

Trabalho Final de Inteligência Artificial

Rubens Anderson de Sousa Silva, 362984

Ciência da Computação, Universidade Federal do Ceará, Fortaleza/Ceará, Brasil. 2017

rubsandsu@gmail.com

www.github.com/rubensandersonn

1 Otimização

Dada a função ... e seus gráficos plot e contour, podemos observar que há vários pontos de mínimo local e pontos de sela ($Df = \{x,y \mid -3 \leq x \leq 3, -2 \leq y \leq 2\}$).

Figura 1: função f no domínio

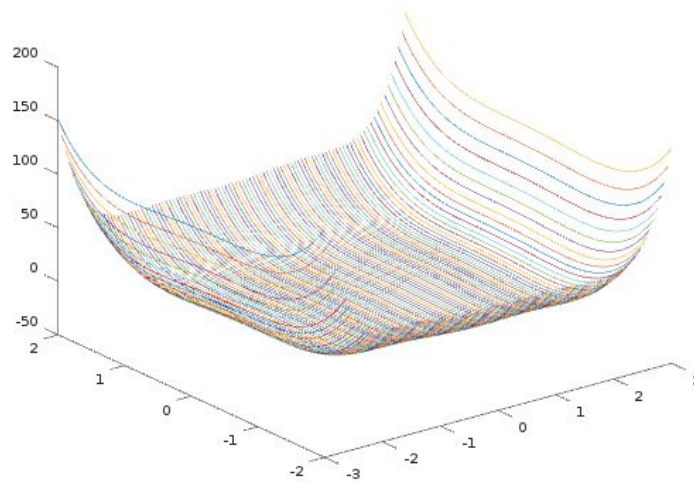
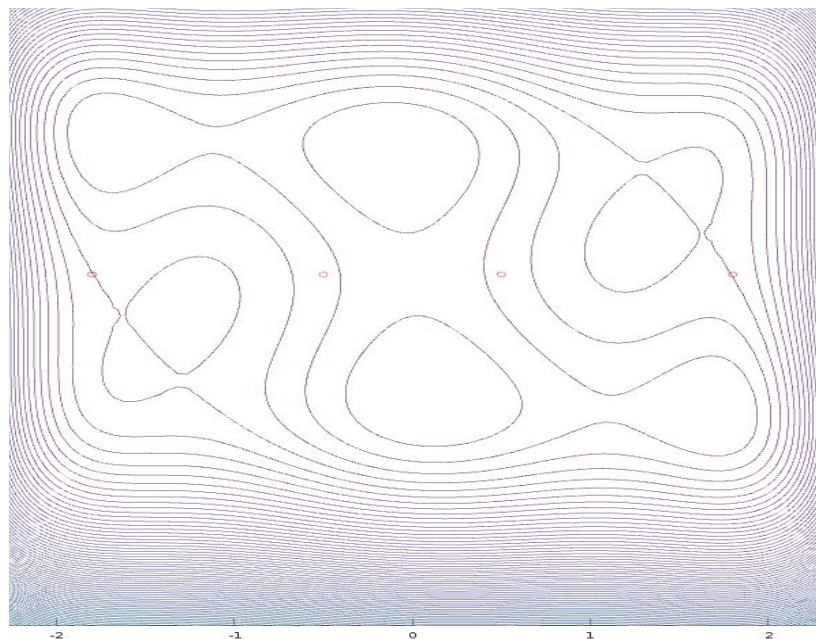


Figura 2: Linhas de contorno de f no mesmo domínio com os pontos iniciais de teste.



Observando as linhas de contorno da função, uma escolha de $0.05 = \alpha$ parece razoável, visto que a função não varia tão rapidamente (para diminuir o valor de α) nem é aproximadamente linear (para manter um α grande).

Direções com Gradiente

Para testar a função, escolhi 3 pontos iniciais e observei o seu comportamento.

Tabela 1: pontos iniciais para teste de f .

Pontos	x_i	y_i
ponto 1	0.5	0.0
ponto 2	-0.5	0.0
ponto 3	-1.8	0.0

Utilizando uma função de minimização com direções de atualização baseadas no gradiente, obtemos o seguinte resultado:

Figura 3: para o ponto 1, o valor de f no ponto final foi de -1.0316 com número de passos igual a 18 e ponto ótimo em $[0.08956, -0.712554]$.

