3. Quadratic and Cubic Search for a Minimum

Background for Search Methods

An approach for finding the minimum of $f(x)$ in a given interval is to evaluate the function many times and search for a local minimum. To reduce the number of function evaluations it is important to have a good strategy for determining where $f(x)$ is to be evaluated. Two efficient bracketing methods are the golden ratio and Fibonacci searches. To use either bracketing method for finding the minimum of $f(x)$, a special condition must be met to ensure that there is a proper minimum in the given interval.

Definition (Unimodal Function) The function $f(x)$ is unimodal on $I = [a, b]$, if there exists a unique number $p \in I$ such that

$f(x)$ is decreasing on $[a, p]$,
and
$f(x)$ is increasing on $[p, b]$.

Minimization Using Derivatives

Suppose that $f(x)$ is unimodal over $[a, b]$ and has a unique minimum at $x = p$. Also, assume that $f'(x)$ is defined at all points in (a, b). Let the starting value p_0 lie in (a, b). If $f'(p_0) < 0$, then the minimum point p lies to the right of p_0. If $f'(p_0) > 0$, then the minimum point p lies to the left of p_0.

Background for Bracketing the Minimum

Our first task is to obtain three test values,

(1) $p_0,\ p_1 = p_0 + h,\ \text{and}\ p_2 = p_0 + 2h,$
so that
(2) $f(p_0) > f(p_1),\ \text{and}\ f(p_1) < f(p_2).$

Suppose that $f'(p_0) < 0$; then $p_0 < p$ and the step size h should be chosen positive. It is an easy task to find a value of h so that the three points in (1) satisfy (2). Start with $h = 1$ in formula (1) (provided that $a + 1 < b$); if not, take $h = \frac{1}{2}$, and so on.

Case (i) If (2) is satisfied we are done.

Case (ii) If $f(p_0) > f(p_1),\ \text{and}\ f(p_1) > f(p_2)$, then $p_2 < p$.
We need to check points that lie farther to the right. Double the step size and repeat the process.

Case (iii) If $f(p_0) < f(p_1)$, we have jumped over p and h is too large.
We need to check values closer to p_0. Reduce the step size by a factor of $\frac{1}{2}$ and repeat the process.

When $f'(p_0) > 0$, the step size h should be chosen negative and then cases similar to (i), (ii), and (iii) can be used.

Quadratic Approximation to Find p

Finally, we have three points (1) that satisfy (2). We will use quadratic interpolation to find p_{min}, which is an approximation to p. The Lagrange polynomial based on the nodes in (1) is

$$(3) \qquad Q(x) = y_0 \frac{(x-p_1)(x-p_2)}{2h^2} - y_1 \frac{(x-p_0)(x-p_2)}{h^2} + y_2 \frac{(x-p_0)(x-p_1)}{2h^2},$$

where $y_i = f(p_i)$ for $i = 1, 2, 3$.

The derivative of $Q(x)$ is

$$(4) \qquad Q'(x) = y_0 \frac{(2x-p_1-p_2)}{2h^2} - y_1 \frac{(2x-p_0-p_2)}{h^2} + y_2 \frac{(2x-p_0-p_1)}{2h^2}.$$

Solving $Q'(x) = 0$ in the form $Q(p_0 + h_{min}) = 0$ yields

$$(5) \qquad 0 = y_0 \frac{(2(p_0+h_{min})-p_1-p_2)}{2h^2} - y_1 \frac{(4(p_0+h_{min})-2p_0-2p_2)}{2h^2} + y_2 \frac{(2(p_0+h_{min})-p_0-p_1)}{2h^2}.$$

Multiply each term in (5) by $2h^2$ and collect terms involving h_{min}:

$$0 = h_{min}(2y_0 - 4y_1 + 2y_2) + y_0(2p_0 - p_1 - p_2) - y_1(4p_0 - 2p_0 - 2p_2) + y_2(2p_0 - p_0 - p_1)$$

$$0 = h_{min}(2y_0 - 4y_1 + 2y_2) + y_0(2p_0 - p_1 - p_2) - y_1(2p_0 - 2p_2) + y_2(p_0 - p_1)$$

$$0 = h_{min}(2y_0 - 4y_1 + 2y_2) - 3hy_0 + 4hy_1 - hy_2$$

$$0 = -h_{min}2(2y_1 - y_0 - y_2) + h(4y_1 - 3y_0 - y_2)$$

This last quantity is easily solved for h_{min}:

$$h_{min} = \frac{h(4y_1 - 3y_0 - y_2)}{2(2y_1 - y_0 - y_2)}.$$

The value $p_{min} = p_0 + h_{min}$ is a better approximation to p than p_0. Hence we can replace p_0 with p_{min} and repeat the two processes outlined above to determine a new h and a new h_{min}. Continue the iteration until the desired accuracy is achieved. In this algorithm the derivative of the objective function $f(x)$ was used implicitly in (4) to locate the minimum of the interpolatory quadratic. The reader should note that the subroutine makes no explicit use of the derivative.

