Newton's method

Dominik Koszkul
Michał Oleszczyk

May 23, 2015

1 Introduce

1.1 Formulating optimization problem

Problem, which needs to be solved, is to find minimum point in set non-linear, multidimensional function. If function has more than one minimum, algorithm is looking for the nearest local minimum from initial point. In this project method to find these special points is Newton's method. This algorithm allows to find local minimum points, but this particular method can be optimized by using other algorithm to find optimal step, which is used in the main optimization program. Thanks to this, whole algorithm can find the solution in a faster way.

1.2 Newton's method

Newton's method is a local optimization type algorithm. To implement the method, stop criteria should be also known. In this case we have three main stop criteria for Newton's method and one additional that ensures that program stops counting after exceeding maximum compute time.

a. $< gradf(x) \cdot gradf(x) > \leqslant \epsilon_1$

b. $||x_i - x_{i-1}|| \leqslant \epsilon_2$

c. $|f(x_i) - f(x_{i-1})| \leqslant \epsilon_3$

d. max number of iterations

If one from these four criteria

Newton's method in optimization is a numeric method, which is used to find local extrema in a defined, differentiable function f. In this method we need to construct a sequence x_n from initial point x_0 to x_* such that $f'(x_*) = 0$. Last point is a local extremum point which we are looking for.

1.2.1 The Newton's method iteration

Let x_0 be a point

2 Examples

2.1 Function with four local minima

Function equation:

$$y = x_1^4 + x_2^4 - 0.62x_1^2 - 0.62x_2^2 \tag{1}$$

Function figures:

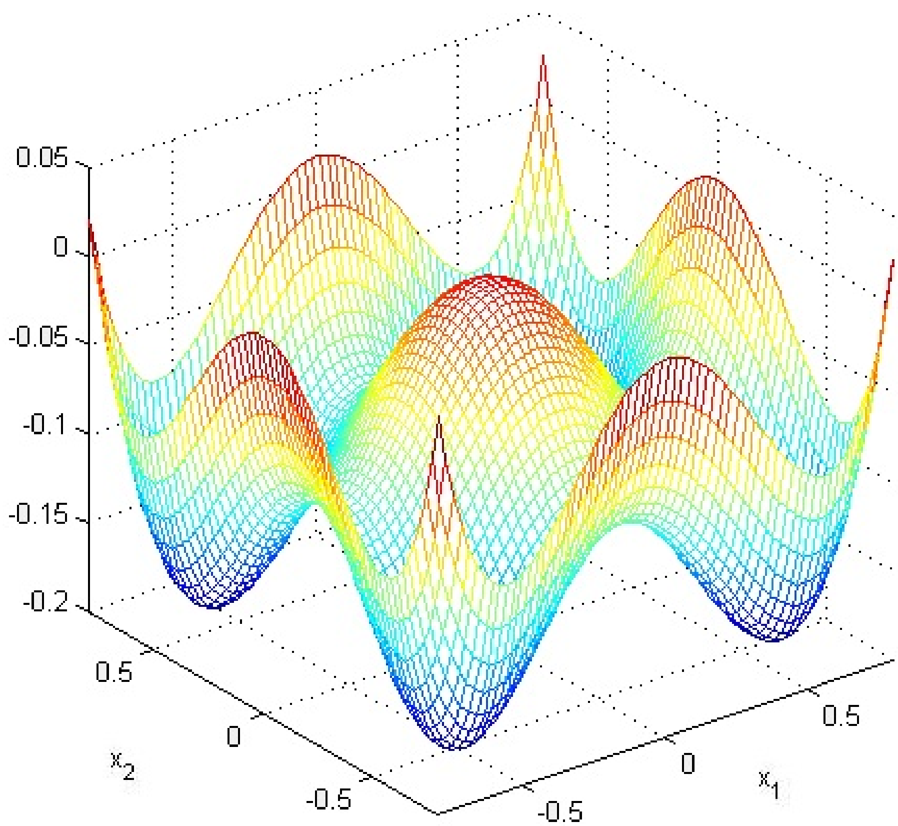

Figure 1: Analyzed function 3D view.

