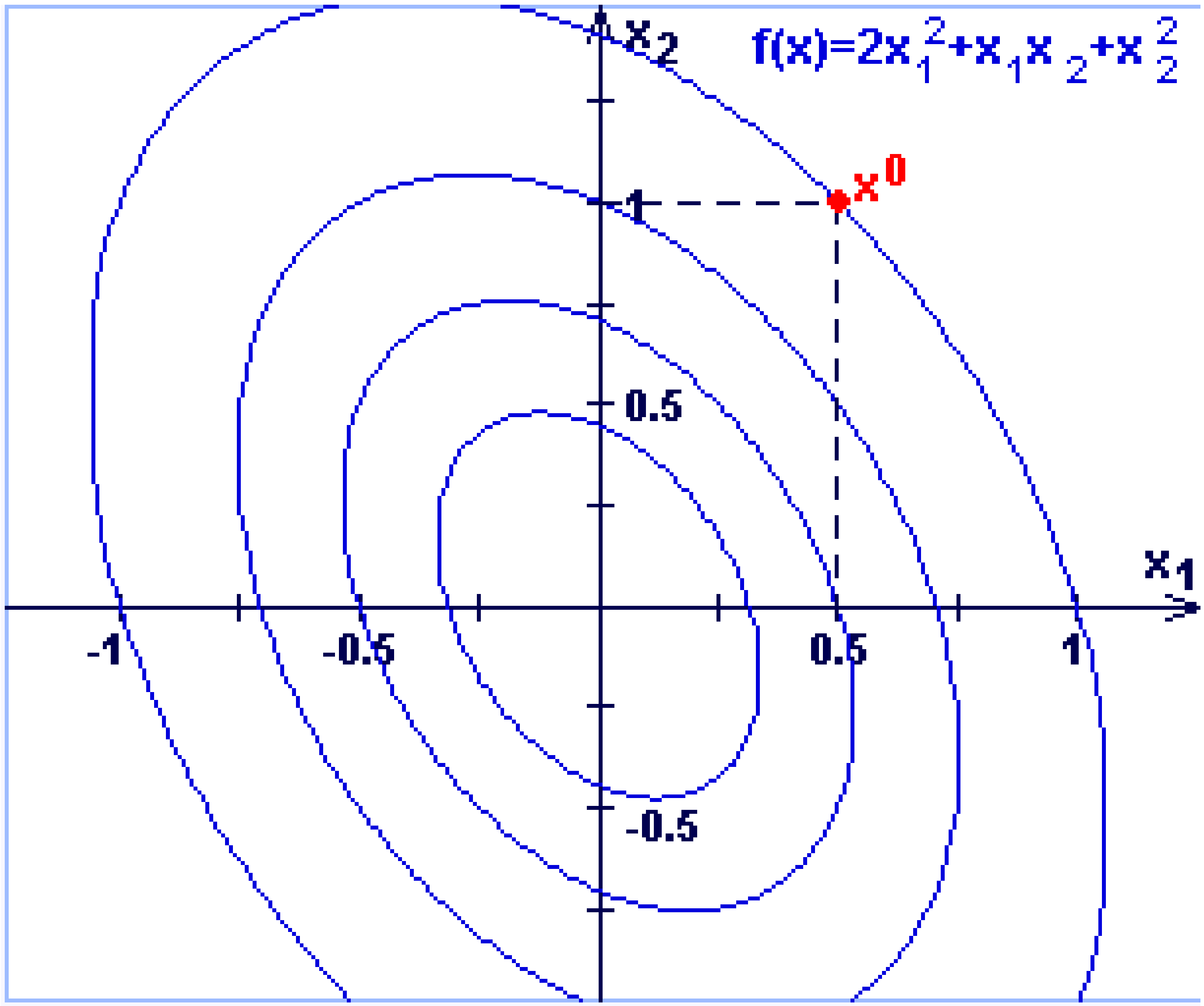
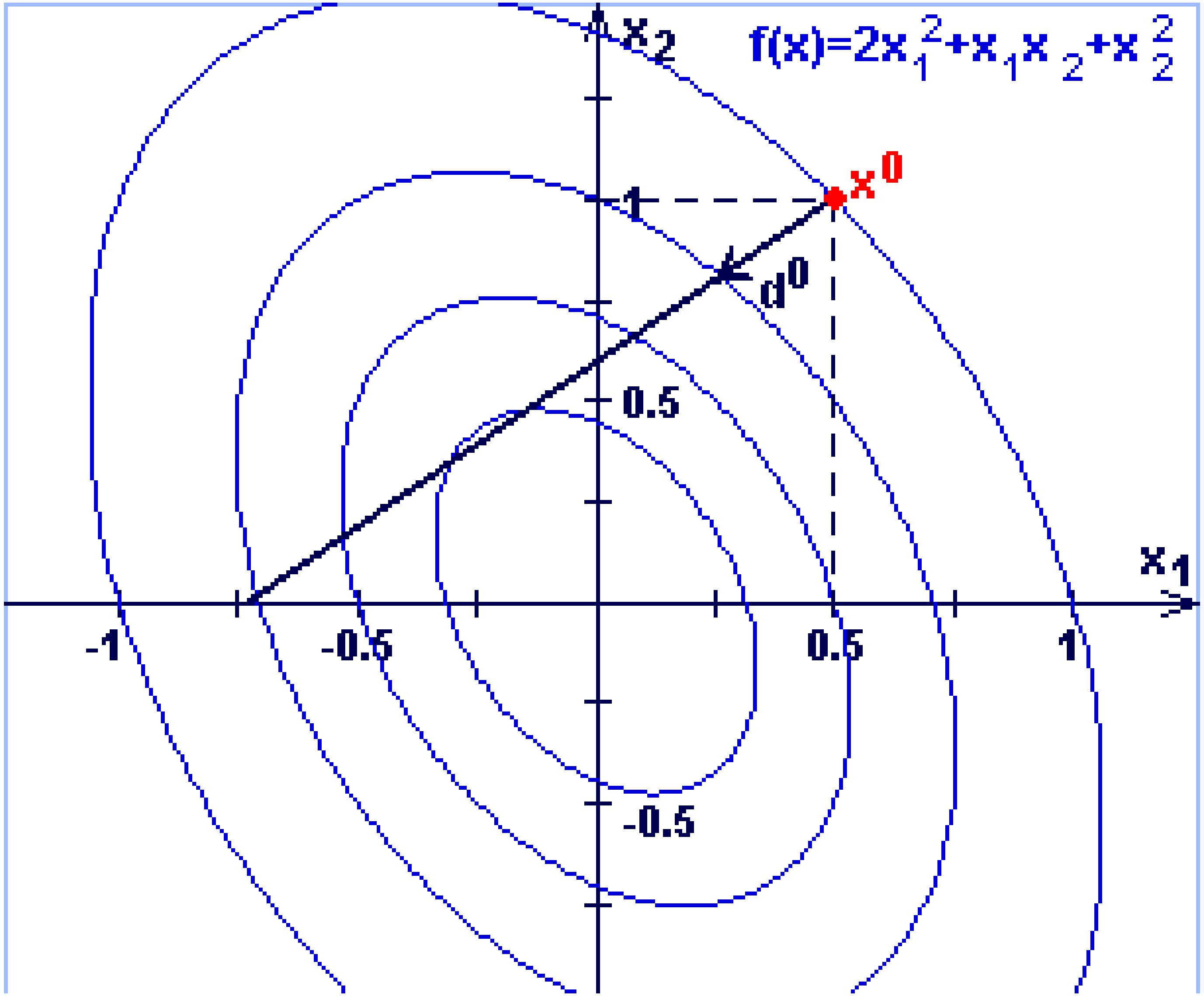
Optimization