Cubic Approximation to Find p

We now consider an approach that utilizes functional evaluations of both $f(x)$ and $f'(x)$. An alternative approach that uses both functional and derivative evaluations explicitly is to find the minimum of a third-degree polynomial that interpolates the objective function $f(x)$ at two points. Assume that $f(x)$ is unimodal and differentiable on $[a, b]$, and has a unique minimum at $x = p$. Let $p_0 = a$. Any good step size h can be used to start the iteration. The Mean Value Theorem could be used to obtain $f'(c) = \frac{f(a+h) - f(a)}{h}$ and if $a+h$ was just to the right of the minimum, then the slope $f'(a)$ might be twice $f'(c)$ which would mean that $\frac{1}{2} f'(a) = \frac{f(a+h) - f(a)}{h}$ we do not know how much further to the right $c = a+h$ lies, so we can imagine that $f(a+h)$ is close to $f(b)$ and estimate h with the formula:

$$h = \frac{2(f[b] - f[a])}{f'[a]}.$$

Thus $p_1 = p_0 + h$. The cubic approximating polynomial $P[x]$ is expanded in a Taylor series about $x = p_2$ (which is the abscissa of the minimum). At the minimum we have $P'[p_2] = 0$, and we write $P[x]$ in the form:

$$(6) \qquad P[x] = \frac{\alpha}{h^3}(x - p_2)^3 + \frac{\beta}{h^2}(x - p_2)^2 + f[p_2],$$

and

(7) $$P'[x] = \frac{3\alpha}{h^3}(x - p_2)^2 + \frac{2\beta}{h^2}(x - p_2).$$

The introduction of $h = p_1 - p_0$ in the denominators of (6) and (7) will make further calculations less tiresome. It is required that $P[p_0] = f[p_0]$, $P[p_1] = f[p_1]$, $P'[p_0] = f'[p_0]$, and $P'[p_1] = f'[p_1]$. To find p_2 we define:

(8) $$p_2 = p_0 + \gamma h,$$

and we must go through several intermediate calculations before we end up with γ.

Use use (6) to obtain

$$f[p_1] - f[p_0] = \frac{\alpha (p_1 - p_2)^3 - \alpha (p_0 - p_2)^3}{h^3} + \frac{\beta (p_1 - p_2)^2 - \beta (p_0 - p_2)^2}{h^2}$$

Then use (8) to get

$$f[p_1] - f[p_0] = \alpha \gamma^3 + \frac{\alpha (-h\gamma - p_0 + p_1)^3}{h^3} + \frac{\beta (-h\gamma - p_0 + p_1)^2}{h^2} - \beta \gamma^2$$

Then substitute $h = p_1 - p_0$ and we have

(9) $$F = f[p_1] - f[p_0] = \alpha (3\gamma^2 - 3\gamma + 1) + \beta (1 - 2\gamma)$$

Use use (7) to obtain

$$f'[p_1] - f'[p_0] = \frac{3\alpha (p_1 - p_2)^2 - 3\alpha (p_0 - p_2)^2}{h^3} + \frac{2\beta (p_1 - p_2) - 2\beta (p_0 - p_2)}{h^2}$$

$$h(f'[p_1] - f'[p_0]) = \frac{3\alpha (p_1 - p_2)^2 - 3\alpha (p_0 - p_2)^2}{h^2} + \frac{2\beta (p_1 - p_2) - 2\beta (p_0 - p_2)}{h}$$

Then use (8) to get

$$h(f'[p_1] - f'[p_0]) = \frac{3\alpha (-h\gamma - p_0 + p_1)^2 - 3h^2\alpha\gamma^2}{h^2} + \frac{2\beta (-h\gamma - p_0 + p_1) + 2h\beta\gamma}{h}$$

Then substitute $h = p_1 - p_0$ and we have

(10) $$G = h(f'[p_1] - f'[p_0]) = 3\alpha (1 - 2\gamma) + 2\beta$$

Finally, use (7) and write

$$h f'[p_0] = h\left(\frac{3\alpha (p_0 - p_2)^2}{h^3} + \frac{2\beta (p_0 - p_2)}{h^2}\right)$$