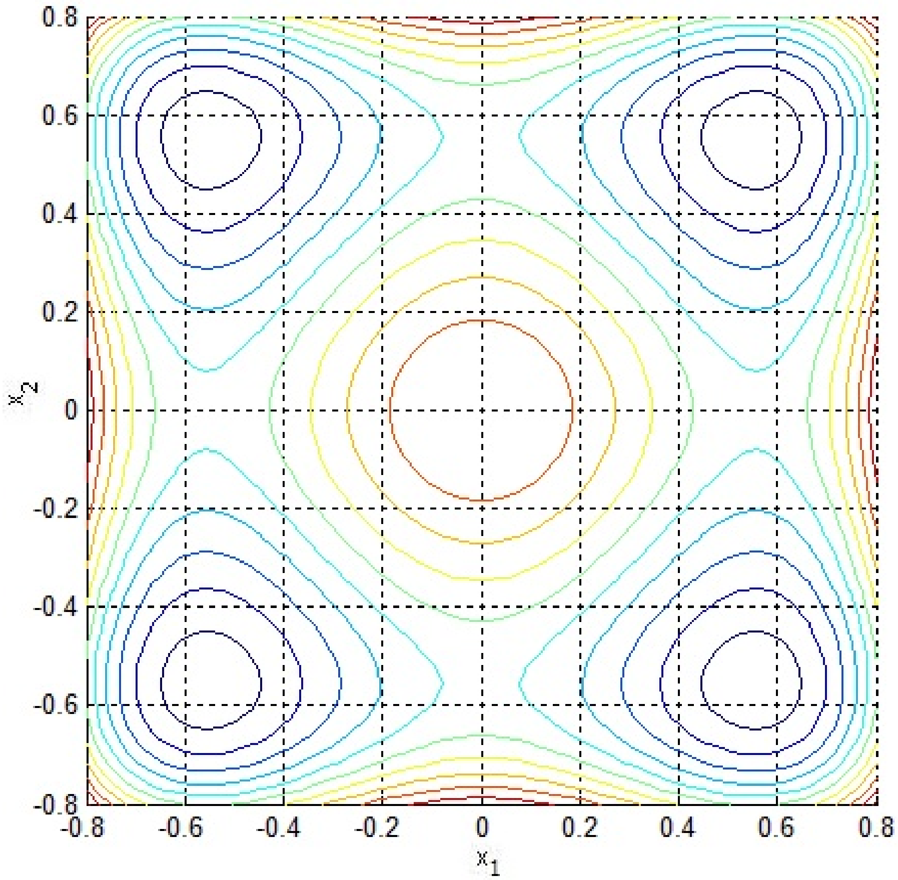

Figure 2: Analyzed function contour view.

Results:

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$1, 1$	0.76	0.132	-	-
1	$0.743, 0.743$	-0.074	0.132	0.363	0.834
2	$0.61, 0.61$	-0.185	0.0358	0.189	0.11
3	$5.63 \cdot 10^{-1}, 5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$4.36 \cdot 10^{-3}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$
4	$5.63 \cdot 10^{-1}, 5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$7.35 \cdot 10^{-5}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-1, 1$	0.76	0.132	-	-
1	$-0.743, 0.743$	-0.074	0.132	0.363	0.834
2	$-0.61, 0.61$	-0.185	0.0358	0.189	0.11
3	$-5.63 \cdot 10^{-1}, 5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$4.36 \cdot 10^{-3}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$
4	$-5.63 \cdot 10^{-1}, 5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$7.35 \cdot 10^{-5}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-1, -1$	0.76	0.132	-	-
1	$-0.743, -0.743$	-0.074	0.132	0.363	0.834
2	$-0.61, -0.61$	-0.185	0.0358	0.189	0.11
3	$-5.63 \cdot 10^{-1}, -5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$4.36 \cdot 10^{-3}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$
4	$-5.63 \cdot 10^{-1}, -5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$7.35 \cdot 10^{-5}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$1, -1$	0.76	0.132	-	-
1	$0.743, -0.743$	-0.074	0.132	0.363	0.834
2	$0.61, -0.61$	-0.185	0.0358	0.189	0.11
3	$5.63 \cdot 10^{-1}, -5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$4.36 \cdot 10^{-3}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$
4	$5.63 \cdot 10^{-1}, -5.63 \cdot 10^{-1}$	$-1.92 \cdot 10^{-1}$	$7.35 \cdot 10^{-5}$	$6.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$

When we chose starting point at $x_1 = 0, x_2 = 0$ then program returned: 'local maximum. Computation stopped'. This is caused by checking Silvester's criterion. Even if algorithm find local extrema it can distinguish maxima or minima.