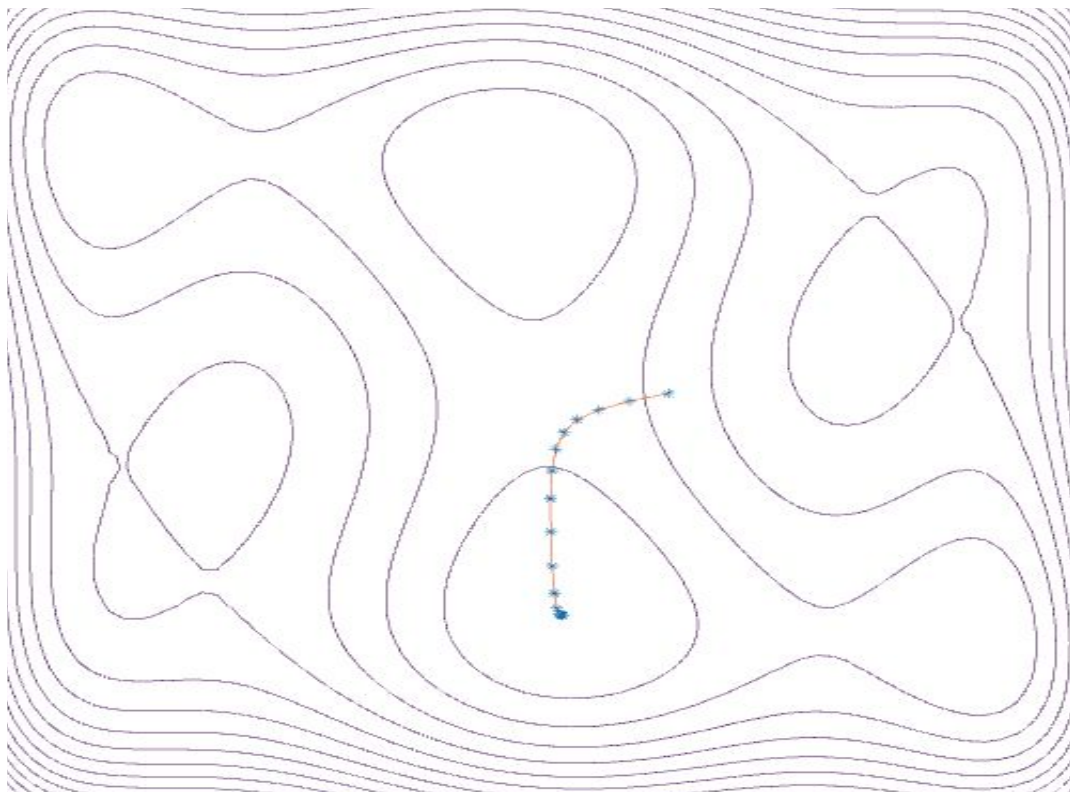


Figura 4: para o ponto 2, o valor de f no ponto final foi de -1.0316 com número de passos igual a 18 e ponto ótimo em $[-0.088956, 0.712554]$.

