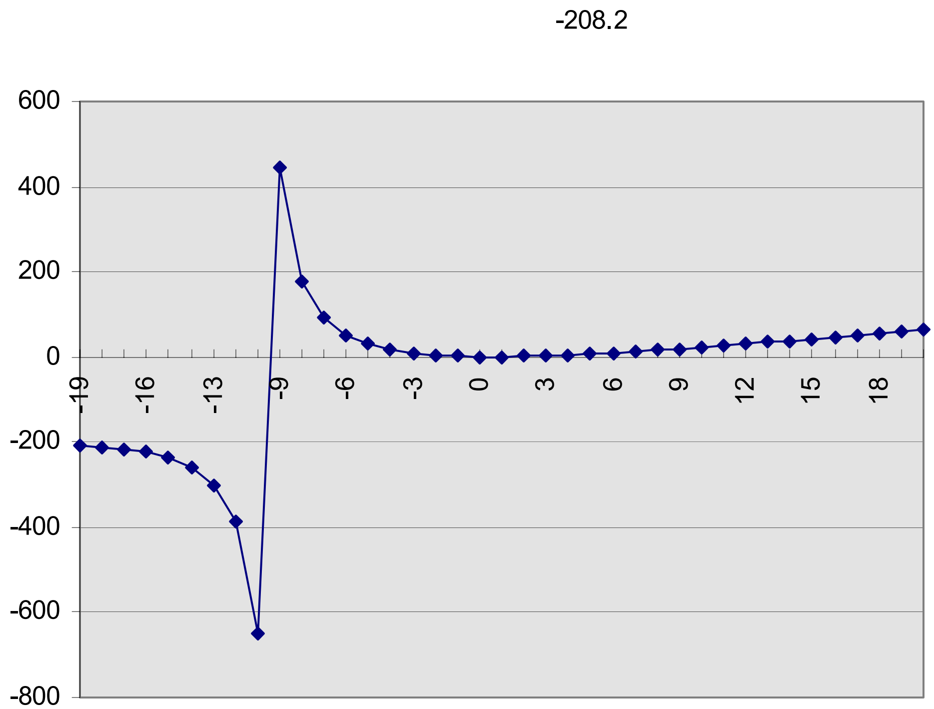
$$\frac{\partial U(x)}{\partial x} = U(x)\left[10x - 4\right] = 0$$
$$x = \frac{2}{5}.$$

2. One possibility is the rational function

$$f(x) = \frac{5x^2 - 4x + 2}{x + 10}.$$



However, these functions typically have bizarre discontinuities. For example, plotting the above function over a broader range indicates that



If we restrict our attention to the range of x's such that x is greater than -10 the problem becomes more tractable.

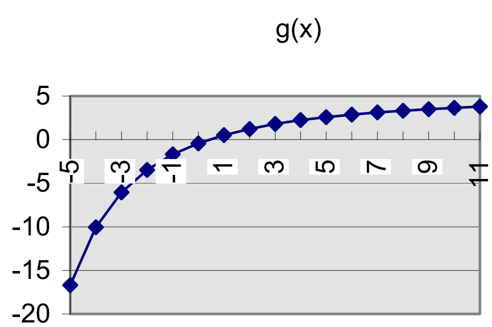
$$\frac{\partial f(x)}{\partial x} = \frac{10x-4}{10-x} - \frac{5x^2-4x+2}{(x+10)^2} = 0$$

$$x = \frac{-100+2\sqrt{2710}}{10}.$$

$$x = 0.411532.$$

B. Solving the Zero

1. As I have previously stated, the trick to optimization is to find the zero of the gradient. Plotting the gradient of the rational function, we see



2. The Method of bisection

a. The method of bisection involves shrinking the interval. Specifically, suppose that the original interval was [-8,20]. The midpoint of this interval is x=6. At x=6, the value of the derivative is 2.88281. Therefore, we throw away to higher bisection yielding an interval of [-8,6]. The midpoint of this interval is -1 with a derivative value of -1.69136. At this point we throw out the lower interval.
b. The sequence of determining the point were the derivative is equal to zero yields

	Lower	Upper	Midpoint	Gradient	Replace
0	-8.0000	20.0000	6.0000	2.8828	Upper
1	-8.0000	6.0000	-1.0000	-1.6914	Lower
2	-1.0000	6.0000	2.5000	1.5312	Upper
3	-1.0000	2.5000	0.7500	0.3099	Upper
4	-1.0000	0.7500	-0.1250	-0.5581	Lower
5	-0.1250	0.7500	0.3125	-0.0965	Lower
6	0.3125	0.7500	0.5313	0.1130	Upper
7	0.3125	0.5313	0.4219	0.0099	Upper
8	0.3125	0.4219	0.3672	-0.0429	Lower
9	0.3672	0.4219	0.3945	-0.0164	Lower
10	0.3945	0.4219	0.4082	-0.0032	Lower
11	0.4082	0.4219	0.4150	0.0034	Upper
12	0.4082	0.4150	0.4116	0.0001	Upper
13	0.4082	0.4116	0.4099	-0.0016	Lower
14	0.4099	0.4116	0.4108	-0.0007	Lower

c. The stopping place is somewhat arbitrary.