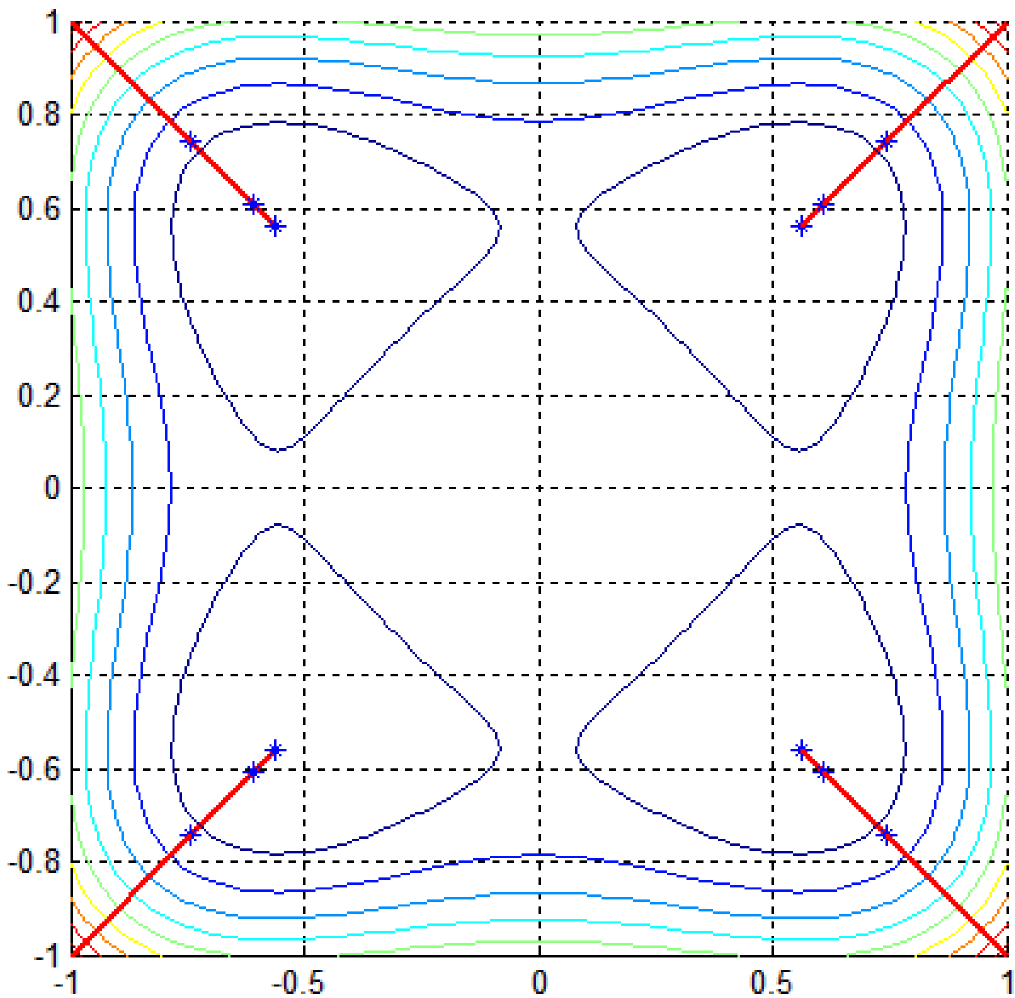

Figure 3: Location of four minima.

2.2 Himmelblau modification function

Function equation:

$$y = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 \tag{2}$$

Function figures:

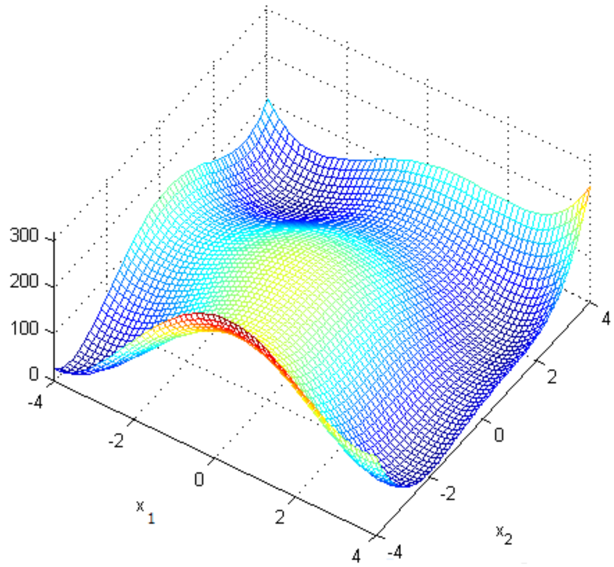

Figure 4: Analyzed function 3D view.

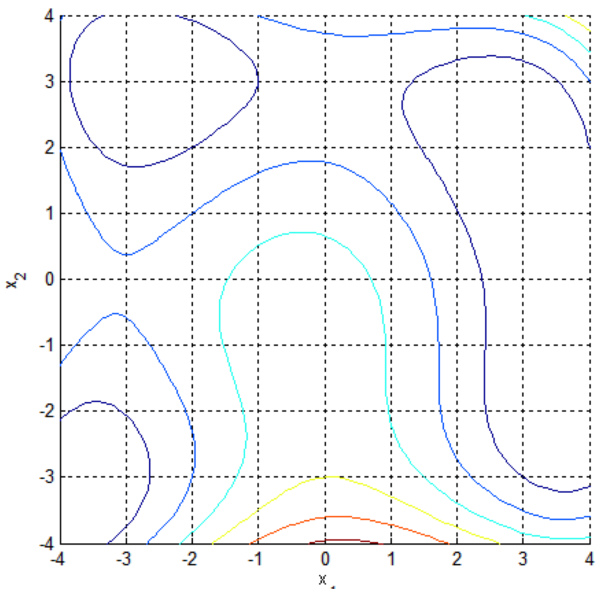

Figure 5: Analyzed function contour view.

Results:

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$4, 4$	250	1.867	-	-
1	$3.1, 2.901$	25.451	1.867	1.367	224.55
2	$2.957, 2.309$	1.666	0.403	0.635	23.785
3	$2.987, 2.055$	$4.53 \cdot 10^{-2}$	$6.52 \cdot 10^{-2}$	$2.55 \cdot 10^{-1}$	1.62
4	$3, 2$	$7.29 \cdot 10^{-5}$	$3 \cdot 10^{-3}$	$5.47 \cdot 10^{-2}$	$4.53 \cdot 10^{-2}$
5	$3, 2$	$7.29 \cdot 10^{-5}$	$5.28 \cdot 10^{-6}$	$5.47 \cdot 10^{-2}$	$4.53 \cdot 10^{-2}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$4, -4$	170	1.379	-	-
1	$3.687, -2.868$	24.212	1.379	1.174	145.788
2	$3.615, -2.215$	2.331	0.432	0.657	21.881
3	$3.591, -1.921$	$7.99 \cdot 10^{-2}$	$8.68 \cdot 10^{-2}$	0.295	2.251
4	$3.585, -1.852$	$2.13 \cdot 10^{-4}$	$4.81 \cdot 10^{-3}$	$6.94 \cdot 10^{-2}$	$7.97 \cdot 10^{-2}$
5	$3.585, -1.852$	$2.13 \cdot 10^{-4}$	$1.47 \cdot 10^{-5}$	$6.94 \cdot 10^{-2}$	$7.97 \cdot 10^{-2}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-4, -4$	26	3.43	-	-
1	$-3.822, -3.442$	1.079	0.343	0.586	24.921
2	$-3.782, -3.294$	$4.61 \cdot 10^{-3}$	$2.36 \cdot 10^{-2}$	0.154	1.074
3	$-3.779, -3.283$	$1.13 \cdot 10^{-7}$	$1.19 \cdot 10^{-4}$	$1.09 \cdot 10^{-2}$	$4.61 \cdot 10^{-3}$
4	$-3.779, -3.283$	$1.13 \cdot 10^{-7}$	$2.93 \cdot 10^{-9}$	$1.09 \cdot 10^{-2}$	$4.61 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-4, 4$	106	1.078	-	-
1	$-3.193, 3.347$	7.469	1.078	1.038	98.531
2	$-2.867, 3.15$	0.139	0.145	0.381	7.329
3	$-2.807, 3.132$	$1.23 \cdot 10^{-4}$	$3.92 \cdot 10^{-3}$	$6.26 \cdot 10^{-2}$	0.139
4	$-2.807, 3.132$	$1.23 \cdot 10^{-4}$	$3.8 \cdot 10^{-6}$	$6.26 \cdot 10^{-2}$	0.139

