

VBC

ZS 2021/22

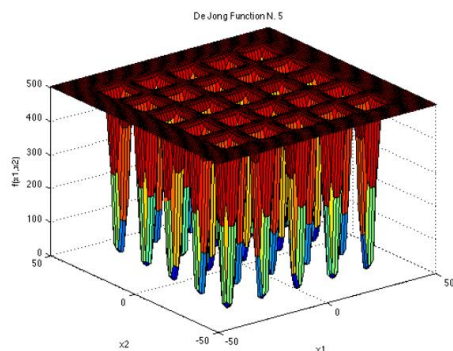
Tomáš Benda, 200604

task #2

Zadání

Pro dané funkce otestovat metaheuristiky genetický algoritmus a simulované žíhání. U genetického algoritmu vyzkoušet různá nastavení parametrů, simulované žíhání testovat s defaultním nastavením z toolboxu. Výsledky zpracovat pomocí statistiky, 2D grafy vykreslit s nalezeným minimem.

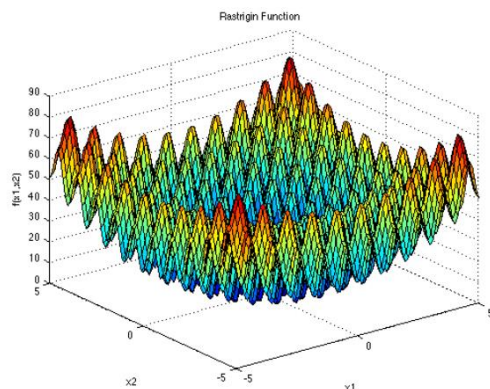
De Jongova funkce N. 5



$$f(\mathbf{x}) = \left(0.002 + \sum_{i=1}^{25} \frac{1}{i + (x_1 - a_{1i})^6 + (x_2 - a_{2i})^6} \right)^{-1}, \text{ where}$$
$$\mathbf{a} = \begin{pmatrix} -32 & -16 & 0 & 16 & 32 & -32 & \dots & 0 & 16 & 32 \\ -32 & -32 & -32 & -32 & -32 & -16 & \dots & 32 & 32 & 32 \end{pmatrix}$$

De Jongova funkce N.5 je definovaná pro $x_i \in [-65.536, 65.536]$ pro $i = 1, 2$. Globální minimum pro 2D je pak v bodě $x = [-32, -32]$ a funkční hodnota je $y = 0.998$.

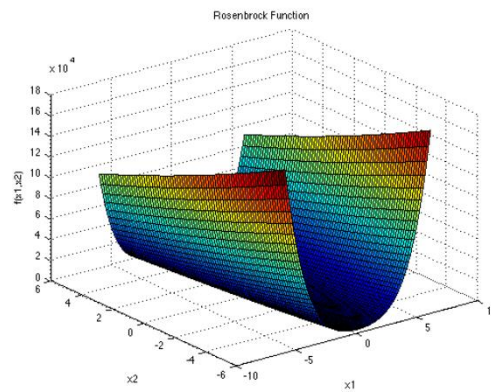
Rastriginova funkce



$$f(\mathbf{x}) = 10d + \sum_{i=1}^d [x_i^2 - 10 \cos(2\pi x_i)]$$

Rastriginova funkce je definovaná pro $x_i \in [-5.12, 5.12]$ pro všechna $i = 1, 2, \dots, d$. Globální minimum pro 2D je pak v bodě $x = [0, 0]$ a funkční hodnota je $y = 0$.

Rosenbrockova funkce



$$f(\mathbf{x}) = \sum_{i=1}^{d-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$$

Rosenbrockova funkce je definovaná pro $x_i \in [-5, 10]$ pro všechna $i = 1, 2 \dots, d$. V tomto protokolu je omezena na $x_i \in [-2.048, 2.048]$ pro všechna $i = 1, 2 \dots, d$. Globální minimum pro 2D je pak v bodě $x = [1, 1]$ a funkční hodnota je $y = 0$.

Postup řešení

K realizaci funkce De Jong N. 5 a Rastrigin byla využita implementace z knihovny Matlabu. Pro funkci Rosenbrock byl vytvořen skript. Po formální stránce je obdobou implementace funkce De Jong N. 5.

Simulované žíhání

Pro simulované žíhání bylo ponecháno, dle zadání, defaultní nastavení solveru z Global Optimization Toolboxu. Pro každý restart byl vygenerován náhodný počáteční bod v rámci omezení oboru hodnot daných funkcí, ve kterém bylo zahájeno hledání minima.

Pro každou zadanou funkci a příslušný počet dimenzí (počet proměnných) bylo provedeno 1000 restartů. Výjimku tvoří dimenze 50D a 100D, jejichž počet restartů byl omezen na 100 z důvodu dlouhého běhu výpočtu.

Dané výsledky, podobně jako u druhé heuristiky, byly zpracovány do tabulek v části Výsledky. Z restartů byla vybráno to řešení, jehož funkční hodnota byla minimální (sloupec best). Pro parametry uvedené v tabulkách byla zjištěno také minimální hodnota, maximální hodnota a průměr.

Pro 2D funkce byl vykreslen následně graf se zobrazeným globálním minimem nalezeným solverem. Grafy byly také zpracovány pro čas běhu restartů, velikost minimální funkční hodnoty pro jednotlivé restarty a počet evaluací fitness funkce.