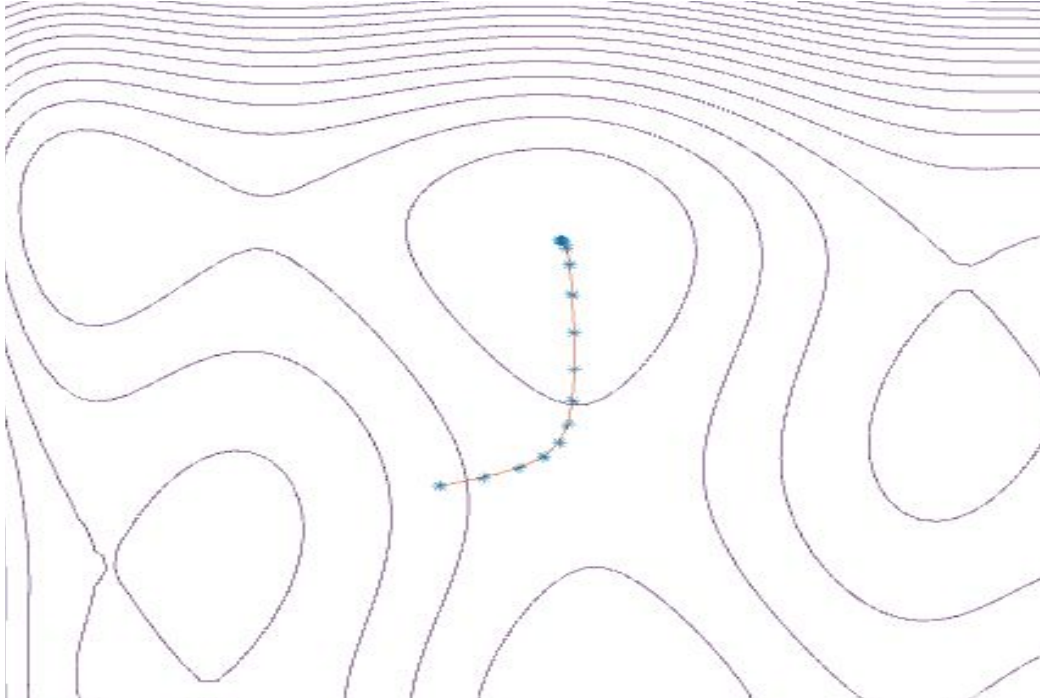
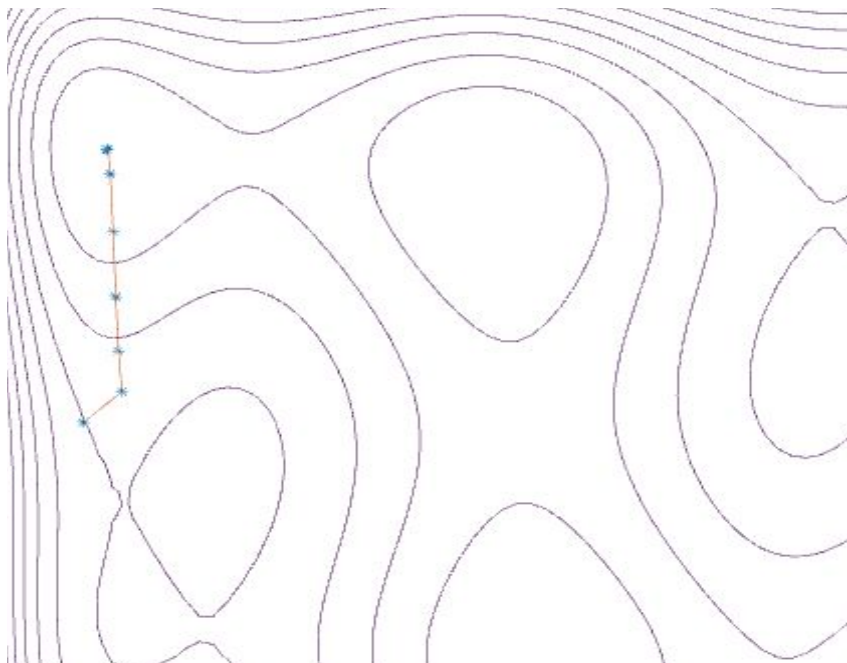


Figura 5: para o ponto 3, o valor de f no ponto final foi de -0.21546 com número de passos igual a 8 e ponto ótimo em $[-1.70359, 0.79607]$.



Preferi codificar a função de minimização. O código implementado em matlab/octave para a função de minimização está disponível no link https://drive.google.com/open?id=1t_VMNeHYBv4E4tx0mSFXrEzj3llenEkx. Deve ser chamada a função main2 sem argumentos.

A seguir, mostramos as tabelas de passo a passo de cada ponto.

Tabela 2: evolução do ponto 1 a cada iteração.

ponto 1	x	y	f(x,y)
1	0.5	0.0	0.87396
2	0.349375	-0.025	0.44634
3	0.228266	-0.052456	0.17982
4	0.144516	-0.084737	0.041865
5	0.092208	-0.79602	-0.039585
6	0.061922	-0.178552	-0.11921
7	0.04618	-0.248515	-0.23474
8	0.040175	-0.337852	-0.4118
9	0.04103	-0.444263	-0.64516
10	0.04686	-0.553872	-0.86784
11	0.055853	-0.641833	-0.99238
12	0.065677	-0.689837	-1.0258
13	0.074017	-0.706435	-1.0304
14	0.07992	-0.710672	-1.0312
15	0.083689	-0.711793	-1.0315
16	0.086049	-0.712191	-1.0316
17	0.087506	-0.712382	-1.0316
18	0.088956	-0.71249	-1.0316
19	0.088956	-0.712554	-1.0316

Conclusão: x^* é um ponto de mínimo global.