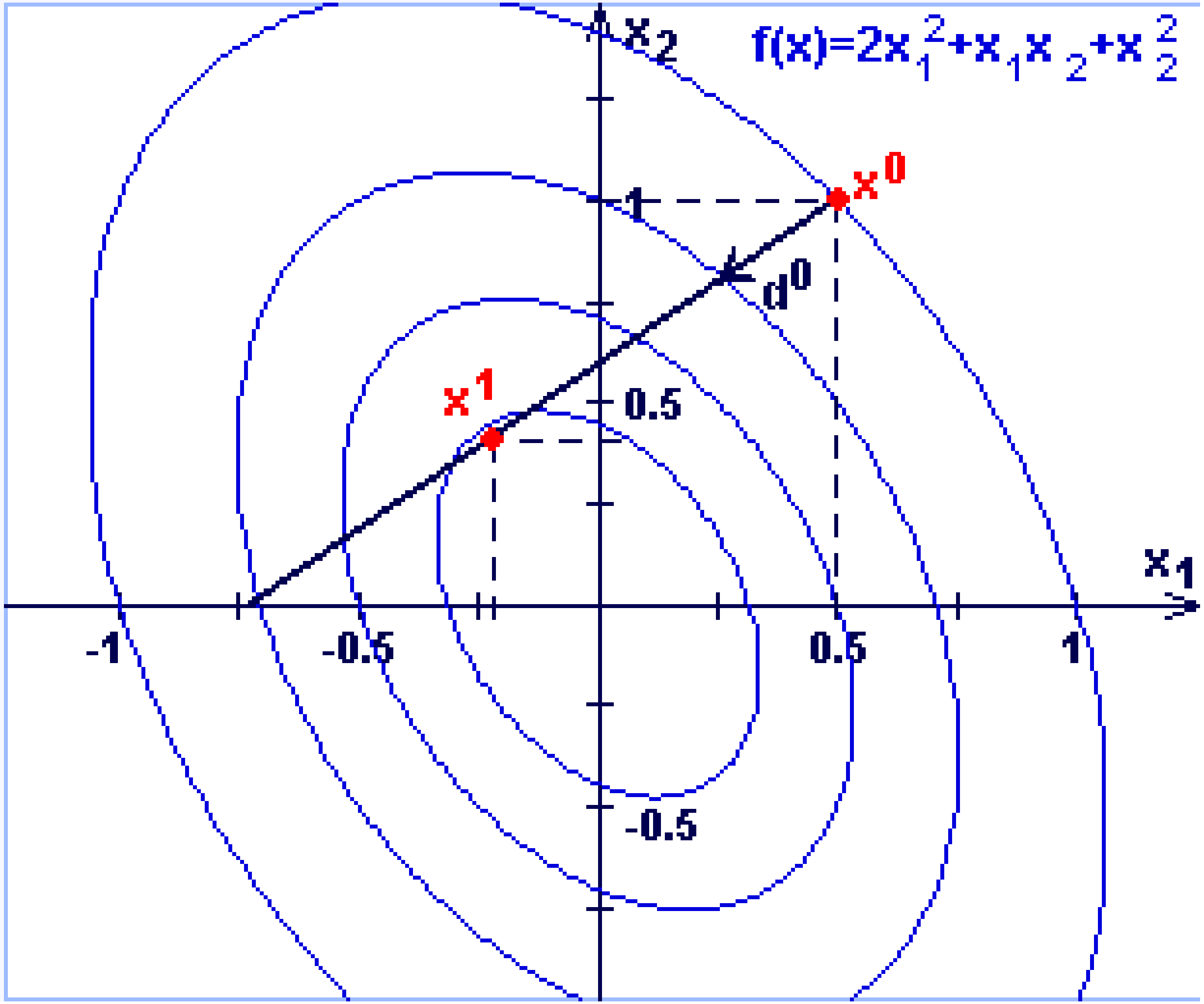
Steepest Descent Example
(on board)