Genetický algoritmus

U genetického algoritmu proběhlo několik nastavení parametrů, které Global Optimization Toolbox pro tuto heuristiku nabízí. Zejména se jednalo o parametry: Selection Function, Max Generations, Population Size, Creation Function. Následně byly vybrány tyto možnosti pro solver:

- Population Size = 300
 - Velikost populace byla zvolena jako kompromis mezi rychlostí výpočtu a přesností nalezení globálního minima
- Selection Function = Selection Tournament
 - Daná funkce pro selekci jedinců, výběr podporuje nejlepší jedince
- Max Generation = 10000
 - Zvoleno vysoké číslo tak, aby nebyl proces zastaven kvůli vysokému počtu generací dle zadání (defaultně $100 \cdot \text{počet_proměnných}$ – pro 2D tedy $100 \cdot 2 = 200$)
- Elite Count = $0.08 \cdot \text{Population Size}$
 - Oproti defaultnímu nastavení zvýšení počtu jedinců, kteří přežijí do následující generace
- Creation Function = Crossover LaPlace
 - Vytvoření původní populace pomocí uniformního dělení
- Crossover Function = Crossover LaPlace
 - S danými parametry v průběhu testování dobré výsledky

Výsledky pokusů

Simulované žíhání

Následující tři tabulky ukazují statistické zpracování hodnot x_1 , x_2 a $fval$ (funkční hodnota) pro všechny požadované funkce a dimenze. V tabulkách je zobrazena maximální, minimální hodnota, průměr a hodnota best – hodnota nejlepšího výsledku pro daný parametr.

		x1			
	D	max.	min.	průměr	best
Rastrigin	2	1.9899	-1.9901	0.0112	0.0000
	5	4.9748	-3.9806	0.0319	0.9953
	10	7.9587	-6.9636	-0.0269	-0.9928
	50	6.9631	-7.9598	0.6167	-2.9881
	100	4.9746	-5.9732	-0.0496	1.9890
Rosenbrock	2	1.0032	0.9972	1.0000	1.0000
	5	1.0015	0.9986	1.0000	1.0000
	10	1.0018	0.9981	1.0000	0.9999
	50	1.0020	0.9971	1.0000	1.0010
	100	1.0024	0.9971	1.0000	1.0001
De Jong N. 5	2	789.5631	-521.7322	-3.4396	-31.9784

Tabulka 1: Statistika pro x_1 – simulované žíhání

		x2			
	D	max.	min.	průměr	best
Rastrigin	2	1.9904	-2.0127	-0.0001	0.0000
	5	4.9748	-4.9758	-0.0239	0.0002
	10	6.9657	-6.9636	0.0936	-0.0019
	50	6.9630	-6.9663	-0.1295	-3.9791
	100	5.9646	-6.9674	0.1092	0.9936
Rosenbrock	2	1.0078	0.9938	1.0000	1.0000
	5	1.0016	0.9985	1.0000	1.0000
	10	1.0016	0.9978	1.0000	0.9998
	50	1.0022	0.9972	1.0000	1.0014
	100	0.9998	0.9962	0.9998	1.0000
De Jong N. 5	2	869.1658	-500.1944	-13.3752	-31.9783

Tabulka 2: Statistika pro x_2 – simulované žíhání

		fval			
	D	max.	min.	průměr	best
Rastrigin	2	4.9749	0.0000	0.8289	0.0000
	5	48.7528	0.9951	13.4759	0.9951
	10	148.2497	9.9533	60.4528	9.9533
	50	696.5479	343.3299	478.8785	343.3299
	100	1286.6181	747.3712	924.7070	747.3712
Rosenbrock	2	0.0028	0.0000	0.0000	0.0000
	5	0.0043	0.0000	0.0003	0.0000
	10	0.0087	0.0001	0.0014	0.0001
	50	0.0380	0.0127	0.0245	0.0127
	100	0.0924	0.0511	0.0664	0.0511
De Jong N. 5	2	499.9994	0.9980	181.1802	0.9980

Tabulka 3: Statistika pro fval (funkční hodnota) – simulované žihání

Následují tabulky se statistikami pro počet evaluací fitness funkce a čas výpočtu.

		počet evaluací fitness			
	D	max.	min.	průměr	best
Rastrigin	2	5457	1067	2050.5360	2214
	5	15001	2736	5653.4560	6400
	10	30001	5995	11053.0800	17063
	50	110607	37415	58177.8500	61610
	100	266422	90232	136126.5200	159546
Rosenbrock	2	4758	1147	2046.5480	1600
	5	15001	2758	7624.5890	7229
	10	30001	7824	21257.8920	18668
	50	150014	78574	126851.2000	122941
	100	300099	148891	263036.4900	300008
De Jong N. 5	2	5797	1009	1623. 5690	1417

Tabulka 4: Statistika pro počet evaluací fitness – simulované žihání

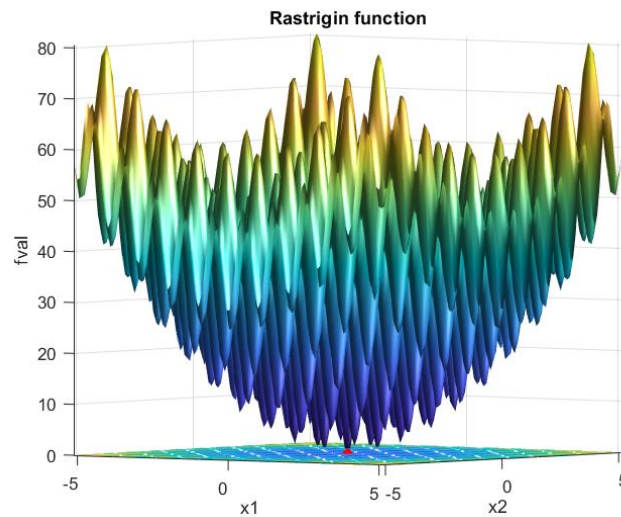
		čas			
	D	max.	min.	průměr	best
Rastrigin	2	1.7656	0.1406	0.3013	0.3125
	5	1.4688	0.2031	0.4950	0.6250
	10	2.9219	0.4844	0.9874	1.5469
	50	30.1719	5.2031	12.4923	14.2500
	100	75.5156	14.3906	30.6636	39.0000
Rosenbrock	2	1.2031	0.1406	0.3325	0.2500
	5	1.3750	0.2188	0.6660	0.6250
	10	3.1094	0.6719	1.9646	1.6875
	50	45.7656	20.1094	36.6836	36.2500
	100	100.2813	35.6094	76.8803	88.7813
De Jong N. 5	2	1.3594	0.0781	0.1790	0.1563

Tabulka 5: Statistika pro čas běhu výpočtu – simulované žihání

Pro 2D funkce bylo za pomoci simulovaného žíhání nalezeny následující optima:

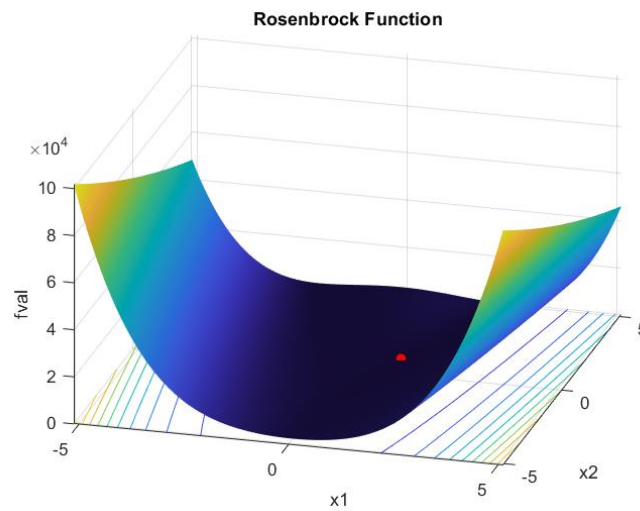
Rastrigin

$x_1 = 0.0000$; $x_2 = 0.0000$; $fval = 0.0000$



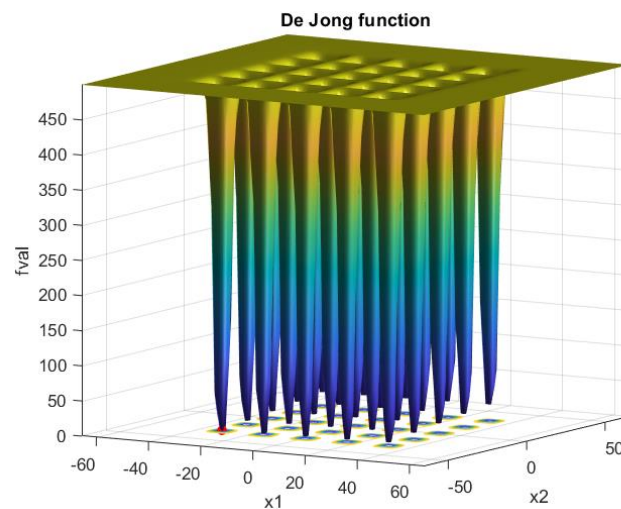
Rosenbrock

$x_1 = 1.0000$; $x_2 = 1.0000$; $fval = 0.0000$;



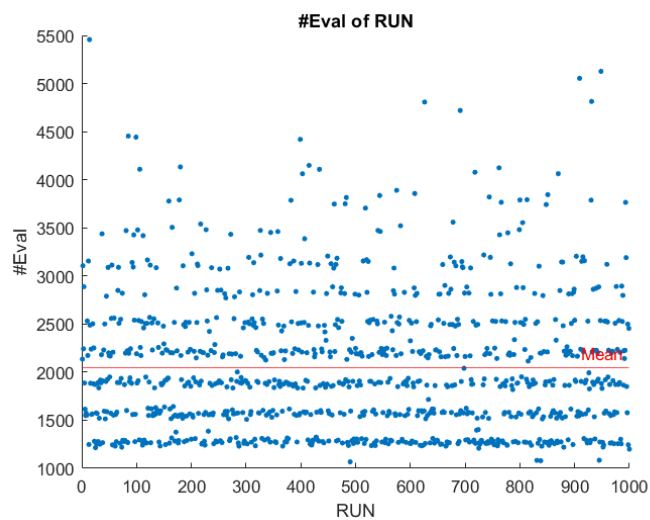
De Jong N. 5

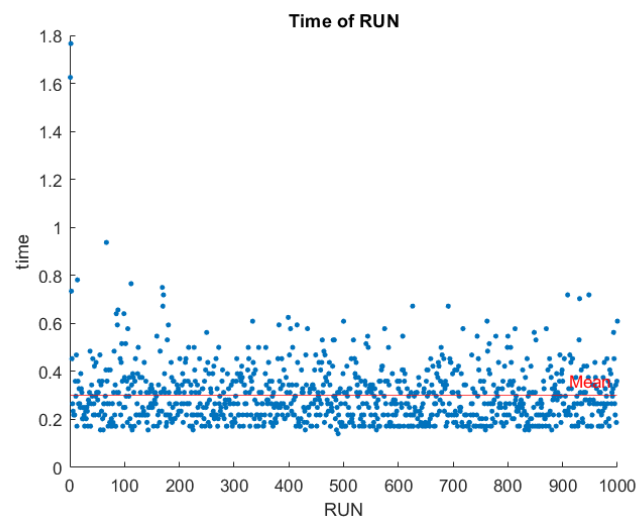
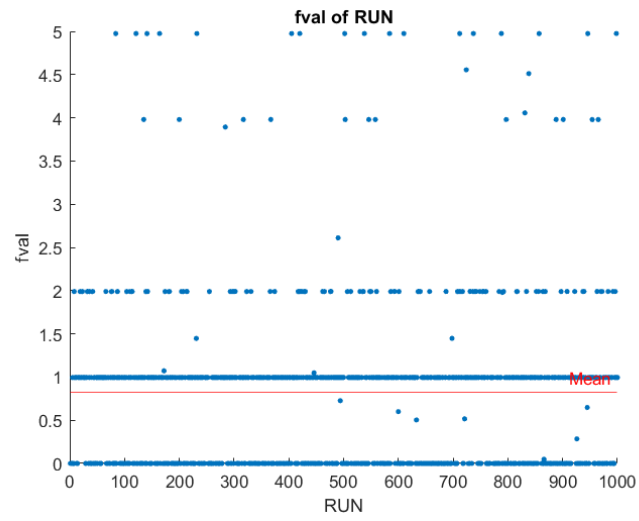
$x_1 = -31.9782$; $x_2 = -31.9782$; $fval = 0.9980$;