When we chose starting point at $x_1 = -3, x_2 = 0$ or $x_1 = 0.1, x_2 = 2.8$ then program returned: 'local minimum or maximum does not exist'. This is caused by saddle points there. Situation can be seen on figure 6. And at point $x_1 = 0, x_2 = 0$ we found local maxima. Program returned: 'local maximum. Computation stopped'.

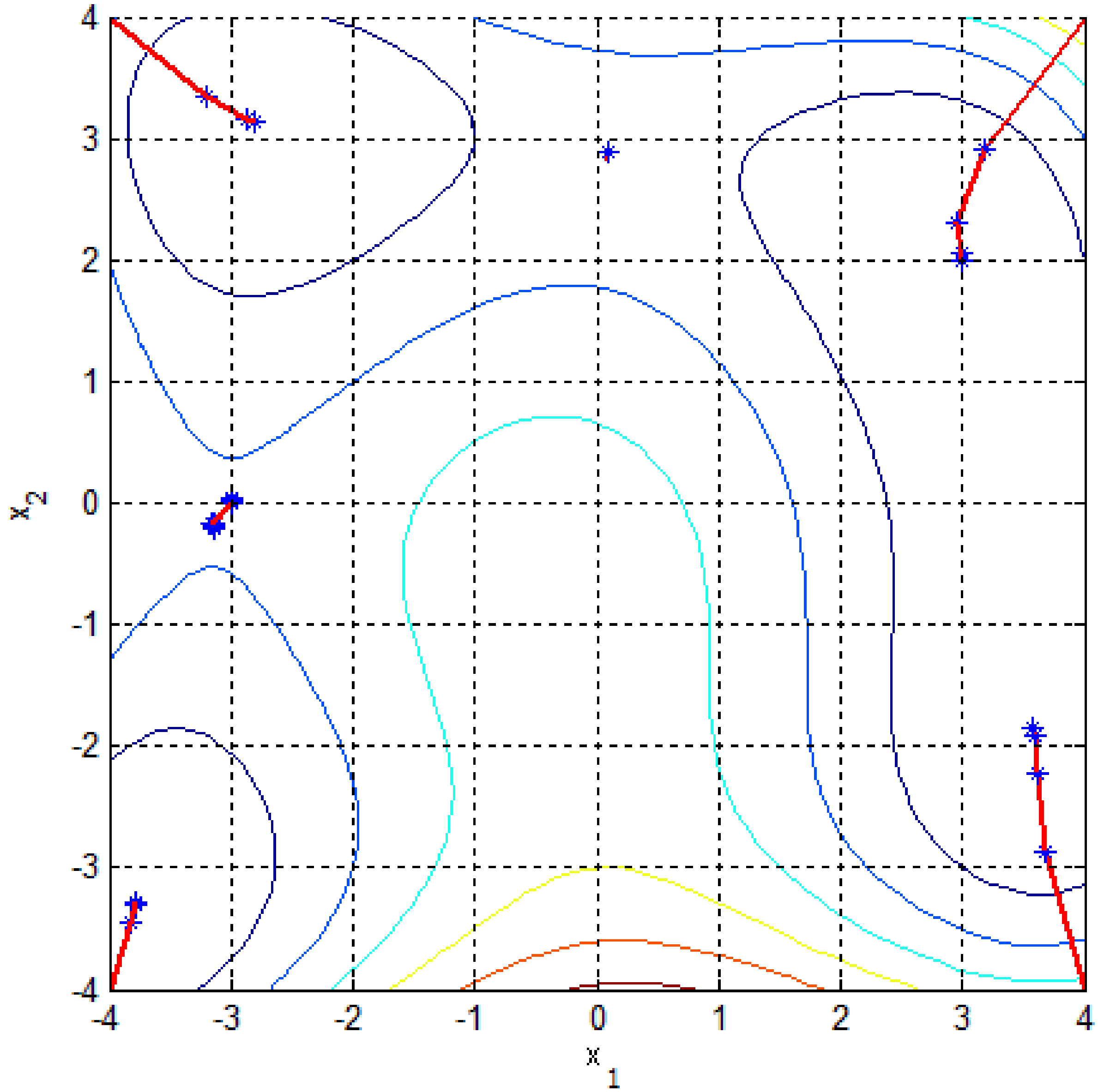

Figure 6: Location of four minima and saddle point.

