Formulário de ONL

Classificação de mínimos (e máximos)

$$\max f(x) = -\min(-f(x)) \qquad x^* = \underbrace{\arg\max\left(f\left(x\right)\right)}_{\text{maximizante}} = \underbrace{\arg\min\left(-f\left(x\right)\right)}_{\text{minimizante}}$$

Método de Davies, Swann e Campey - DSC

Dados: x_1, $\delta > 0$, $M < 1$, ε

1. procura no sentido positivo

A partir do x_1 e no sentido positivo, calcula-se uma sequência de pontos $x_2 = x_1 + \delta$, $x_3 = x_2 + 2\delta$, $x_4 = x_3 + 4\delta$, $\ldots$, $x_k = x_{k-1} + 2^{k-2}\delta$ até que no ponto x_k se tenha $f\left(x_k\right) > f\left(x_{k-1}\right)$.

- Calcula-se o ponto médio do último intervalo: $x_m = \frac{x_k + x_{k-1}}{2}$.
- Obter 4 pontos igualmente espaçados: $x_{k-2} < x_{k-1} < x_m < x_k$
- Para a aproximação quadrática, seleccionar três dos quatro pontos igualmente espaçados, comparando os valores de $f\left(x\right)$ nos dois pontos interiores do intervalo:
  - Se $f\left(x_{k-1}\right) \leq f\left(x_m\right)$ então escolhem-se os pontos x_{k-2}, x_{k-1} e x_m
  - Senão $f\left(x_{k-1}\right) > f\left(x_m\right)$ escolhem-se os pontos x_{k-1}, x_m e x_k
  - Redefinir e ordenar os pontos como $x_1 < x_2 < x_3$. Ir para **3. Minimizante da quadrática**.

2. procura no sentido negativo

Quando, a partir de x_1, o valor de $f\left(x_2\right) > f\left(x_1\right)$ (para $x_2 = x_1 + \delta$) a procura deve voltar-se para o sentido negativo, a começar novamente por x_1. O próximo ponto, na procura, é $x_{-1} = x_1 - \delta$.

- Se $f\left(x_{-1}\right) > f\left(x_1\right)$, então $[x_{-1}, x_2]$ contém o minimizante desejado. Os 3 pontos x_{-1}, x_1 e x_2 são e selecionados. Ir para **3. Minimizante da quadrática**.
- Se $f\left(x_{-1}\right) < f\left(x_1\right)$ então a procura deve continuar no sentido negativo, calculando $x_{-2} = x_{-1} - 2\delta$, $x_{-3} = x_{-2} - 4\delta$, $\ldots$, $x_{-k} = x_{-(k-1)} - 2^{k-1}\delta$, até que no ponto x_{-k} se tenha $f\left(x_{-k}\right) > f\left(x_{-(k-1)}\right)$.
- Calcula-se o ponto médio do último intervalo: $x_m = \frac{x_{-k} + x_{-(k-1)}}{2}$
- Obter 4 pontos igualmente espaçados: $x_{-k}, x_m, x_{-(k-1)}, x_{-(k-2)}$
  - Se $f\left(x_m\right) < f\left(x_{-(k-1)}\right)$ então escolhem-se os pontos x_{-k}, x_m e $x_{-(k-1)}$
  - Senão $\left(f\left(x_m\right) \geq f\left(x_{-(k-1)}\right)\right)$ escolhem-se os pontos $x_m, x_{-(k-1)}$ e $x_{-(k-2)}$
  - Redefinir e ordenar os pontos como $x_1 < x_2 < x_3$. Ir para **3. Minimizante da quadrática**.

3. Minimizante da quadrática

- Considerando $x_1 < x_2 < x_3$, determinar o minimizante da quadrática

$$x^*\left(q\right) = x_2 + \Delta \frac{f\left(x_1\right) - f\left(x_3\right)}{2\left(f\left(x_3\right) - 2f\left(x_2\right) + f\left(x_1\right)\right)}$$

  com $\Delta = (x_2 - x_1) = (x_3 - x_2)$.

Critério de paragem:

- Se $\Delta \leq \varepsilon$ então o processo iterativo termina, e $x^*\left(q\right)$ é a melhor aproximação à solução
- Senão o processo iterativo repete-se iniciando com $x_1 = x^*\left(q\right)$ e $\delta = M\delta$.

Condições de otimalidade

- **Condição necessária (e suficiente) de 1^a ordem**: $\nabla f(x) = 0$
- **Condições de 2^a ordem:**
  - $\nabla^2 f(x^*)$ é definida positiva $\Rightarrow x^*$ é um minimizante de $f(x)$ (condição suficiente).
  - $\nabla^2 f(x^*)$ é definida negativa $\Rightarrow x^*$ é um maximizante de $f(x)$ (condição suficiente).
  - $\nabla^2 f(x^*)$ é semi-definida positiva $\Rightarrow x^*$ é um minimizante ou ponto de sela de $f(x)$ (cond. necessária).
  - $\nabla^2 f(x^*)$ é semi-definida negativa $\Rightarrow x^*$ é um maximizante ou ponto de sela de $f(x)$ (cond. necessária).
  - $\nabla^2 f(x^*)$ é indefinida $\Rightarrow x^*$ é ponto de sela de $f(x)$ (condição suficiente).
- Uma matriz diz-se **definida positiva** se os determinantes das submatrizes principais são positivos
- Uma matriz diz-se **definida negativa** se os determinantes das submatrizes principais têm sinais alternados, sendo o determinante da primeira submatriz negativo
- Uma matriz diz-se **semi-definida positiva** se pelo menos um dos determinantes das submatrizes principais é zero e os outros são positivos
- Uma matriz diz-se **semi-definida negativa** se pelo menos um dos determinantes das submatrizes principais é zero e os outros têm sinais alternados, sendo o determinante da primeira submatriz negativo
- Uma matriz diz-se **indefinida** se os sinais dos determinantes das submatrizes principais não verificam nenhuma das 4 situações acima mencionadas.