Then use (8) to get

(11) $$h f'[p_0] = 3\alpha\gamma^2 - 2\beta\gamma$$

Now we will use the three nonlinear equations (9), 10), (11) listed below in (12). The order of determining the variables will be F, G, α, γ (the variable β will be eliminated).

$$F = \alpha (3\gamma^2 - 3\gamma + 1) + \beta (1 - 2\gamma)$$

(12) $$G = 3\alpha(1-2\gamma)+2\beta$$
$$h f'[p_0] = 3\alpha\gamma^2 - 2\beta\gamma$$

First, we will find α which is accomplished by combining the equation in (12) as follows:

$$G-2(F-h f'[p_0]) = 3\alpha(1-2\gamma)+2\beta - 2(\alpha(3\gamma^2-3\gamma+1)+\beta(1-2\gamma)-(3\alpha\gamma^2-2\beta\gamma))$$

Straightforward simplification yields $G-2(F-h f'[p_0]) = \alpha$, therefore α is given by

(13) $$\alpha = G-2(F-h f'[p_0]).$$

Second, we will eliminate β by combining the equation in (12) as follows, multiply the first equation by γ and add it to the third equation

$$G\gamma = 3\alpha\gamma - 6\alpha\gamma^2 + 2\beta\gamma$$
$$h f'[p_0] = 3\alpha\gamma^2 - 2\beta\gamma$$

$$G\gamma + h f'[p_0] = 3\alpha\gamma - 3\alpha\gamma^2$$

which can be rearranged in the form

$$G\gamma - 3\alpha\gamma + 3\alpha\gamma^2 + h f'[p_0] = 0$$

Now the quadratic equation can be used to solve for γ

$$\gamma = \frac{(G-3\alpha)-\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}{-6\alpha}$$

It will take a bit of effort to simplify this equation into its computationally preferred form.

$$\frac{(G-3\alpha)-\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}{-6\alpha} = \frac{(G-3\alpha)+\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}{(G-3\alpha)+\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}\ \frac{(G-3\alpha)-\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}{-6\alpha}$$

$$\frac{(G-3\alpha)-\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}{-6\alpha} = \frac{12h\alpha f'[p_0]}{\left((G-3\alpha)+\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}\right)(-6\alpha)}$$

$$\frac{(G-3\alpha)-\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}{-6\alpha} = \frac{-2h f'[p_0]}{(G-3\alpha)+\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}$$

Hence,

(14) $$\gamma = \frac{-2h f'[p_0]}{(G-3\alpha)+\sqrt{(G-3\alpha)^2-12h\alpha f'[p_0]}}$$

Therefore, the value of p_2 is found by substituting the calculated value of γ in (14) into the formula $p_2 = p_0 + \gamma h$. To continue the iteration process, let $h = p_2 - p_1$ and replace p_0 and p_1 with p_1 and p_2, respectively, in formulas (12), (13), and (14). The algorithm outlined above is not a bracketing method. Thus determining stopping criteria becomes more problematic. One technique would be to require that $|f'(p_n)| < \epsilon$, since $f'(p) = 0$.

Example 1. Find the minimum of the function $f[x] = x^2 - \mathrm{Sin}[x]$ on the interval $[0, 1]$ using the quadratic search method.
Solution 1.

Algorithm (Cubic Search for a Minimum). To numerically approximate the minimum of $f(x)$ on the interval $[a, b]$ by using a "cubic interpolative" search. Proceed with the method only if $f(x)$ is a unimodal function on the interval $[a, b]$.

Start with $p_0 = a$, $p_1 = b$, and $h_0 = \frac{2(f[p_1] - f[p_0])}{f'[p_0]}$ and use the following sequence steps in the iteration:

$$F = f[p_{k+1}] - f[p_k]$$

$$G = h_k (f'[p_{k+1}] - f'[p_k])$$

$$\alpha = G - 2(F - h_k f'[p_k])$$

$$\gamma = \frac{-2 h_k f'[p_k]}{(G - 3\alpha) + \sqrt{(G - 3\alpha)^2 - 12 h_k \alpha f'[p_k]}}$$