2.3 Geem test function

Function equation:

$$y = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4 \tag{3}$$

Function figures:

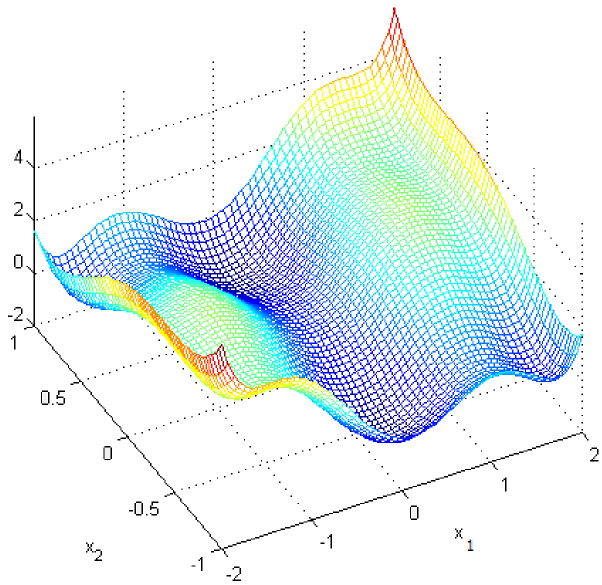

Figure 7: Analyzed function 3D view.

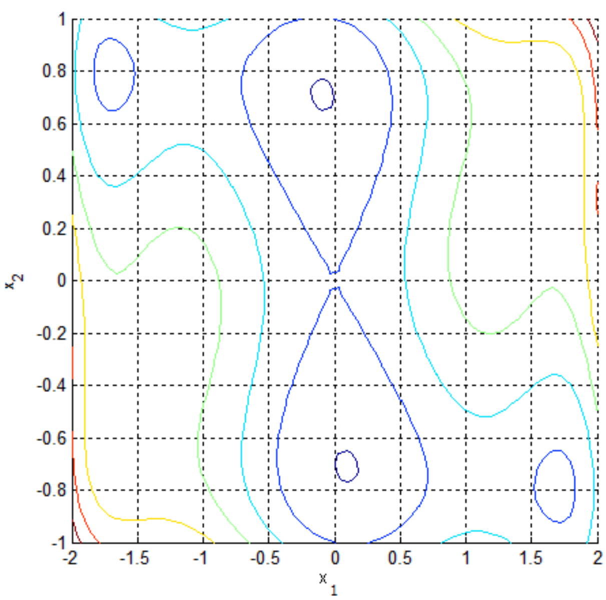

Figure 8: Analyzed function contour view.

Results:

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$0.5, 0.9$	0.708	2.69	-	-
1	$-0.318, 0.846$	-0.699	2.69	0.82	1.407
2	$-7.47 \cdot 10^{-2}, 0.75$	-1.017	$8.942 \cdot 10^{-2}$	0.262	0.319
3	$-8.82 \cdot 10^{-2}, 0.72$	-1.031	$1.47 \cdot 10^{-3}$	$3.36 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$
4	$-8.82 \cdot 10^{-2}, 0.72$	-1.031	$5.11 \cdot 10^{-5}$	$3.36 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-0.5, -0.9$	0.708	2.69	-	-
1	$0.318, -0.846$	-0.699	2.69	0.82	1.407
2	$7.47 \cdot 10^{-2}, -0.75$	-1.017	$8.942 \cdot 10^{-2}$	0.262	0.319
3	$8.82 \cdot 10^{-2}, -0.72$	-1.031	$1.47 \cdot 10^{-3}$	$3.36 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$
4	$8.82 \cdot 10^{-2}, -0.72$	-1.031	$5.11 \cdot 10^{-5}$	$3.36 \cdot 10^{-2}$	$1.33 \cdot 10^{-2}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-2, 1$	1.733	$5.55 \cdot 10^{-2}$	-	-
1	$-1.822, 0.846$	$-2.47 \cdot 10^{-2}$	$5.55 \cdot 10^{-2}$	0.236	1.758
2	$-1.732, 0.8$	-0.207	$1.02 \cdot 10^{-2}$	0.101	0.183
3	$-1.705, 0.796$	-0.215	$6.92 \cdot 10^{-4}$	$2.63 \cdot 10^{-2}$	$7.95 \cdot 10^{-3}$
4	$-1.705, 0.796$	-0.215	$4.21 \cdot 10^{-6}$	$2.63 \cdot 10^{-2}$	$7.95 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$2, -1$	1.733	$5.55 \cdot 10^{-2}$	-	-
1	$1.822, -0.846$	$-2.47 \cdot 10^{-2}$	$5.55 \cdot 10^{-2}$	0.236	1.758
2	$1.732, -0.8$	-0.207	$1.02 \cdot 10^{-2}$	0.101	0.183
3	$1.705, -0.796$	-0.215	$6.92 \cdot 10^{-4}$	$2.63 \cdot 10^{-2}$	$7.95 \cdot 10^{-3}$
4	$1.705, -0.796$	-0.215	$4.21 \cdot 10^{-6}$	$2.63 \cdot 10^{-2}$	$7.95 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$2, 1$	5.733	0.101	-	-
1	$1.798, 0.755$	2.624	0.101	0.317	3.109
2	$1.677, 0.625$	2.15	$3.17 \cdot 10^{-2}$	0.178	0.474
3	$1.621, 0.576$	2.106	$5.48 \cdot 10^{-3}$	$7.4 \cdot 10^{-2}$	$4.46 \cdot 10^{-2}$
4	$1.608, 0.567$	2.104	$2.3 \cdot 10^{-4}$	$1.52 \cdot 10^{-2}$	$1.31 \cdot 10^{-3}$
5	$1.608, 0.567$	2.104	$6.52 \cdot 10^{-7}$	$1.52 \cdot 10^{-2}$	$1.31 \cdot 10^{-3}$