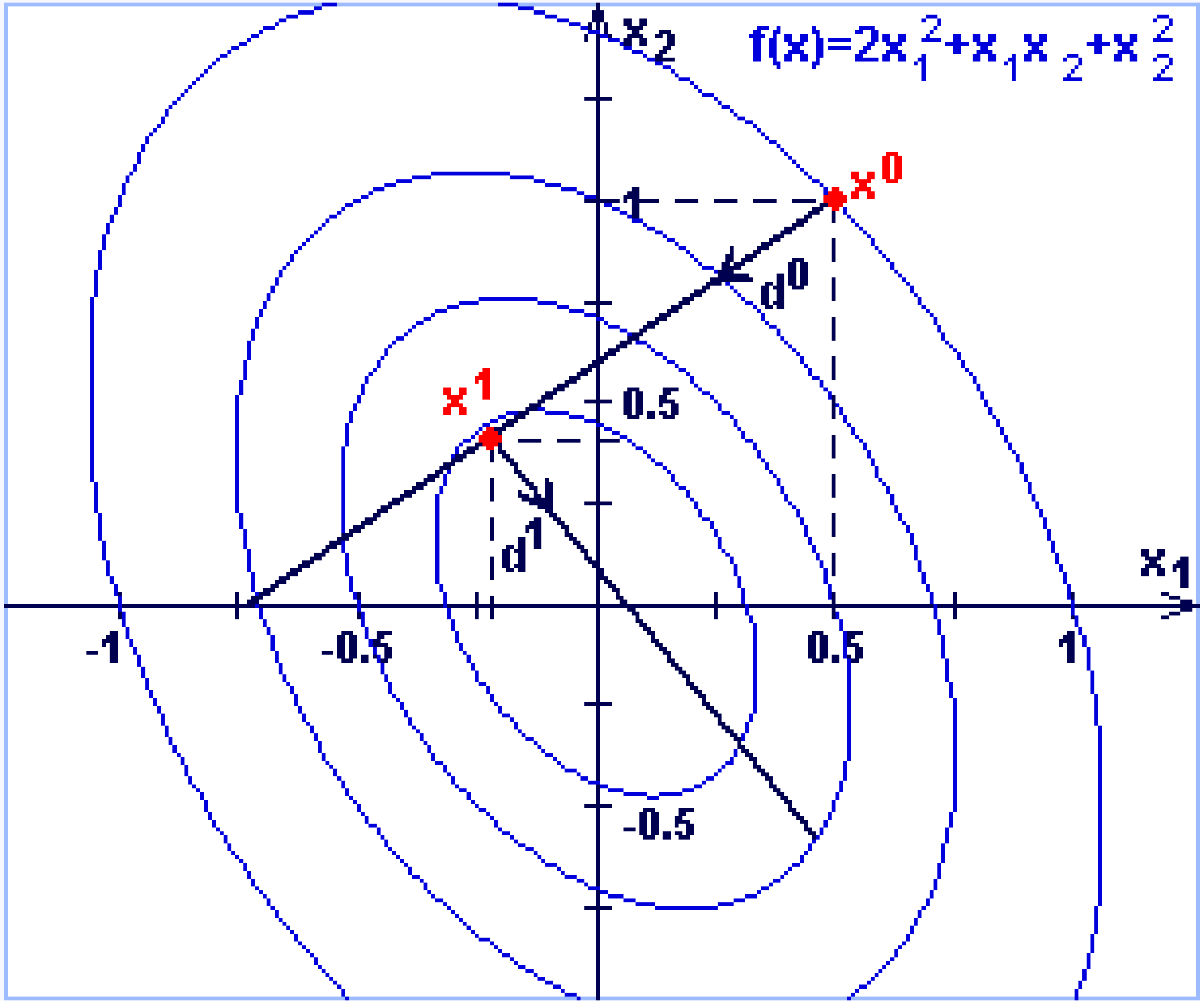
$f(x)=2x_1^2+x_1x_2+x_2^2$
x^0
x_1
x_2
-1
-0.5
0.5
1