$$p_{k+2} = p_k + \gamma h_k$$

$$h_{k+1} = p_{k+2} - p_{k+1}$$

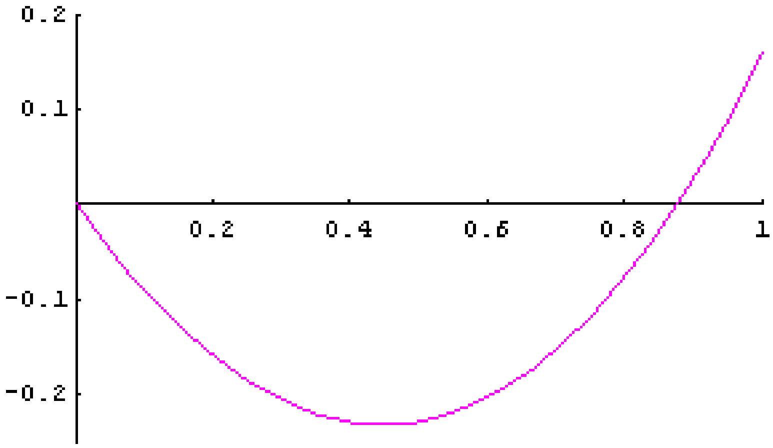
for $k = 0, 1, 2, \ldots, n$.

Example 2. Find the minimum of $f[x] = \frac{1}{2} + x^5 - \frac{4}{5} x$ on the interval $[0, 1]$ using the quadratic search method.

Solution 2.

Example 1. Find the minimum of the function $f[x] = x^2 - Sin[x]$ on the interval $[0, 1]$ using the quadratic search method.

Solution 1.



$$y = f[x] = x^2 - Sin[x]$$

Set $p_0 = a$, and $p_1 = b$, and $h = \frac{b-a}{2}$.

$$f[x] = x^2 - Sin[x]$$
$$h = 0.5$$
$$p_0 = 0.$$
$$p_1 = 0.5$$
$$p_2 = 1.$$

$$Y_0 = f[p_0] = f[\ 0.00000000\ ] = 0.$$
$$Y_1 = f[p_1] = f[\ 0.50000000\ ] = -0.229426$$
$$Y_2 = f[p_2] = f[\ 1.00000000\ ] = 0.158529$$

We must decide whether to use p_0, p_1, p_2 or change the step size to $\frac{h}{2}$ or $2h$ and recompute some of these values.

$f[x] = x^2 - \mathrm{Sin}[x]$

$h = 0.5$

$p_0 = 0.$

$p_1 = 0.5$

$p_2 = 1.$

$Y_0 = f[p_0] = f[\ 0.00000000\ ] = 0.$

$Y_1 = f[p_1] = f[\ 0.50000000\ ] = -0.229426$

$Y_2 = f[p_2] = f[\ 1.00000000\ ] = 0.158529$

The step size has not changed and we now compute $p_{min} = p_0 + H_{min}$.

$f[x] = x^2 - \mathrm{Sin}[x]$

$h = 0.5$

$p_0 = 0.$

$p_1 = 0.5$

$p_2 = 1.$

$Y_0 = f[p_0] = f[\ 0.00000000\ ] = 0.$

$Y_1 = f[p_1] = f[\ 0.50000000\ ] = -0.229426$

$Y_2 = f[p_2] = f[\ 1.00000000\ ] = 0.158529$

$$p_{min} = \frac{h\,(4\,Y_1 - 3\,Y_0 - Y_2)}{2\,(2\,Y_1 - Y_0 - Y_2)}$$

$p_{min} = h\,(4\,Y_1 - 3\,Y_0 - Y_2)\,/\,2\,(2\,Y_1 - Y_0 - Y_2)$

$p_{min} = 0.5(-1.07623)/(2(-0.61738))$

$p_{min} = 0.5(-1.07623)/(-1.23476)$

$p_{min} = p_0 + H_{min}$

$p_{min} = 0. + (0.435806)$

$p_{min} = 0.435806$

$Y_{min} = f[p_{min}] = f[\ 0.43580575\ ] = -0.232214$

We use the "latest three values" and change the step size and continue the iteration with the following three values used in the interpolating quadratic.

$f[x] = x^2 - \mathrm{Sin}[x]$

$h = 0.0641943$

$p_0 = 0.435806$

$p_1 = 0.5$

$p_2 = 0.564194$

$Y_0 = f[p_0] = f[\ 0.43580575\ ] = -0.2322143232$

$Y_1 = f[p_1] = f[\ 0.50000000\ ] = -0.2294255386$

$Y_2 = f[p_2] = f[\ 0.56419425\ ] = -0.2164199625$

The last three values will be labeled p_3, p_4, p_5 by the subroutine. The list of computations obtained by using the `Quadraticearch` subroutine are:

$p_0 = 0.00000000,\ f[\ 0.00000000\ ] = -0.2322143232$

$p_1 = 0.50000000,\ f[\ 0.50000000\ ] = -0.2294255386$

$p_2 = 1.00000000,\ f[\ 1.00000000\ ] = 0.1585290152$

$p_3 = 0.43580575,\ f[\ 0.43580575\ ] = -0.2322143232$

$p_4 = 0.50000000,\ f[\ 0.50000000\ ] = -0.2294255386$

$p_5 = 0.56419425,\ f[\ 0.56419425\ ] = -0.2164199625$

$q_{5,0} = 0.46790287,\ f[\ 0.46790287\ ] = -0.2320824604$

$q_{5,1} = 0.45185431,\ f[\ 0.45185431\ ] = -0.2324621759$

$p_6 = 0.45016827,\ f[\ 0.45016827\ ] = -0.2324655749$

$p_7 = 0.45185431,\ f[\ 0.45185431\ ] = -0.2324621759$

$p_8 = 0.45354035,\ f[\ 0.45354035\ ] = -0.2324518503$

$q_{8,0} = 0.45101129,\ f[\ 0.45101129\ ] = -0.2324647410$

$q_{8,1} = 0.45058978,\ f[\ 0.45058978\ ] = -0.2324653743$

$q_{8,2} = 0.45037903,\ f[\ 0.45037903\ ] = -0.2324655287$

$q_{8,3} = 0.45027365,\ f[\ 0.45027365\ ] = -0.2324655653$

$q_{8,4} = 0.45022096,\ f[\ 0.45022096\ ] = -0.2324655735$

$q_{8,5} = 0.45019462,\ f[\ 0.45019462\ ] = -0.2324655750$

$p_9 = 0.45018361,\ f[\ 0.45018361\ ] = -0.2324655752$

Let us compare these answers with *Mathematica*'s subroutine `FindMinimum`.

$f[x] = x^2 - \mathrm{Sin}[x]$

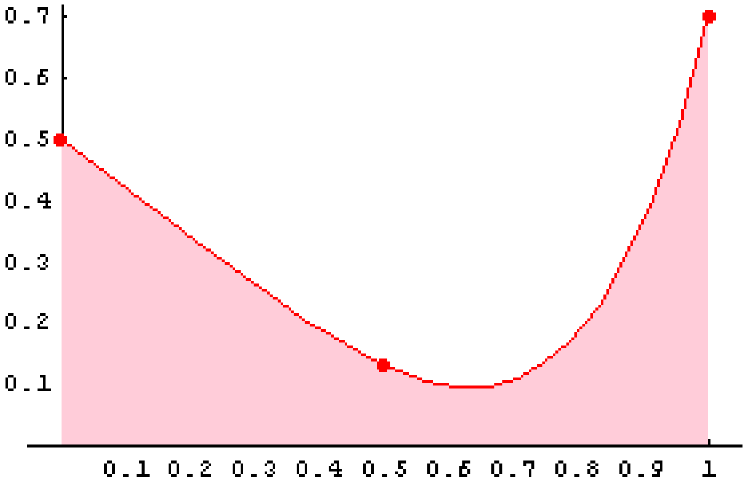
$\{-0.232466, \{x \to 0.450184\}\}$

$p = 0.45018361$

$f[p] = -0.23246558$

Example 2. Find the minimum of $f[x] = \frac{1}{2} + x^5 - \frac{4}{5}x$ on the interval $[0, 1]$ using the quadratic search method.

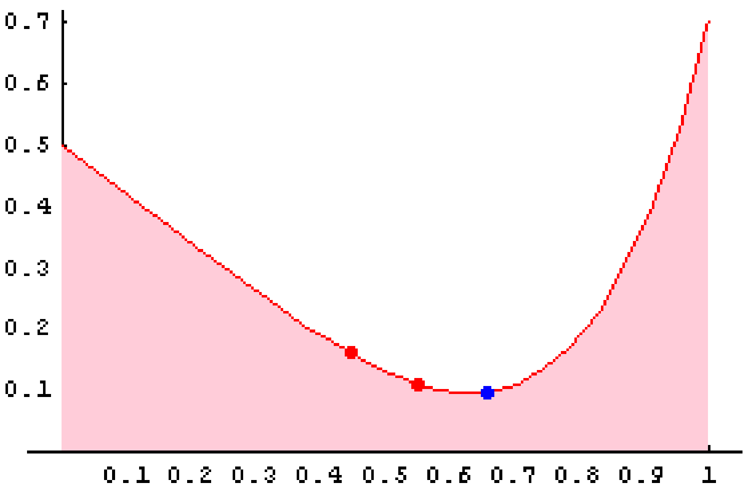
Solution 2.