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$-2, -1$	5.733	0.101	-	-
1	$-1.798, -0.755$	2.624	0.101	0.317	3.109
2	$-1.677, -0.625$	2.15	$3.17 \cdot 10^{-2}$	0.178	0.474
3	$-1.621, -0.576$	2.106	$5.48 \cdot 10^{-3}$	$7.4 \cdot 10^{-2}$	$4.46 \cdot 10^{-2}$
4	$-1.608, -0.567$	2.104	$2.3 \cdot 10^{-4}$	$1.52 \cdot 10^{-2}$	$1.31 \cdot 10^{-3}$
5	$-1.608, -0.567$	2.104	$6.52 \cdot 10^{-7}$	$1.52 \cdot 10^{-2}$	$1.31 \cdot 10^{-3}$

When we chose starting point at $x_1 = -1, x_2 = 0.7$ or $x_1 = 1, x_2 = -0.7$ or $x_1 = 0, x_2 = 0$ then program returned: 'local minimum or maximum does not exist'. This is caused by saddle points there. Situation can be seen on figure 9. On the other hand start points at $x_1 = 1.25, x_2 = 0.2$ and $x_1 = -1.25, x_2 = -0.2$ program described as neighbourhood of local maxima.

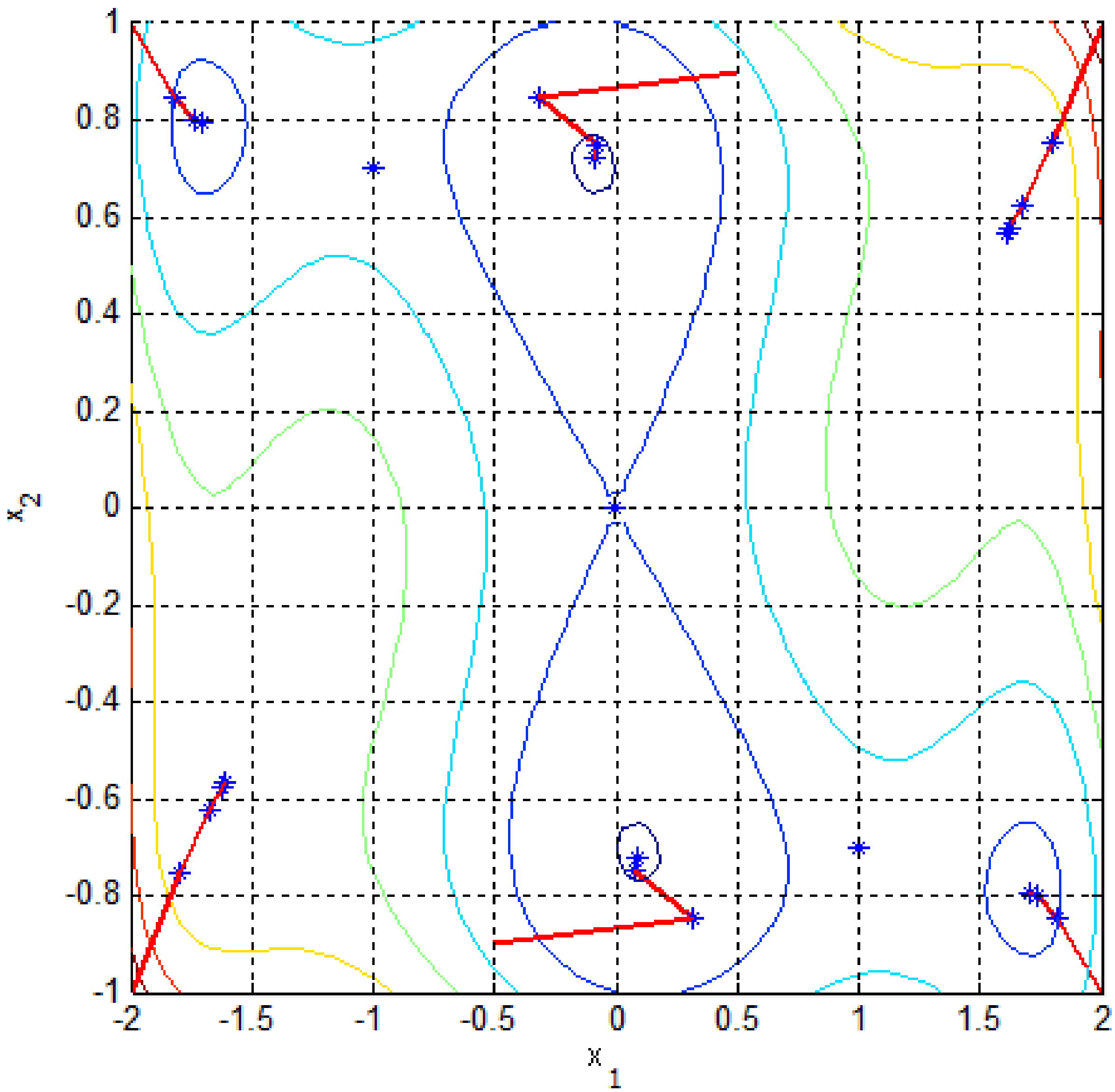

Figure 9: Location of four minima and saddle points.

2.4 Practical example - spring-carts system

Interpretation:
Figure 10 shows two frictionless rigid bodies (carts) A and B connected by three linear elastic springs having spring constants $k_1 n k_2, k_3$. The springs are at their natural positions when the applied force P is zero. The task is to find the displacement x_1 and x_2 under the force P by using principle of minimum potential energy.