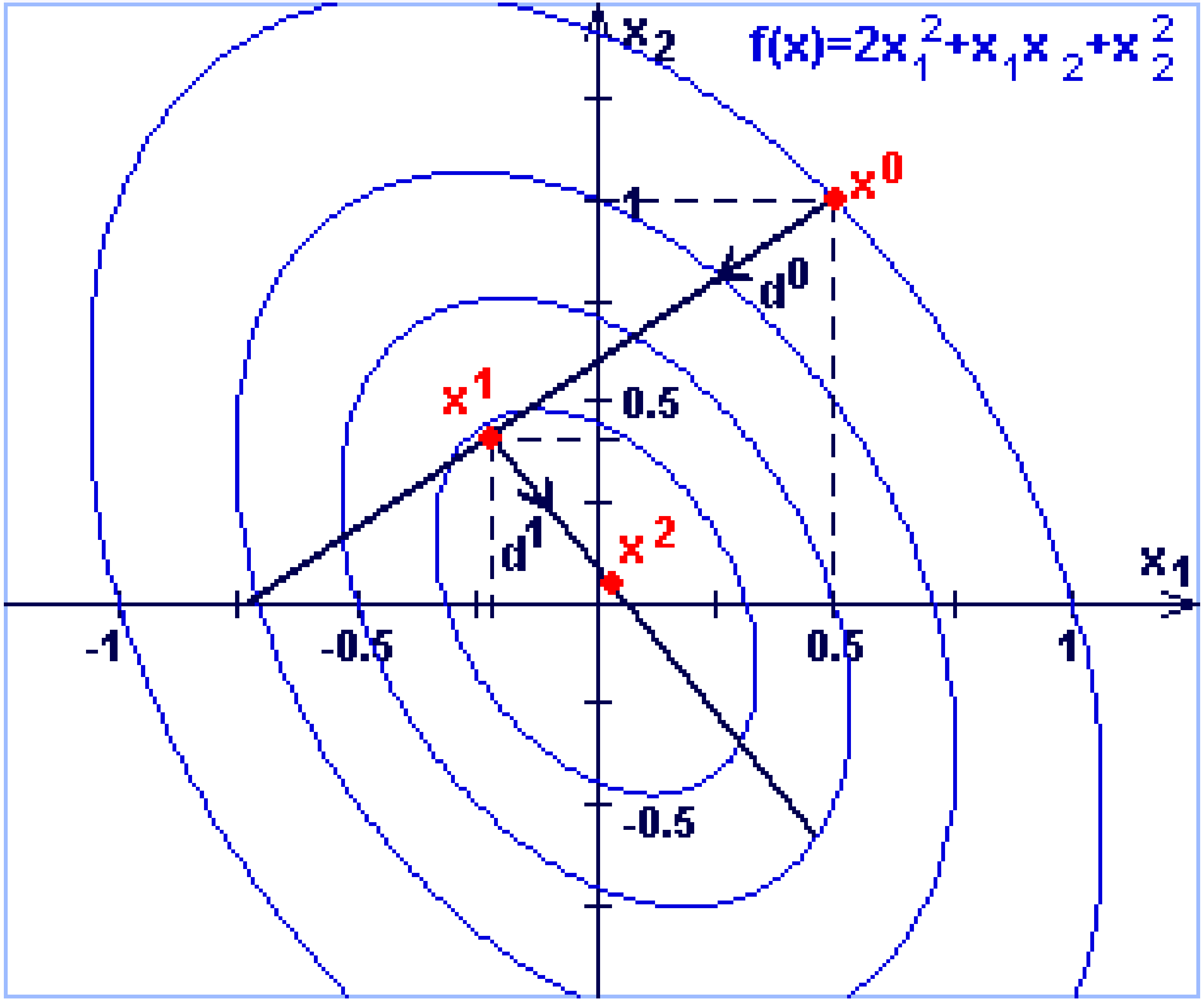
$f(x)=2x_1^2+x_1x_2+x_2^2$
x_2
x_1
x^0
d^0
1
0.5
-0.5
-1
-0.5
0.5
1



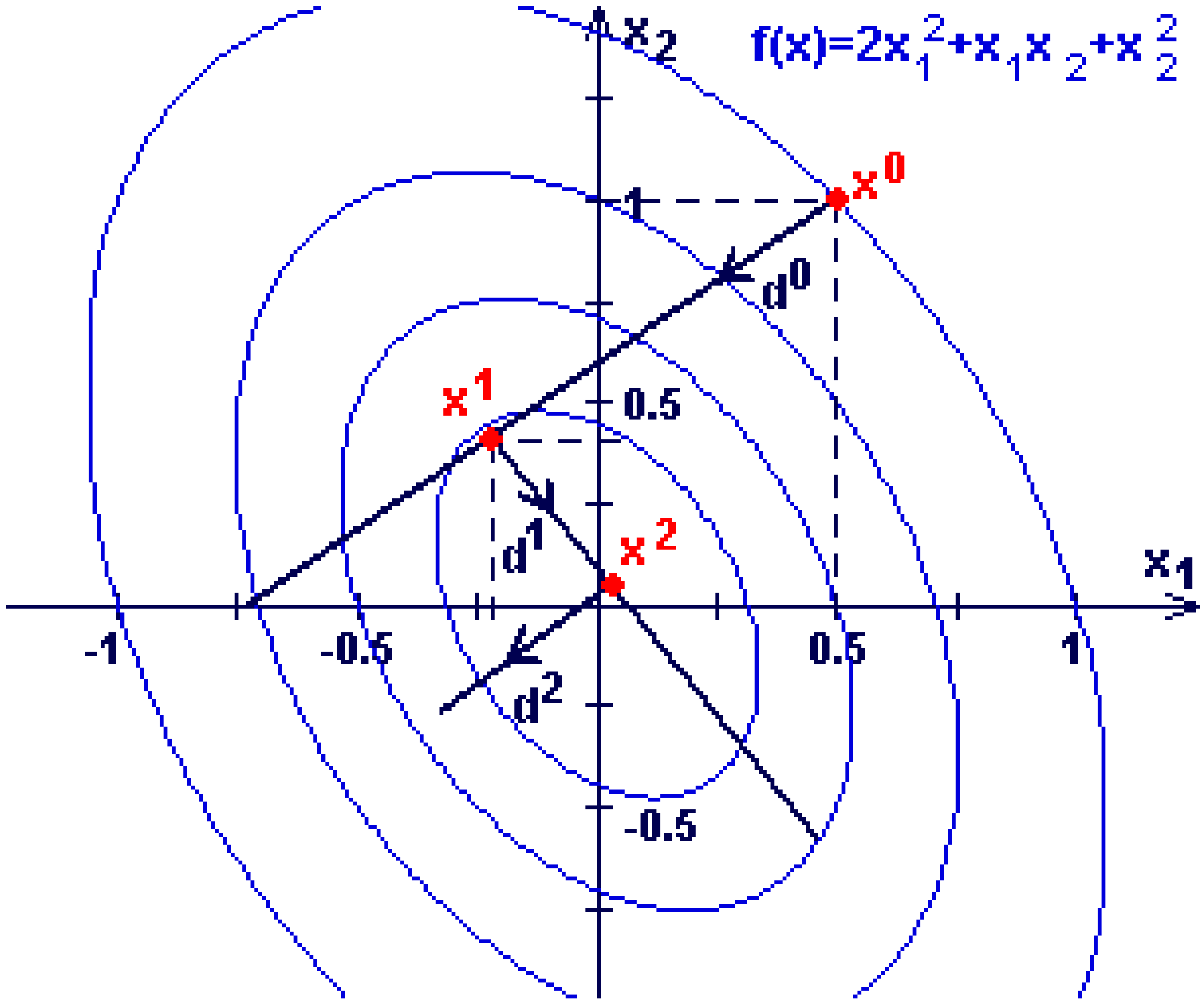
f(x)=2x_1^2+x_1x_2+x_2^2
x2
x1
x0
x1
d0
-1
-0.5
0.5
1
0.5
1
-0.5