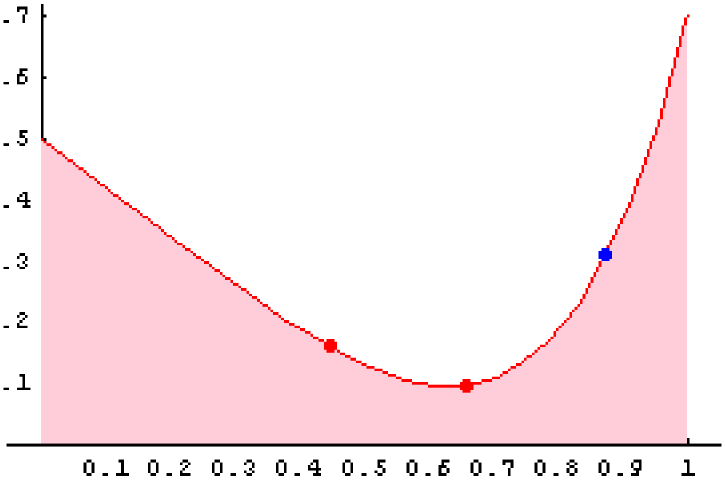
p_0 = 0.00000000, f[0.00000000] = 0.5000000000

p_1 = 0.50000000, f[0.50000000] = 0.1312500000

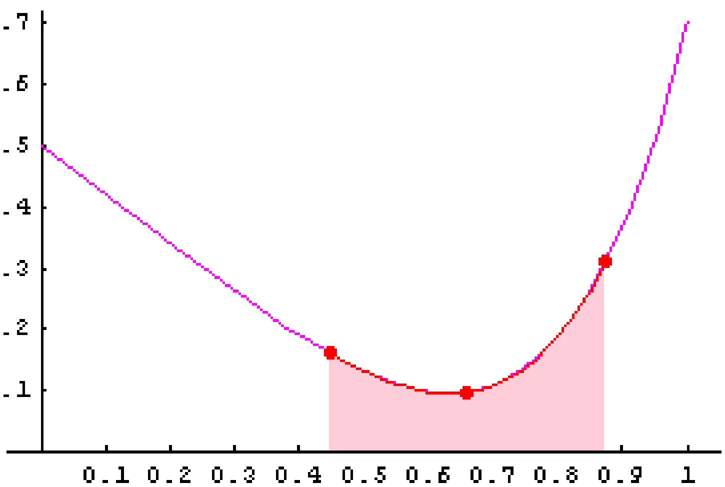
p_2 = 1.00000000, f[1.00000000] = 0.7000000000



$q_{5,0}q_{5,0}$ = 0.66000000, f[0.66000000] = 0.0972332576



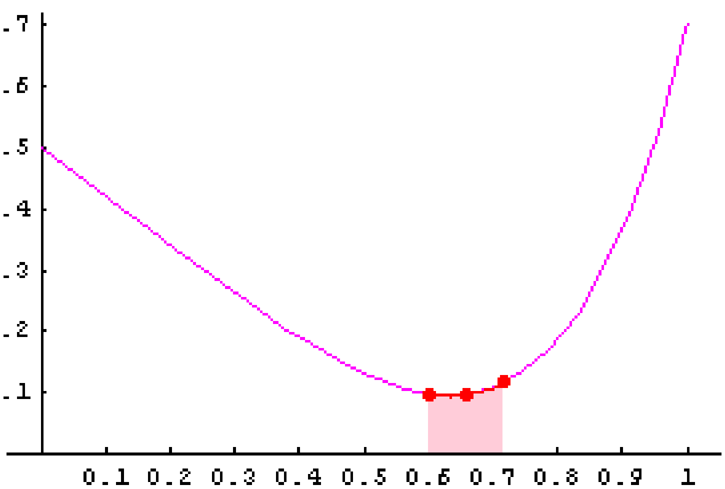
$q_{5,1}q_{5,1}$ = 0.87333333, f[0.87333333] = 0.3093759954