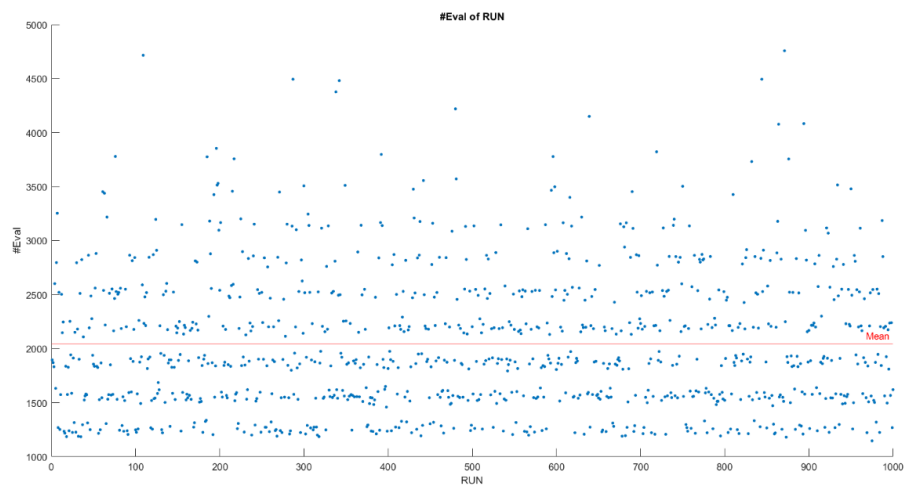
Dále byly zpracovány grafy, které přibližují průběh restartů. Konkrétně to jsou grafy pro počet evaluací fitness v průběhu restartů, hodnoty fval v průběhu restartů a čas restartů. Grafy jsou prezentovány pouze pro 2D funkce.

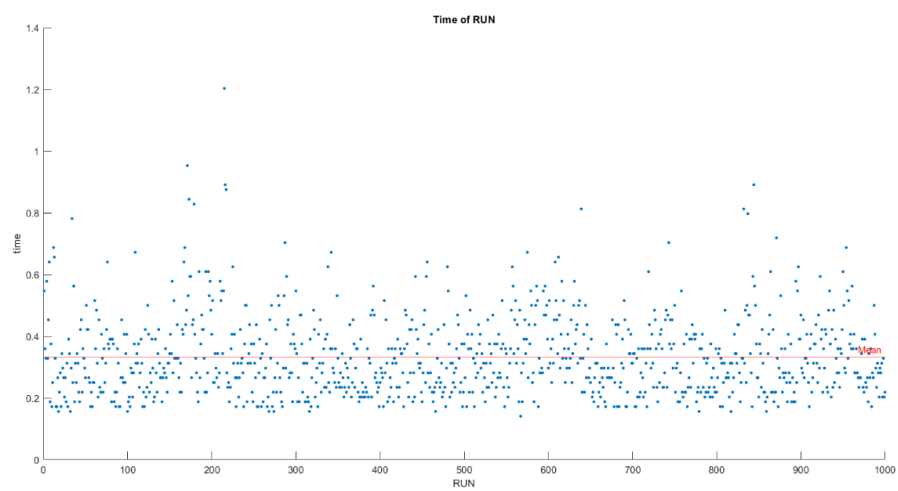
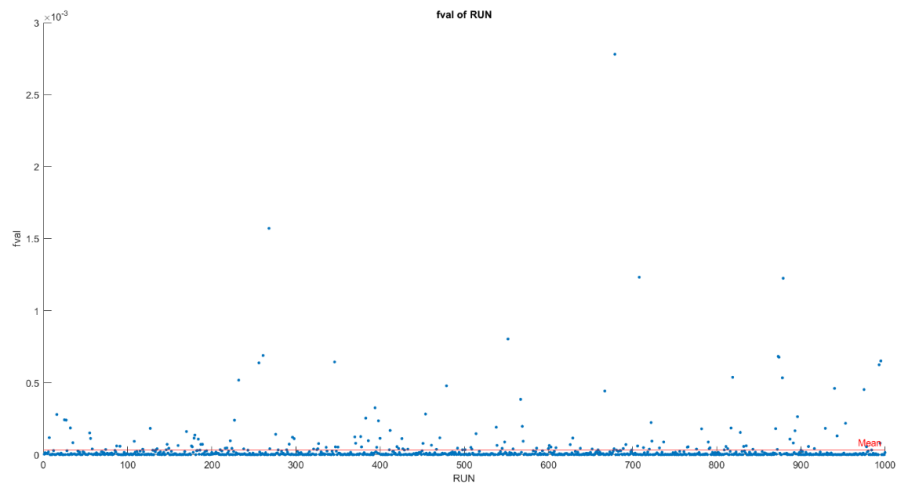
Rastrigin 2D