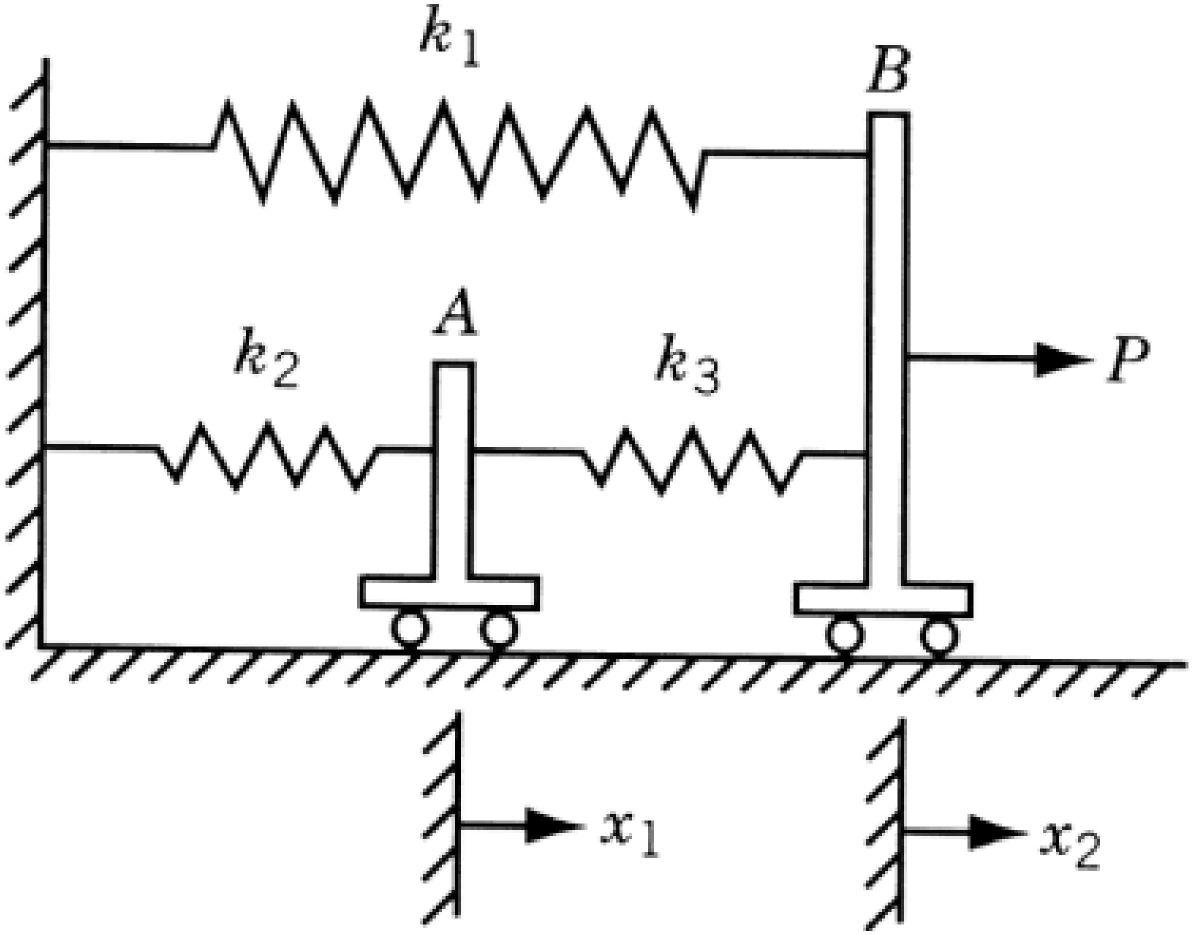

Figure 10: Spring-cart system.

Function equation:

$$y = (\frac{1}{2}k_2x_1^2 + \frac{1}{2}k_3(x_2 - x_1)^2 + \frac{1}{2}k_1x_2^2) - Px_2 \tag{4}$$

Lets assume: $k_1 = 12, k_2 = 30, k_3 = 15, P = 100$.

Function figures:

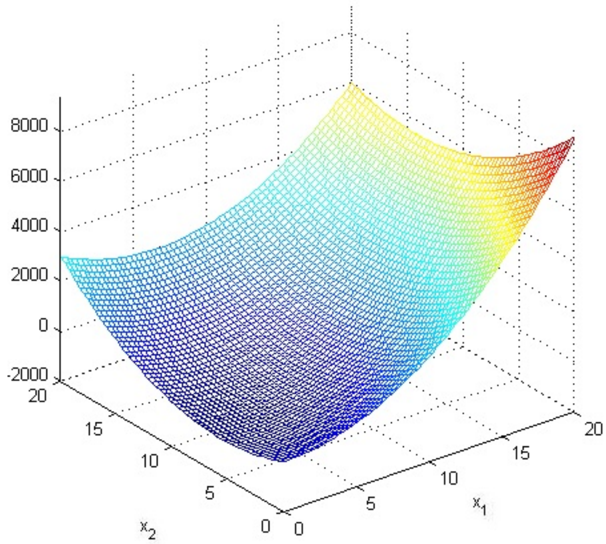

Figure 11: Analyzed function 3D view.