p_3 = 0.44666667, f[0.44666667] = 0.1604460919

p_4 = 0.66000000, f[0.66000000] = 0.0972332576

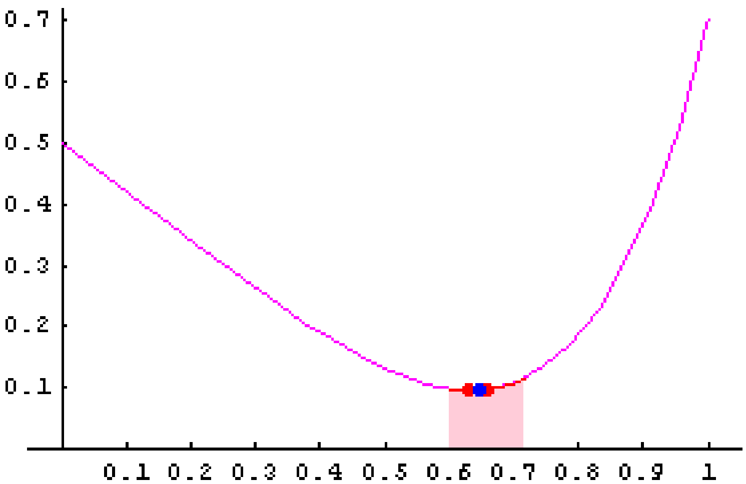
p_5 = 0.87333333, f[0.87333333] = 0.3093759954



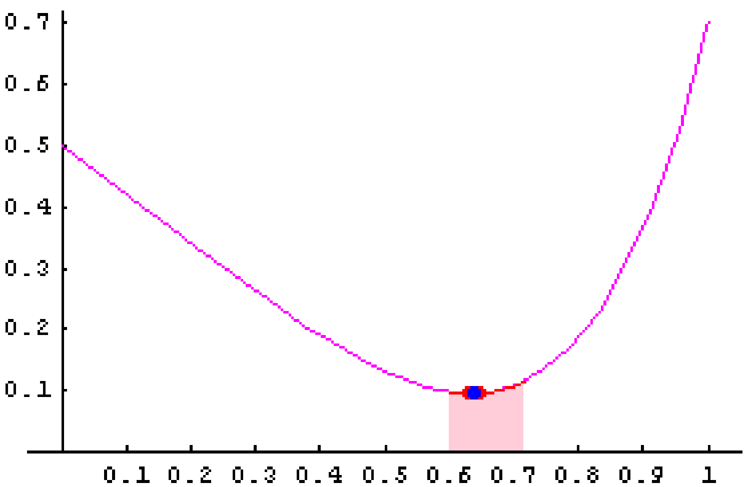
p_6 = 0.60230784, f[0.60230784] = 0.0974207564

p_7 = 0.66000000, f[0.66000000] = 0.0972332576

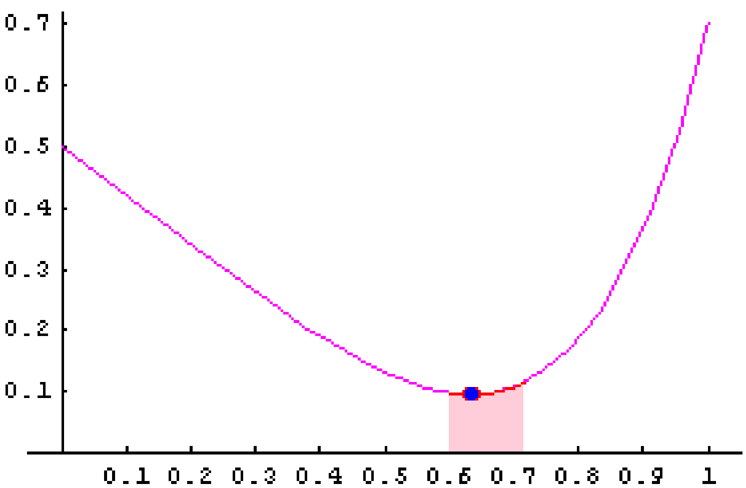
p_8 = 0.71769216, f[0.71769216] = 0.1162568207



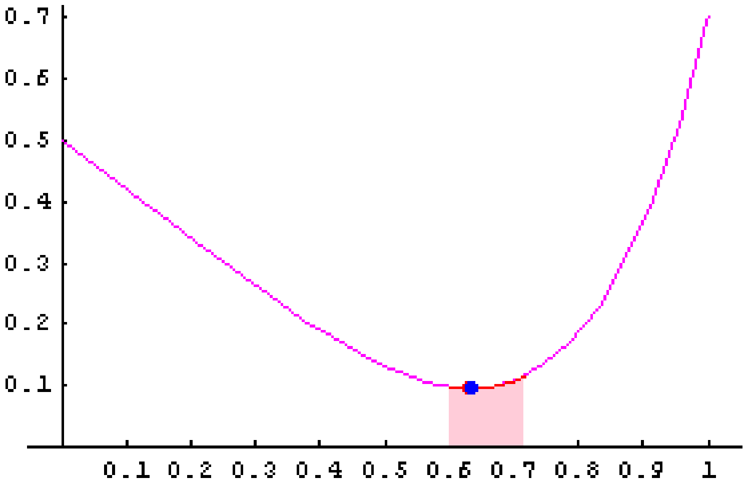
$q_{11,0}$ = 0.64585850, f[0.64585850] = 0.0956926487