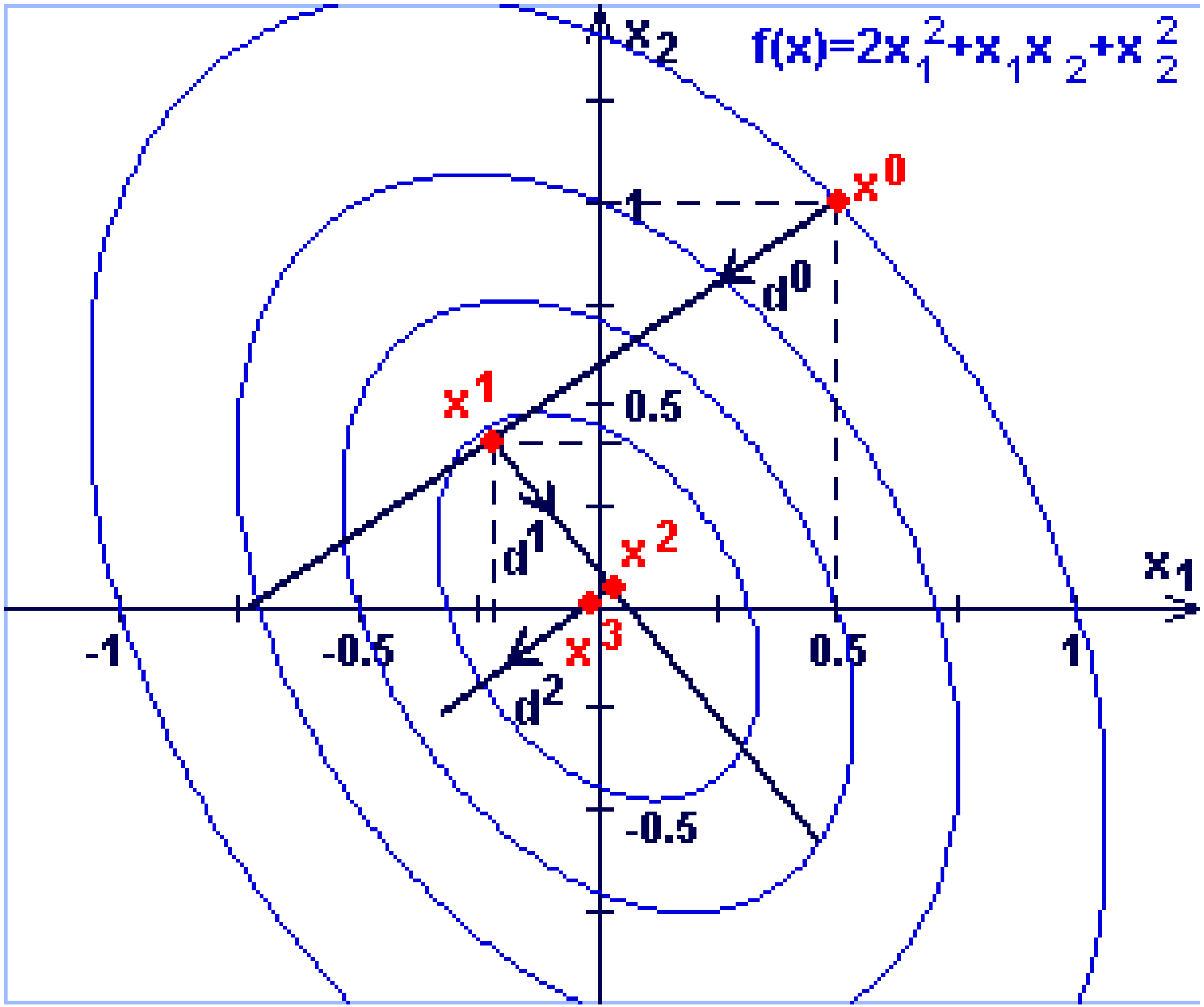
$f(x)=2x_1^2+x_1x_2+x_2^2$
x_2
x_1
x^0
x^1
d^0
d^1
-1
-0.5
0.5
1
-0.5
0.5
1