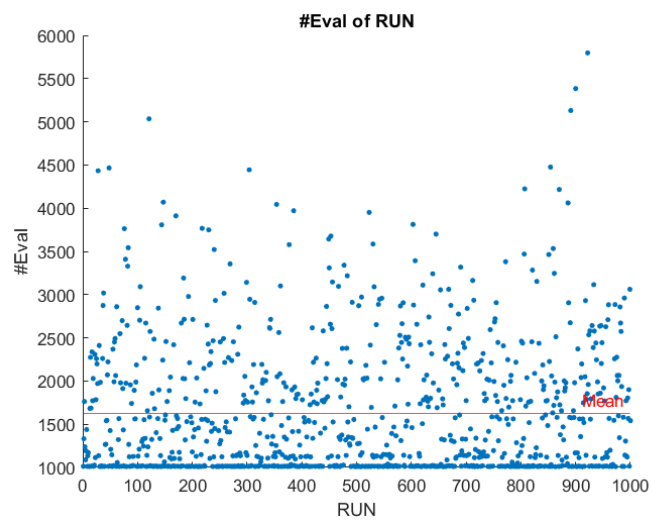


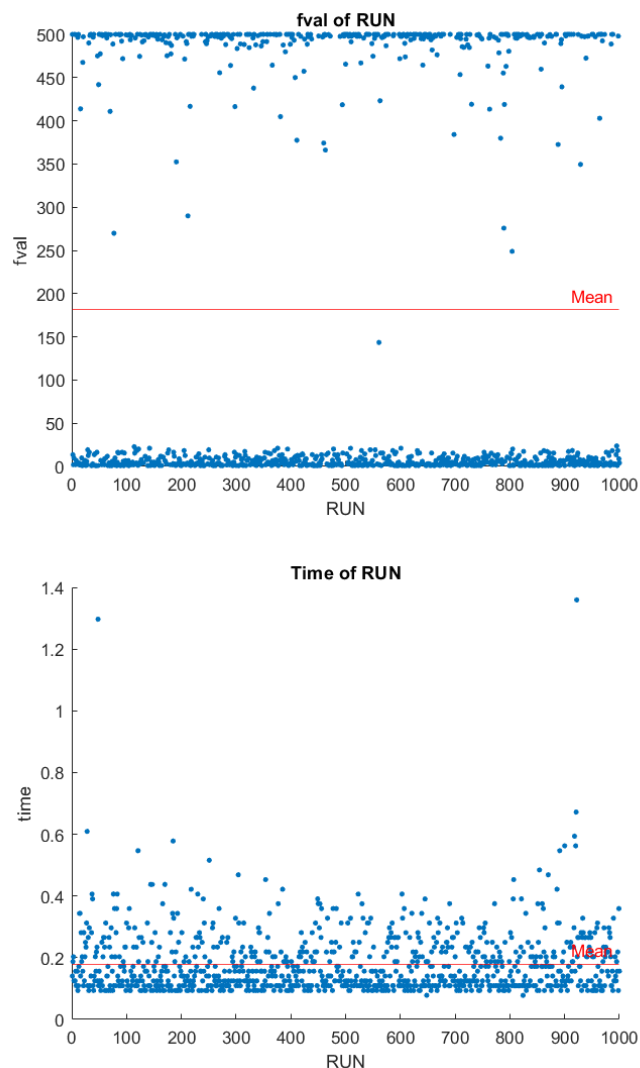
Rosenbrock 2D





De Jong N. 5





Genetický algoritmus

Podobně jako u simulovaného žíhání byly vytvořeny tabulky s maximální, minimální, průměrnou a nejlepší hodnotou pro x_1 , x_2 , fval, počet evaluací fitness, čas. Navíc byla přidána tabulka s počtem generací se stejnými parametry.

V některých případech (zejména pro funkce 2D) genetický algoritmus z Global Optimization Toolboxu našel globální minimum v několika restartech. Například pro Rastrigin 2D bylo nalezeno globální minimum s funkční hodnotou 0 v 999 restartech z 1000. V tabulce je uvedena vždy jedna hodnota jako nejlepší.

		x1			
	D	max.	min.	průměr	best
Rastrigin	2	0.0000	0.0000	0.0000	0.0000
	5	0.9950	-0.9950	0.0109	0.0000
	10	1.9899	-0.9950	0.0199	0.0000
	50	2.9849	-1.9900	0.0328	0.0000
	100	4.9747	-2.9850	-0.0268	0.0000
Rosenbrock	2	1.0138	0.9825	1.0000	1.0000
	5	1.0245	0.9870	0.9997	1.0000
	10	1.0002	0.9994	1.0000	1.0000

	50	1.0008	0.9790	0.9999	1.0000
	100	1.0001	0.9998	1.0000	1.0000
De Jong N.5	2	-15.9864	-31.9783	-31.5945	-31.9783

Tabulka 6: Statistika pro x1 – genetický algoritmus

		x2			
	D	max.	min.	průměr	best
Rastrigin	2	0.9950	0.0000	0.0010	0.0000
	5	0.9950	-0.9950	0.0000	0.0000
	10	0.9950	-1.9899	-0.0199	0.0000
	50	1.9899	-1.9900	0.0030	0.0000
	100	2.9850	-2.9850	0.0149	-0.0002
Rosenbrock	2	1.0346	0.9653	1.0000	1.0000
	5	1.0617	0.9675	0.9992	1.0000
	10	1.0007	0.9993	1.0000	1.0000
	50	1.0009	0.9918	0.9999	1.0000
	100	1.0001	0.9997	1.0000	1.0000
De Jong N.5	2	- 31.9703	- 31.9783	- 31.9781	-31.9783

Tabulka 7: Statistika pro x2 – genetický algoritmus

		fval			
	D	max.	min.	průměr	best
Rastrigin	2	0.9950	0.0000	0.0010	0.0000
	5	2.9849	0.0000	0.3104	0.0000
	10	9.9496	0.0000	2.4566	0.0000
	50	76.6118	15.9193	38.6447	15.9193
	100	185.0631	54.7228	109.9088	54.7228
Rosenbrock	2	0.0306	0.0000	0.0001	0.0000
	5	22.7473	0.0000	0.1477	0.0000
	10	60.9790	0.0000	0.4725	0.0000
	50	715.1319	0.0000	12.3844	0.0000
	100	169.1091	0.0000	8.2173	0.0000
De Jong N.5	2	1.9920	0.9980	1.0219	0.9980

Tabulka 7: Statistika pro fval (funkční hodnota) – genetický algoritmus

		počet evaluací fitness			
	D	max.	min.	průměr	best
Rastrigin	2	19368	15228	17667.5640	17988
	5	30132	21576	24079.5960	24060
	10	44208	26820	31296.7200	32340
	50	102168	53592	63034.7880	65460
	100	155436	81744	96153.4080	94992
Rosenbrock	2	56076	15780	18018.9120	18264
	5	178344	22128	40752.4800	24612
	10	205668	31512	76288.8600	33720

	50	447720	143292	229122.4800	147156
	100	567228	277152	334938.1200	321864
De Jong N.5	2	34548	14952	20747.1720	25716

Tabulka 8: Statistika pro počet evaluací fitness – genetický algoritmus

		počet generací			
	D	max.	min.	průměr	best
Rastrigin	2	69	54	62.8390	64
	5	108	77	86.0710	86
	10	159	96	112.2200	116
	50	369	193	227.2130	236
	100	562	295	347.2080	343
RosenBrock	2	202	56	64.1120	65
	5	645	79	146.4800	88
	10	744	113	275.2350	121
	50	1621	518	828.9800	532
	100	2054	1003	1212.3700	1165
De Jong N.5	2	124	53	73.9970	92