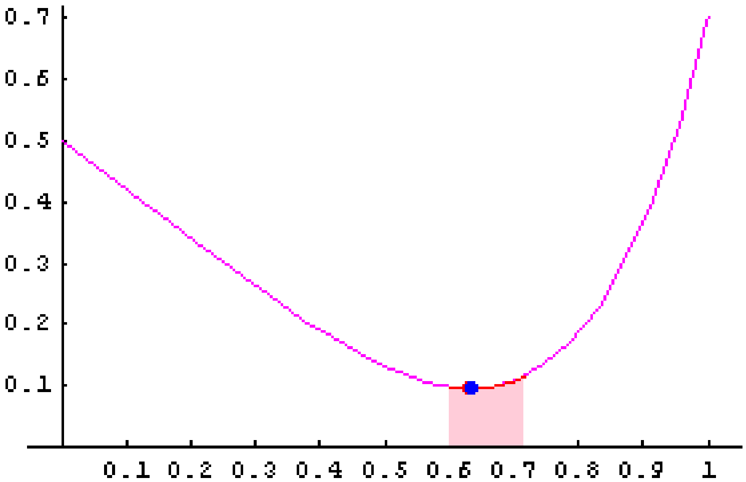
$q_{11,1}$ = 0.63878775, f[0.63878775] = 0.0953309183



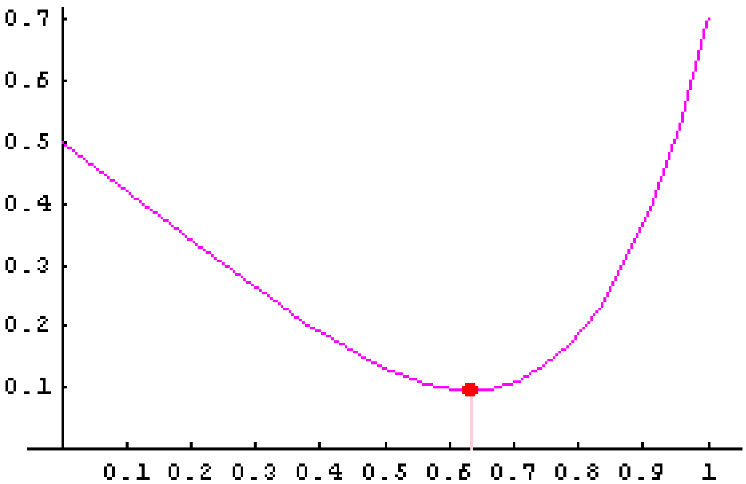
$q_{11,2} = 0.63525237, \quad f[\ 0.63525237\ ] = 0.0952483362$



$q_{11,3} = 0.63348468, \quad f[\ 0.63348468\ ] = 0.0952311433$



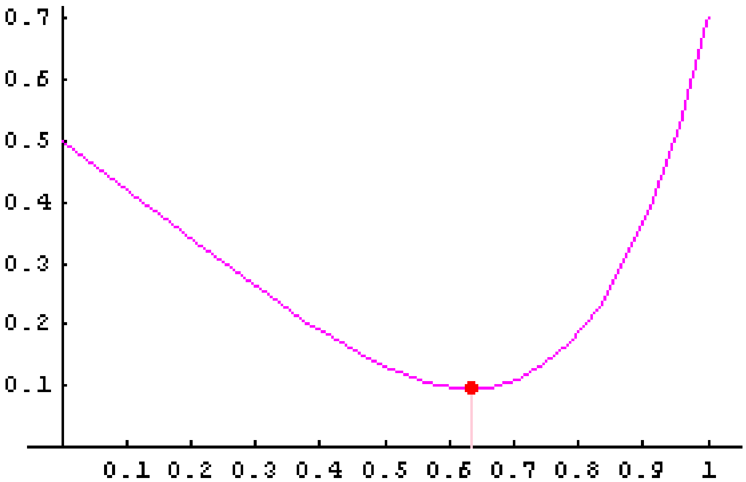
$q_{11,4} = 0.63260084, \quad f[\ 0.63260084\ ] = 0.0952285129$



p_9 = 0.63171699, f[0.63171699] = 0.0952298377

p_{10} = 0.63260084, f[0.63260084] = 0.0952285129

p_{11} = 0.63348468, f[0.63348468] = 0.0952311433



p_{12} = 0.63245496, f[0.63245496] = 0.0952284595

$$f[x] = \frac{1}{2} - \frac{4x}{5} + x^5$$