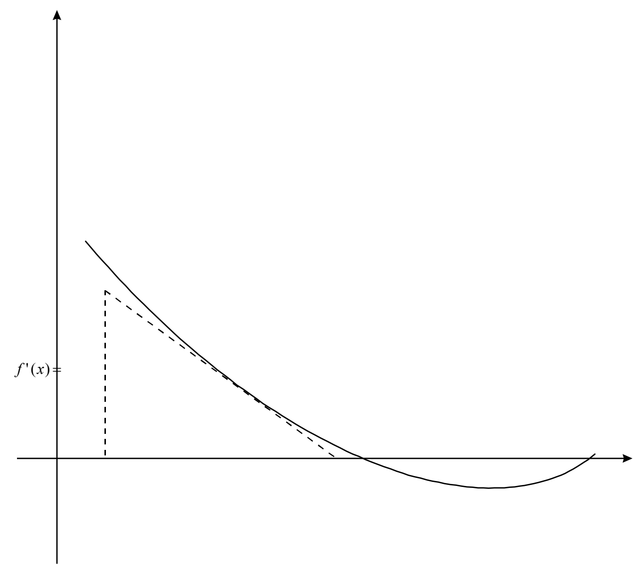
3. Newton's Method
   a. The theoretical foundation of Newton's method involves inscribing triangles inside the function. Mathematically

$$\Delta y \cong \frac{\partial f(x)}{\partial x} \Delta x$$

$$\Delta x = \frac{\Delta y}{\frac{\partial f(x)}{\partial x}} .$$

$$x_{t+1} = x_t - \frac{f(x)}{\frac{\partial f(x)}{\partial x}}$$

Graphically, this result can be depicted:

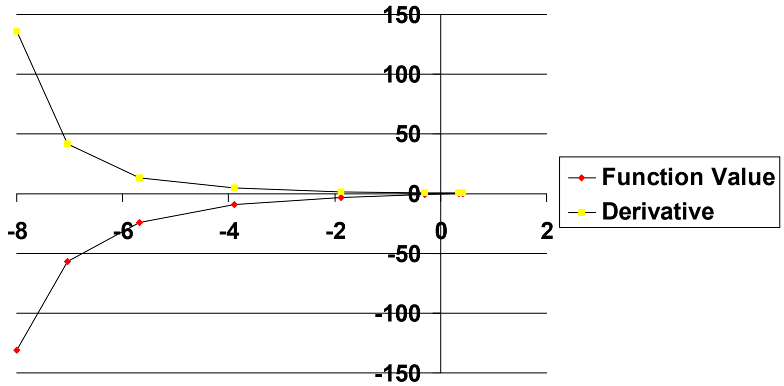


Thus, the Newton search points are

Search Point	Function Value	Derivative	Step
-8.0000	-130.5000	135.5000	0.9631
-7.0369	-56.7315	41.6668	1.3616
-5.6754	-23.9799	13.4022	1.7893
-3.8861	-9.4998	4.7432	2.0028
-1.8833	-3.2269	2.0272	1.5919
-0.2914	-0.7503	1.1846	0.6334
0.3419	-0.0675	0.9800	0.0689
0.4108	-0.0007	0.9607	0.0007
0.4115	0.0000	0.9605	0.0000

Graphically, the search path can be depicted

Newton Search



II. Finding the Multivariate Maximum

A. The basic difference between univariate and multivariate optimization is the number of equations which we want to solve simultaneously for zero. In the multivariate case we want to solve

$$\nabla_x f(x) = 0$$

where x is an n element vector, so we want to solve for n equations equal to zero. Appealing again to the second order Taylor series expansion

$$\nabla_x f(x) = \nabla_x f(x^*) + \nabla^2_{xx} f(x^*)(x - x^*) = 0$$

$$(x - x^*) = \left(\nabla^2_{xx} f(x^*)\right)^{-1} \nabla_x f(x)$$

which implies that

$$x_{t+1} = x_t - \left(\nabla^2_{xx} f(x_t)\right)^{-1} \nabla_x f(x_t)$$

B. A Simple Problem

1. As a first problem, consider a Cobb-Douglas utility function with a budget constraint imposed

$$\max_x x_1^{.2} x_2^{.3} x_3^{.4} x_4^{.1}$$
$$st\ x_1 + 2x_2 + 3x_3 + x_4 = 100 .$$

By substitution, this problem becomes

$$\max_x x_2^{.3} x_3^{.4} (100 - 2x_2 - 3x_3 - x_4)^{.2} x_4^{.1}$$

Starting with point x=(1,1,1), we have

$$\nabla_x f(x) = \begin{pmatrix} .7337 & .9766 & .2428 \end{pmatrix}$$

$$\nabla^2_{xx} f(x) = \begin{pmatrix} -.5275 & .2885 & .0717 \\ .2885 & -.6085 & .0954 \\ .0717 & .0954 & -.2244 \end{pmatrix}$$

$$x_{t+1} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} - \nabla_x f(x) \left(\nabla^2_{xx} f(x) \right)^{-1} = \begin{pmatrix} 5.3559 & 5.3479 & 5.3226 \end{pmatrix}$$

C. Steepest Descent

1. Replacing the inverse of the Hessian matrix in the above example with the negative of the identity matrix yields the steepest descent algorithm