Tabulka 9: Statistika pro počet generací – genetický algoritmus

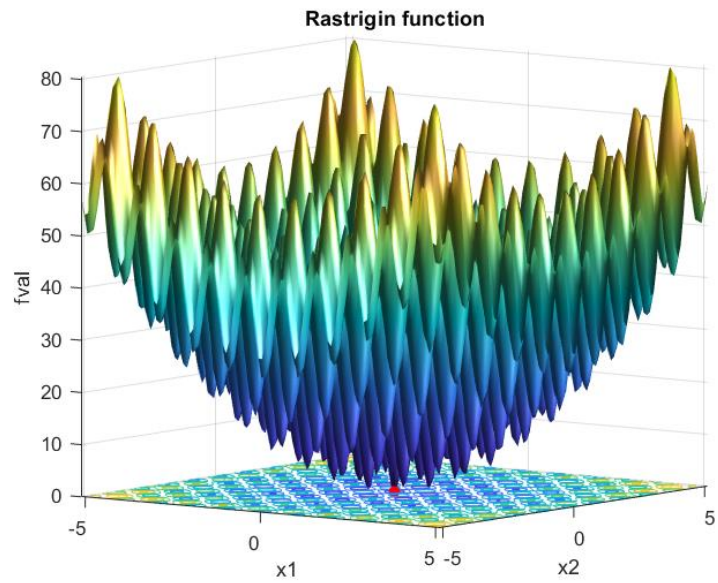
		čas			
	D	max.	min.	průměr	best
Rastrigin	2	1.1719	0.2969	0.3615	0.5469
	5	2.1406	0.3438	0.5155	0.4688
	10	1.5469	0.5469	0.7474	0.6875
	50	10.1250	1.6719	2.1256	2.2969
	100	9.5156	2.3906	3.4702	3.2656
Rosenbrock	2	1.3906	0.2188	0.2951	0.3750
	5	3.4063	0.3125	0.6976	0.4219
	10	3.8906	0.5000	1.3716	0.5938
	50	10.4063	3.1406	5.4192	3.4063
	100	17.3438	8.1563	10.1316	9.2969
De Jong N.5	2	0.9375	0.2969	0.4733	0.6406

Tabulka 5: Statistika pro čas běhu výpočtu – genetický algoritmus

Optima pak byla vykreslena v následujících grafech.

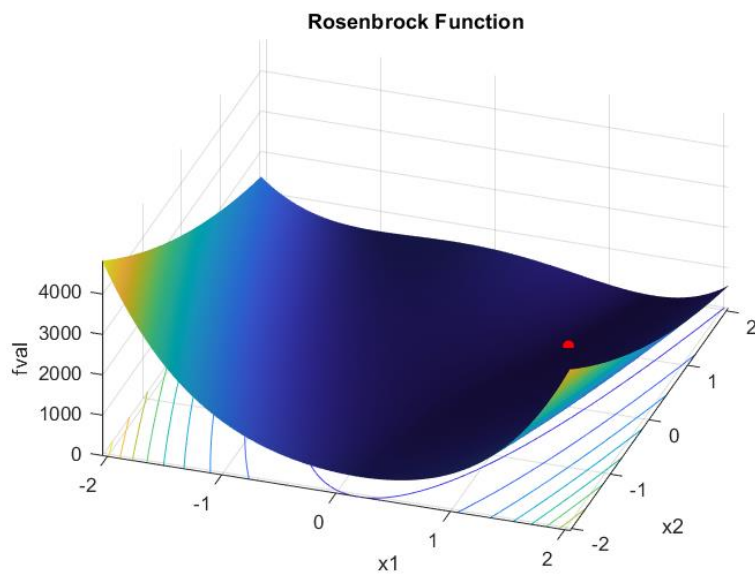
Rastrigin

x1 = 0.0000; x2 = 0.0000; fval = 0.0000;



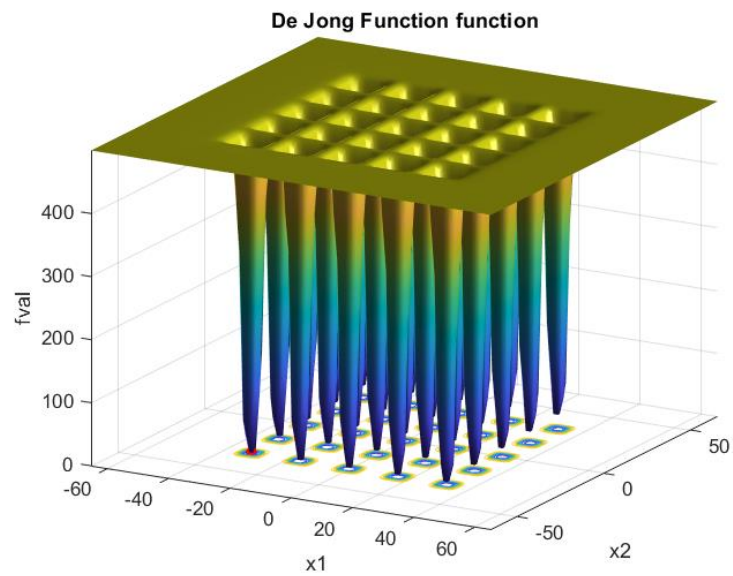
Rosenbrock

$x_1 = 1.0000$; $x_2 = 1.0000$; $fval = 0.0000$;