Tabela 3: evolução do ponto 2 a cada iteração.

ponto 2	x	y	f(x,y)
1	-0.5	0.0	0.87396
2	-0.349375	0.025	0.44634
3	-0.228266	0.052456	0.17982
4	-0.144516	0.084737	0.041865
5	-0.092208	0.79602	-0.039585
6	-0.061922	0.178552	-0.11921
7	-0.04618	0.248515	-0.23474
8	-0.040175	0.337852	-0.4118
9	-0.04103	0.444263	-0.64516
10	-0.04686	0.553872	-0.86784
11	-0.055853	0.641833	-0.99238
12	-0.065677	0.689837	-1.0258
13	-0.074017	0.706435	-1.0304
14	-0.07992	0.710672	-1.0312
15	-0.083689	0.711793	-1.0315
16	-0.086049	0.712191	-1.0316
17	-0.087506	0.712382	-1.0316
18	-0.088956	0.71249	-1.0316
19	-0.088956	0.712554	-1.0316
Conclusão: x^* é um ponto de mínimo global.			

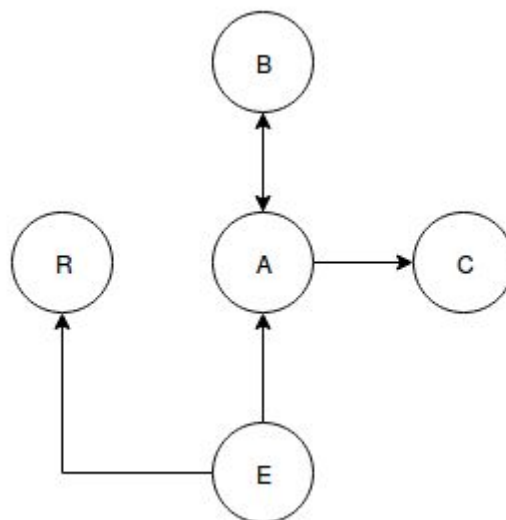
Tabela 4: evolução do ponto 3 a cada iteração.

ponto 3	x	y	f(x,y)
1	-1.8	0.0	2.2524
2	-1.63972	0.09	1.8729
3	-1.65468	0.20741	1.5431
4	-1.66555	0.36597	0.9782
5	-1.67646	0.55642	0.26643
6	-1.68838	0.725	-0.16223
7	-1.69872	0.79456	-0.21522
8	-1.70324	0.79602	-0.21546
9	-1.703539	0.79607	-0.21546
Conclusão: x^* é um ponto de mínimo local.			

2 Probabilidade

2.1 Bandido na casa se o vizinho ligou.

Figura 6: belief network para as variáveis do problema.



Analisando as variáveis, temos as seguintes cobertas de markov para cada variável.

Tabela 4: Markov-blanket para cada variável

Variáveis	Markov-Blanket
A	Pais: {B, E}, filhos: {C}, pais dos filhos: {}
B	Pais: {A}, filhos: {A}, pais dos filhos: {E}
E	Pais: {}, filhos: {A, R}, pais dos filhos: {B}
R	Pais: {E}, filhos: {}, pais dos filhos: {}
C	Pais: {A}, filhos: {}, pais dos filhos: {}

Dadas as condições iniciais do problema, o objetivo é encontrar a probabilidade condicional de $B = 1$ dado que $C = 1$ (a probabilidade de ter um bandido na casa dado que o vizinho ligou). Ou seja,

$$p(B = 1 \mid C = 1).$$

Tendo em vista que C só acontece quando A acontece ($C \neq A \rightarrow P(C) = 0$), temos que

$$p(C) = p(A)$$

logo,

$$p(B = 1 \mid C = 1) = p(B = 1 \mid A = 1) = p(A = 1, B = 1) / p(A = 1).$$

Temos ainda que

$$\begin{aligned} p(A = 1 \mid B = 1) &= 0.99 \text{ e } p(B = 1) = 0.001 \rightarrow \\ p(A = 1, B = 1) &= p(A = 1 \mid B = 1) p(B = 1) = \\ &= 0.99 * 0.001 = 0.00099 = 0.099\%. \end{aligned}$$