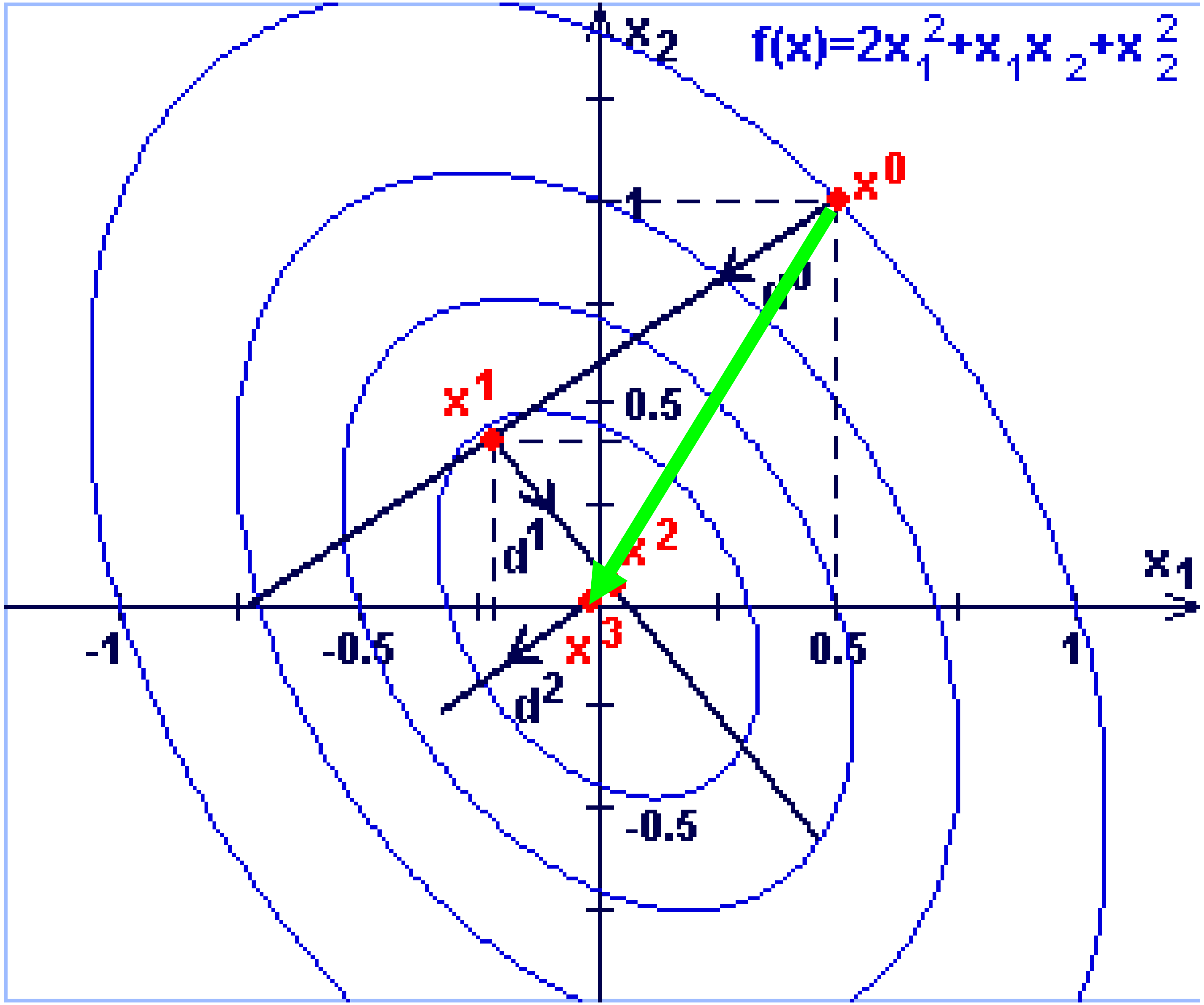


f(x)=2x 2 1+x 1x 2+x 2 2
x2
x1
x0
x1
x2
d0
d1
1
0.5
-0.5
-1



$f(x)=2x_1^2+x_1x_2+x_2^2$
x_1
x_2
x^0
x^1
x^2
d^0
d^1
d^2
-1
-0.5
0.5
1


f(x)=2x
1
2
+x
1
x
2
+x
2
2
x2
x1
x0
x1
x2
x3
d0
d1
d2
-1
-0.5
0.5
1
0.5
-0.5
1



$f(x)=2x_1^2+x_1x_2+x_2^2$
x_1
x_2
x^0
x^1
x^2
x^3
d^0
d^1
d^2
-1
-0.5
0.5
1

Newton Method

Step 1: Set an initial point x^0, $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, M- maximum number of iterations. Find the gradient $\nabla f(x)$ and Hessian matrix $H(x)$.

Step 2: Set $k=0$.

Step 3: Compute $\nabla f(x^k)$.

Step 4: Check up the the stopping criteria:
a) if $||\nabla f(x^k)|| < \varepsilon_1$, set $x^* = x^k$ and finish the search process;
b) if $||\nabla f(x^k)|| > \varepsilon_1$, go to step 5.

Step 5: Check the condition $k \geq M$:
a) if it is satisfied, finish the search process and set $x^* = x^k$;
b) if it is not satisfied, go to step 6.