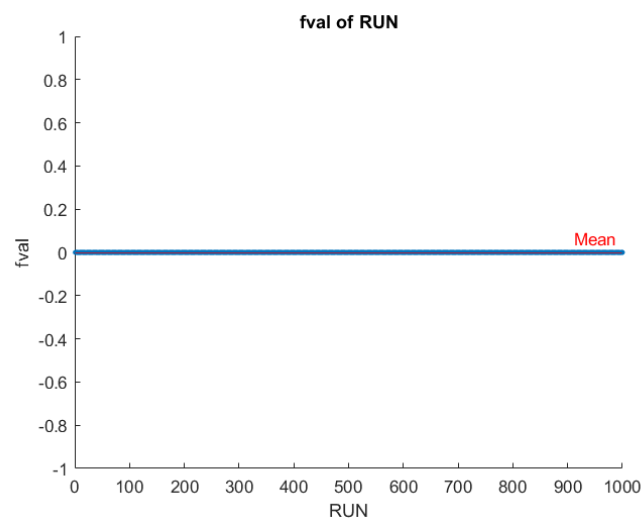
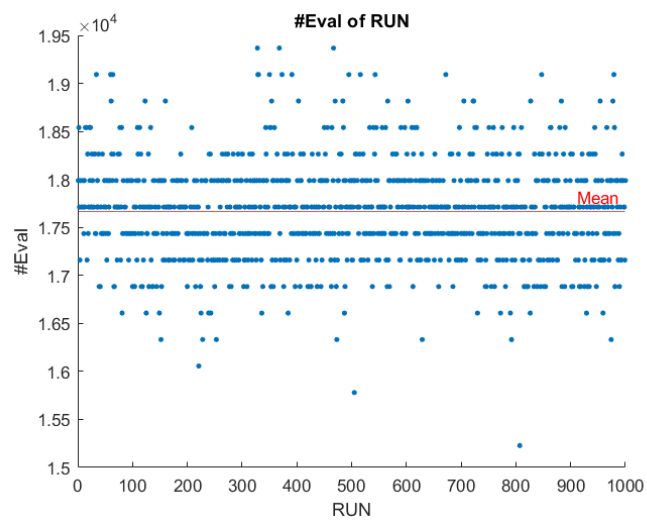
De Jong N. 5

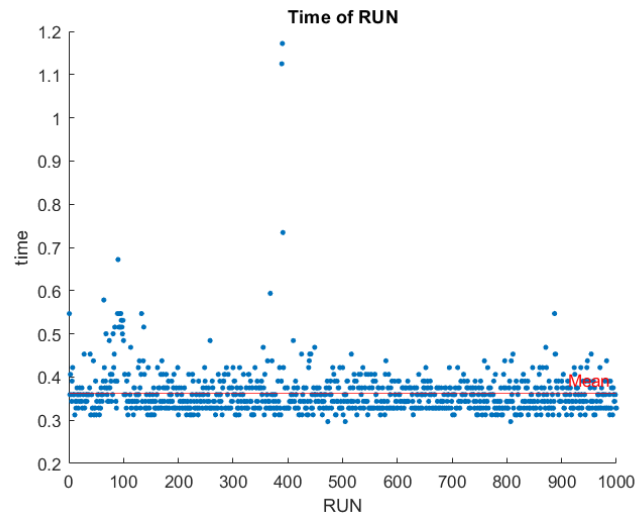
$x_1 = -31.9783$; $x_2 = -31.9783$; $fval = 0.9980$;



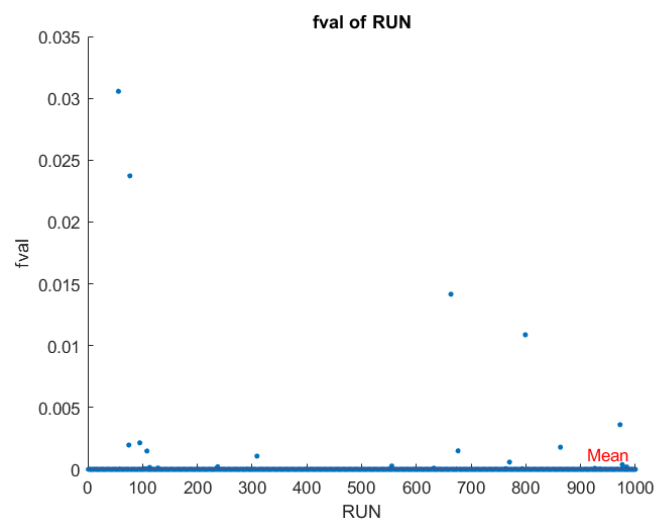
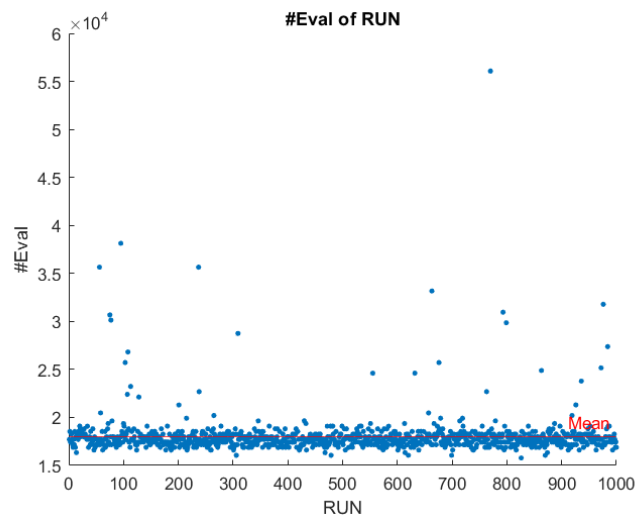
Dále grafy zobrazující funkční hodnotu v průběhu restartů, počet evaluací fitness a čas běhu výpočtu.