Logo, basta achar $p(A=1)$. Mas

$$p(A = 1) = 1 - p(A = 0).$$

Os dados de entrada úteis para resolver esse problema são

$$p(E = 0) = 1 - e = 0.999$$

$$p(E = 1) = e = 0.001$$

$$p(B = 0) = 1 - b = 0.999$$

$$p(B = 1) = b = 0.001$$

$$p(A = 1 \mid F = 1) = f = 0.001 \text{ (probabilidade de algum evento aleatório ativar alarme)}$$

$$p(A = 1 \mid B = 1) = ab = 0.99$$

$$p(A = 1 \mid E = 1) = ae = 0.01$$

Todas as descrições destas variáveis estão na descrição do trabalho. Para achar $p(A = 0)$, temos

$$p(A = 0) =$$

$$\begin{aligned}
&= \sum_{B,E} p(A=0, B, E) = \\
&= p(A=0 | B=0, E=0) p(B=0) p(E=0) + \\
& p(A=0 | B=1, E=0) p(B=1) p(E=0) + \\
& p(A=0 | B=0, E=1) p(B=0) p(E=1) + \\
& p(A=0 | B=1, E=1) p(B=1) p(E=1) \\
&= (1-f)(1-b)(1-e) + \\
& (1-f)(1-ab)(b)(1-e) + \\
& (1-f)(1-ae)(1-b)(e) + \\
& (1-f)(1-ab)(1-ae)(b)(e) \\
&= 0.999 * 0.999 * 0.999 + \\
& 0.999 * 0.01 * 0.001 * 0.999 + \\
& 0.999 * 0.99 * 0.999 * 0.001 + \\
& 0.999 * 0.01 * 0.99 * 0.001 * 0.001 \\
&= 0.998.
\end{aligned}$$

Como $p(A=0) = 0.998$, $p(A=1) = 0.002$. Portanto,

$$\begin{aligned}
p(B=1 | C=1) &= p(B=1 | A=1) = \\
p(A=1, B=1) / p(A=1) &= 0.00099 / 0.002 = 0.495 = 49.5\%.
\end{aligned}$$

•

Isto nos diz que é bastante provável que haja um bandido dentro da casa dela, dado que o alarme tocou.

2.2 Bandido na casa se a rádio noticiou terremoto.

Tendo em vista que R só acontece quando E acontece ($E \neq R \rightarrow P(R) = 0$), temos que

$$p(R) = p(E)$$

O raciocínio é semelhante ao anterior, o problema se resume a

$$\begin{aligned}
p(B=1 | C=1, R=1) &= p(B=1 | A=1, E=1) \\
&= p(B=1, E=1, A=1) / p(A=1 | E=1) p(E=1).
\end{aligned}$$

Resta encontrar $p(B=1, E=1, A=1)$.

Modo 1:

$$p(B=1, E=1, A=1) = p(B=1, E=1 | A=1) p(A=1)$$

$$= p(B = 1 | A = 1) p(E = 1 | A = 1) p(A = 1) \quad (B \text{ e } E \text{ são independentes})$$

Como

$$p(E = 1 | A = 1) = p(A = 1 | E = 1) p(E = 1) / p(A = 1) \rightarrow$$

$$p(B = 1, E = 1, A = 1) = p(B = 1 | A = 1) p(A = 1 | E = 1) p(E = 1)$$

$$p(B = 1, E = 1, A = 1) = 0.495 * 0.01 * 0.001 = 0.00000495$$

Modo 2: $p(B = 1, E = 1, A = 1) = p(B = 1 | E = 1, A = 1) p(A = 1 | E = 1) p(E = 1)$

$$= p(B = 1 | A = 1) p(A = 1 | E = 1) p(E = 1) \quad (B \text{ e } E \text{ são independentes})$$

$$= p(B = 1, E = 1, A = 1) = 0.495 * 0.01 * 0.001 = 0.00000495$$

Modo 3: $p(B = 1, E = 1, A = 1) = p(E = 1 | B = 1, A = 1) p(A = 1 | B = 1) p(B = 1)$

$$= p(E = 1 | A = 1) p(A = 1 | B = 1) p(B = 1) \quad (B \text{ e } E \text{ são independentes})$$

Falta achar $p(E = 1 | A = 1)$

$$p(E = 1 | A = 1) = p(A = 1 | E = 1) p(E = 1) / p(A = 1)$$

$$= 0.01 * 0.001 / 0.002 = 0.005$$

Daí,

$$= p(B = 1, E = 1, A = 1) = 0.005 * 0.99 * 0.001 = 0.00000495$$

Agora, como os três modos deram o mesmo valor, podemos confiar que as contas estão certas! Chegamos em

$$p(B = 1 | A = 1, E = 1) = p(B = 1, E = 1, A = 1) / p(A = 1 | E = 1) p(E = 1)$$

$$= 0.00000495 / (0.01 * 0.001) = 0.495$$

Isso faz (um certo) sentido pois $p(E = 1) = p(B = 1)$ e ainda são independentes, apesar de que se espere que essa probabilidade diminua dadas as notícias na rádio.