Step 6: Compute the Hessian matrix $H(x^k)$.

Newton Method

Step 7: Compute $(H^{-1})(x^k)$.

Step 8: Check the following condition $(H^{-1})(x^k)>0$:
a) if it is satisfied, go to step 9;
b) if it is not satisfied, set $d^k=-\nabla f(x^k)$ and go to step 10;

Step 9: Compute $d^k=(-(H^{-1})(x^k))\cdot\nabla f(x^k)$.

Step 10: Update $x^{k+1}=x^k+t_k\cdot d^k$.
- if $d^k=(-(H^{-1})(x^k))\cdot\nabla f(x^k)$, set $t_k=1$,
- if $d^k=-\nabla f(x^k)$, choose t_k using the following condition $f(x^{k+1})<f(x^k)$.

Step 11: Check up the finishing conditions:
$||x^{k+1}-x^k|| < \varepsilon_2$
$||f(x^{k+1})-f(x^k)|| < \varepsilon_2$
a) if both conditions are satisfied with numbers k and k-1, finish the search process and set $x^*=x^{k+1}$;
b) if both conditions are not satisfied, set $k=k+1$ and go to step 3.

Newton Method Example
(on board)

Convergence

We have tried to find some x^k that converge to the desired point x^* (most often a local minimizer). The fundamental question is how fast the convergence is

Let's define errors at each iteration step as $\| e_k \| = \| x_k - x^* \|$ and $\| e_{k+1} \| = \| x_{k+1} - x^\star \|$

Linear Convergance : There exist a constant $0 < c_1 < 1$ where $\| e_{k+1} \| \leq c_1 \| e_k \|$

Quadratic Convergance : There exist a constant $c_2 > 0$where $\| e_{k+1} \| \leq c_2 \| e_k \|^2$

Superlinear Convergance : Where $\frac{\|e_{k+1}\|}{\|e_k\|} \to 0$

Convergence Example

Two methods are given; one of them showing linear convergence and the other one showing quadratical convergence.
After certain number of steps errors reach 3 digit precision for both of the methods ($\| e_k \| < 0{,}001$).

How many steps (iterations) will be necessary if we require 12 digits of precision ?

$$c_1 = c_2 = \frac{1}{2}$$

In case of Linear Convergance : $\| e_{k+1} \| \leq \frac{1}{2} \| e_k \|$

$$\| e_{k+2} \| \leq \left(\frac{1}{2}\right)^2 \| e_k \| \| e_{k+1} \|$$

....

$\left(\frac{1}{2}\right)^{30} \approx 10^{-9}$ we already had 3 digits of precision so approximately after 30 steps we can reach 12 digits of precision.

Convergence Example

In case of Quadratic Convergance :

$$\| e_{k+1} \| \leq \frac{1}{2} \| e_k \|^2$$

$$\| e_{k+2} \| \leq \frac{1}{2}\left(\frac{1}{2} \| e_k \|^2\right)^2$$

$$\| e_{k+3} \| \leq \frac{1}{2}\left(\frac{1}{2}\left(\frac{1}{2} \| e_k \|^2\right)^2\right)^2$$

Up to 3rd iteration the method converges....

How to Approximate Derivatives of Univariate Functions ?

Taylor Series Expansion

Taylor's Theorem : Suppose *f* is continuous on the closed interval [a, b] and has *n + 1* continuous derivatives on the open interval (a, b). If *x* and *c* are points in (a, b), then

$$\text{The Taylor series expansion of } f(x) \text{ about } c:$$

$$f(c) + f^{'}(c)(x-c) + \frac{f^{(2)}(c)}{2!}(x-c)^2 + \frac{f^{(3)}(c)}{3!}(x-c)^3 + \dots$$

$$or$$

$$Taylor\ \ Series = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(c)\ (x-c)^k$$

$$\text{If the series converge, we can write:}$$

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(c)\ (x-c)^k$$

Maclaurin Series

Maclaurin series is a special case of Taylor series with the center of expansion $c = 0$.

The Maclaurin series expansion of $f(x)$:

$$f(0) + f^{'}(0)\,x + \frac{f^{(2)}(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \ldots$$

If the series converge, we can write:

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(0)\; x^k$$

What If we change the interval ?

If we change the interval so $x=x+h$ and $c=x$

The Taylor series expansion of $f(x+h)$ about $c=x$

$$f(x+h)=f(x)+f^{'}(x)h+\frac{f^{(2)}(x)}{2!}h^{2}+\frac{f^{(3)}(x)}{3!}h^{3}+\ldots$$

Finite Differences for Derivation

Taylor expansion for $F(x+h)$

$$F(x+h) = F(x) + hF'(x) + \frac{h^2}{2!}F''(x) + ...$$

$$F(x+h) - F(x) = hF'(x) + O(h^2)$$

Use only first two terms of the expansion

$$\frac{F(x+h) - F(x)}{h} = F'(x) + \frac{O(h^2)}{h}$$

$$\boxed{F'(x) = \frac{F(x+h) - F(x)}{h}}$$

Forward Difference Formula

Finite Differences for Derivation

Taylor expansion for $F(x-h)$

$$F(x-h) = F(x) - hF'(x) + \frac{h^2}{2!}F''(x) + ...$$

$$F(x) - F(x-h) = hF'(x) + O(h^2)$$

$$\boxed{F'(x) = \frac{F(x) - F(x-h)}{h}}$$

Backward Difference Formula

Finite Differences for Derivation

$$F(x-h) = F(x) - hF'(x) + \frac{h^2}{2!}F''(x) + \frac{h^3}{3!}F'''(x) + ...$$

$$F(x+h) - F(x-h) = 2hF'(x) + 2\frac{h^3}{3!}F'''(x) + ...$$

$$\frac{F(x+h) - F(x-h)}{2h} = F'(x) + O(h^2)$$

$$\boxed{F'(x) = \frac{F(x+h) - F(x-h)}{2h}}$$

Central Difference Formula

Better approximations

By using Taylor expansions for $F(x+2h)$ and $F(x-2h)$, better approximations can be obtained:

$$F'(x) = \frac{-F(x+2h)+4F(x+h)-3F(x)}{2h} + O(h^2)$$

Forward Difference Formula

$$F'(x) = \frac{3F(x)-4F(x-h)+F(x-2h)}{2h} + O(h^2)$$

Backward Difference Formula

$$F'(x) = \frac{-F(x+2h)+8F(x+h)-8F(x-h)+F(x-2h)}{12h} + O(h^4)$$

Central Difference Formula

Example : Derivation of Forward Difference Formula

$$F(x+2h) = F(x) + 2hF'(x) + \frac{4h^2}{2!}F''(x) + ...$$

$$*\,(-4)\qquad F(x+h) = F(x) + hF'(x) + \frac{h^2}{2!}F''(x) + ...$$

$$+$$

$$F(x+2h) - 4F(x-h) = F(x) - 4F(x) + 2hF'(x) - 4hF'(x) + \frac{4h^2}{2!}F''(x) - .\frac{4h^2}{2!}F''(x)$$

$$F(x+2h) - 4F(x-h) = -3F(x) + 2hF'(x)$$

$$\boxed{F'(x) = \frac{F(x+2h) - 4F(x-h) + 3F(x)}{2h}}$$

Second Derivatives

Similarly various approximations for second derivatives are possible :

$$F''(x) = \frac{F(x+2h) - 2F(x+h) + F(x)}{h^2} + O(h)$$

$$F''(x) = \frac{F(x+h) - 2F(x) + F(x-h)}{h^2} + O(h^2)$$

$$F''(x) = \frac{-F(x+2h) + 16F(x+h) - 30F(x) + 16F(x-h) - F(x-2h)}{12h^2} + O(h^4)$$

Example : Derivation of Forward Difference Formula

$$F(x+h) = F(x) + hF'(x) + \frac{h^2}{2!}F''(x) + ...$$

$$F(x-h) = F(x) - hF'(x) + \frac{h^2}{2!}F''(x) + ...$$

$+$

$$F(x+h) + F(x-h) = 2F(x) + hF'(x) - hF'(x) + \frac{h^2}{2!}F''(x) + \frac{h^2}{2!}F''(x)$$

$$F(x+h) + F(x-h) = 2F(x) + \cancel{2}\frac{h^2}{\cancel{2}!}F''(x)$$

$$\boxed{F''(x) = \frac{F(x+h) - 2F(x) + F(x-h)}{h^2} + O(h^2)}$$

Derivatives of Bivariate & Multivariate Functions

First Order Partial Derivatives

For functions with more variables, the partial derivatives can be approximated by grouping together all of the same variables and applying the univariate approximation for that group.
For example, if *F(x, y)* is our function, then some first order partial derivative approximations are:

$$f_x(x, y) = \frac{F(x+h, y) - F(x-h, y)}{2h}$$

$$f_y(x, y) = \frac{F(x, y+k) - F(x, y-k)}{2k}$$

Second Partial Derivatives

Following formulas can be verified by taking the limits $h \to 0$ and $k \to 0$

$$f_{xx}(x,y) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{F(x+h,y) - 2f(x,y) + F(x-h,y)}{h^2}$$

$$f_{yy}(x,y) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{F(x,y+k) - 2f(x,y) + F(x,y-k)}{k^2}$$

$$\boxed{f_{xy}(x,y) = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{F(x+h,y+k) - F(x+h,y-k) - F(x-h,y+k) + F(x+h,y-k)}{4hk}}$$

The derivatives F_x, F_y, F_{xx} and F_{yy} just use the univariate approximation formulas.

The mixed derivative requires slightly more work. The important observation is that the approximation for F_{xy} is obtained by applying the x–derivative approximation for F_x, then applying the y–derivative approximation to the previous approximation.

1 - the x–derivative approximation for F_x :

$$f_x(x, y) = \frac{F(x+h, y) - F(x-h, y)}{2h}$$

2- the y–derivative approximation:

$$f_y(x, y) = \frac{F(x, y+k) - F(x, y-k)}{2k}$$

3- Apply the 2 nd formula to the 1st one:

$$f_{xy}(x, y) = \frac{F(x+h, y+k) - F(x+h, y-k) - F(x-h, y+k) + F(x+h, y-k)}{4hk}$$

Second Partial Derivatives

Following formulas can be verified by taking the limits $h \to 0$ and $k \to 0$

To find the F_{xy} partial derivative first use the formula for F_x

Then apply the approximation for (f_y)

$$f_{xx}(x, y) = \frac{F(x+h, y) - 2f(x, y) + F(x-h, y)}{h^2}$$

$$f_{yy}(x, y) = \frac{F(x, y+k) - 2f(x, y) + F(x, y-k)}{k^2}$$

$$f_{xy}(x, y) = \frac{F(x+h, y+k) - F(x+h, y-k) - F(x-h, y+k) + F(x+h, y-k)}{4hk}$$