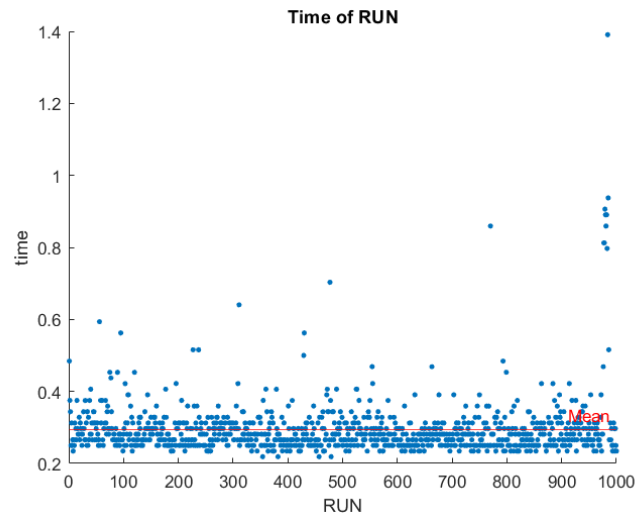
Rastrigin



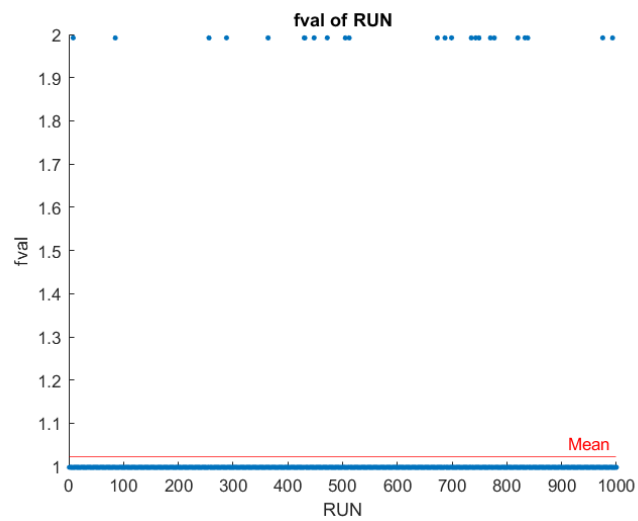
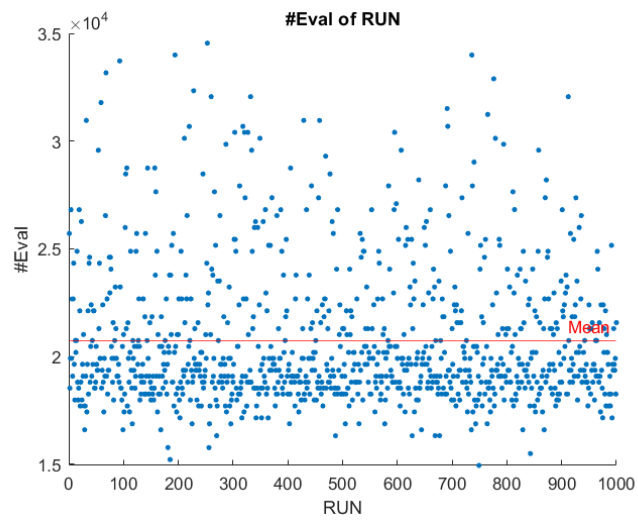


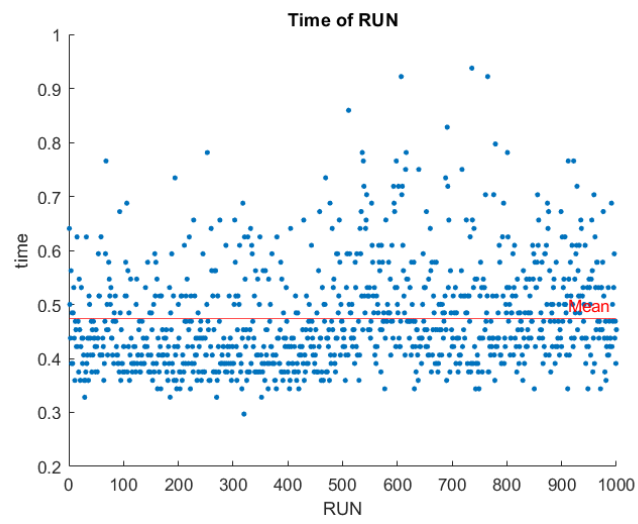
Rosenbrock





De Jong N. 5





Zhodnocení výsledků

Simulované žíhání

Pro simulované žíhání bylo pomocí solveru pro grafy 2D dosaženo globálního minima. Pro Rastriginovu funkci již nebylo pro více dimenzí (5D, 10D, 50D, 100D) nalezeno správné globální minimum. Oproti tomu u funkce Rosenbrock bylo i pro více proměnných nalezeno minimum podobné globálnímu. Toto je způsobeno nejspíše tím, že u Rastriginovy funkce bylo dosaženo lokálního minima a solver s danými parametry nebyl schopný toto minimum opustit. Tvar Rosenbrockovy funkce pak nasvědčuje, že pro daný biologicky inspirovaný algoritmus bylo jednodušší nalézt globální minimum nebo se mu v rámci tolerance přiblížit.

Zajímavé jsou průměrné hodnoty časů restartů (tabulka 5). Zatímco pro například Rastriginovu funkci 2D byl průměrný čas nalezení minima 0.3013 s, pro stejnou funkci ve 100D trval výpočet 30.6636 s. To je více než 100x nárůst hodnoty. Patrnější je toto pro Rosenbrockovu funkci, kde je nárůst až 200x. Z tohoto důvodu bylo provedeno pro 50D a 100D pouze 100 restartů.

Delší čas výpočtu souvisí také s rostoucími hodnotami evaluací fitness (výpočtu funkční hodnoty) pro vyšší dimenze. Jelikož byly evaluace častější a výpočet s více proměnnými trval déle, i výsledný čas byl větší.

Genetický algoritmus

U genetického algoritmu bylo hlavním úkolem najít takové nastavení, které bude kompromisem mezi přesností řešení a rychlostí výpočtu. Některá nastavení, která byla v průběhu zavrhnuta, selhávala již pro deset proměnných, optimum pro 50D či 100D pak bylo například v řádech tisíců pro Rastriginovu funkci.

Postupným zkoušením parametrů pro nastavení solveru bylo vybráno nastavení, které dle mého názoru nabízí zmíněný kompromis.

Pro Rastriginovu funkci jsou výsledky při vyšším počtu proměnných lepší, než u simulovaného žíhání s defaultním nastavením. Čas výpočtu u této funkce je pro 100D až desetkrát lepší – průměr simulovaného žíhání byl 30.66 s zatímco u genetického algoritmu s daným nastavením 3.26 s. Díky tomu nebyl problém provést požadovaných 1000 restartů v rozumném čase.

V příslušném grafu pro funkční hodnoty De Jong N. 5 je vidět, že fval se blíží optimu nebo naopak hodnota je *fval* ~ 2. Podobně je tomu u simulovaného žíhání. Lze předpokládat, že pro danou funkční hodnotu se algoritmy dostaly do lokálního minima přibližně [-15.9864, -31.9703], ze kterého se jim již nepovedlo dostat a nalézt tak globální minimum.