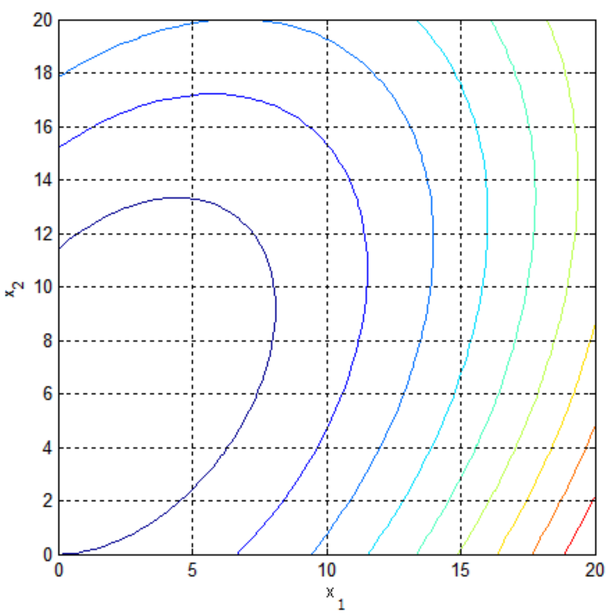

Figure 12: Analyzed function contour view.

iteration	point coordinates	function value	C_1 value	C_2 value	C_3 value
0	$10, 10$	500	101.736	-	-
1	$2.083, 3.75$	-187.5	101.736	10.086	787.5
2	$2.083, 3.75$	-187.5	$5.6.52 \cdot 10^{-31}$	10.086	787.5

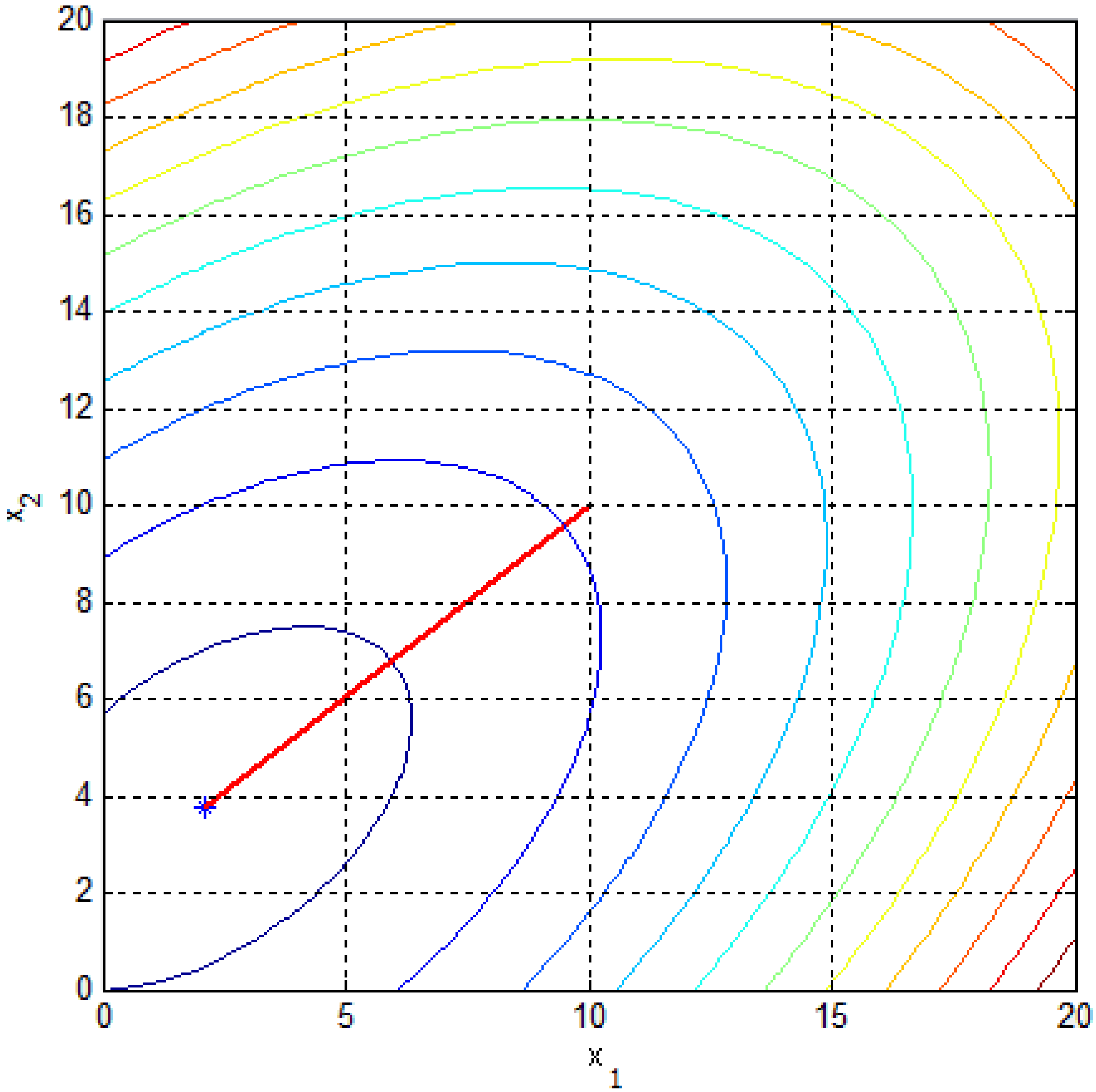

Figure 13: Founded minimum.

Example taken from "Engineering Optimization: Theory and Practice" (Authors: Singiresu S. Rao,S. S. Rao).