Método Nelder-Mead

1. Seja $S_1 = \langle X_1, X_2, \ldots, X_{n+1} \rangle$ o simplex inicial já ordenado por ordem crescente dos valores da função
2. Calcular o **centróide** do simplex $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} X_i$
3. Calcular o **vértice refletido** $x_r = (1+\alpha)\bar{x} - \alpha X_{n+1}$ (com $\alpha = 1$)

   CASO 1: Se $(f(X_1) \leq f(x_r) < f(X_n))$ então x_r é bom $\Rightarrow$ aceitar $x_r \Rightarrow S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_r \rangle$

   CASO 2: Se $(f(x_r) < f(X_1))$ então x_r é muito bom $\Rightarrow$ fazer uma expansão do simplex:

   - calcular o **vértice expandido** $x_e = \gamma x_r + (1-\gamma)\,\bar{x}$ (com $\gamma = 2$)

     - Se $(f(x_e) < f(X_1))$ então x_e é muito bom $\Rightarrow$ aceitar $x_e \Rightarrow S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_e \rangle$
     - Senão aceitar $x_r \Rightarrow S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_r \rangle$

   CASO 3: Se $(f(X_n) \leq f(x_r) < f(X_{n+1}))$ então x_r é fraco $\Rightarrow$ fazer uma contracção para o exterior:

   - calcular o **vértice contraído para o exterior** $\hat{x}_c = \beta x_r + (1-\beta)\,\bar{x}$ (com $\beta = 0.5$)

     - Se $(f(\hat{x}_c) < f(X_n))$ então $\hat{x}_c$ é bom $\Rightarrow$ aceitar $\hat{x}_c \Rightarrow S_{k+1} = \langle X_1, X_2, \ldots, X_n, \hat{x}_c \rangle$
     - Senão **encolher o simplex** $\Rightarrow$ fazer $x_i = \dfrac{X_i + X_1}{2}$ $(i = 2, \ldots, n+1) \Rightarrow S_{k+1} = \langle X_1, x_2, \ldots, x_n, x_{n+1} \rangle$

   CASO 4: Se $(f(x_r) \geq f(X_{n+1}))$ então x_r é muito fraco $\Rightarrow$ fazer uma contracção para o interior:

   - calcular o **vértice contraído para o interior** $x_c = \beta X_{n+1} + (1-\beta)\,\bar{x}$ (com $\beta = 0.5$)

     - Se $(f(x_c) < f(X_n))$ então x_c é bom $\Rightarrow$ aceitar $x_c \Rightarrow S_{k+1} = \langle X_1, X_2, \ldots, X_n, x_c \rangle$
     - Senão **encolher o simplex** $\Rightarrow$ fazer $x_i = \dfrac{X_i + X_1}{2}$ $(i = 2, \ldots, n+1) \Rightarrow S_{k+1} = \langle X_1, x_2, \ldots, x_n, x_{n+1} \rangle$

4. **Critério de paragem**

   - Ordenar o simplex obtido. Se

     $$\frac{\max\limits_{2 \leq i \leq n+1} \|X_i - X_1\|_2}{\max\{1, \|X_1\|_2\}} \leq \varepsilon$$

     então o processo iterativo termina e X_1 é a a melhor aproximação à solução
   - Senão, o processo iterativo repete-se iniciando com o último simplex obtido (voltar a 2).

Métodos do gradiente

Algoritmo geral dos métodos do gradiente

Dados: aproximação inicial $x^{(1)}$, $\varepsilon > 0\,(\approx 0)$, $k \leftarrow 1$
Enquanto $\left\|\nabla f\left(x^{(k+1)}\right)\right\|_2 > \varepsilon$ **fazer**

- calcular $d^{(k)}$ (direção de procura através do Algoritmo de Segurança de Newton ou quasi-Newton)
- calcular $\alpha^{(k)}$ (comprimento do passo através do Critério de Armijo)
- calcular $x^{(k+1)} \leftarrow x^{(k)} + \alpha^{(k)} d^{(k)}$
- $k \leftarrow k+1$

Solução: $\begin{cases} x^* \approx x^{(k+1)} \\ f^* \approx f\left(x^{(k+1)}\right) \end{cases}$

Algoritmo do critério de Armijo para calcular $\alpha^{(k)}$

Dados $x^{(k)}$, $d^{(k)}$, $\nabla f\left(x^{(k)}\right)$, $f\left(x^{(k)}\right)$ e μ

1. $\alpha \leftarrow 1$
2. $\bar{x} \leftarrow x^{(k)} + \alpha d^{(k)}$
3. <u>se</u> $\left(f(\bar{x}) \leq f\left(x^{(k)}\right) + \mu\,\alpha\,\nabla f\left(x^{(k)}\right)^T d^{(k)}\right)$ <u>então</u> $\alpha^{(k)} \leftarrow \alpha$
   <u>senão</u>
   $\alpha \leftarrow \alpha/2$ e voltar a 2.

Algoritmo para o cálculo da <u>direção de Segurança de Newton</u>

Dados $x^{(k)}$ e η,
Resolver o sistema linear Newton $\nabla^2 f\left(x^{(k)}\right) d_N^{(k)} = -\nabla f\left(x^{(k)}\right)$ por EGPP
<u>se</u> (o sistema linear tem solução única - $\exists d_N^{(k)}$) <u>então</u>
  <u>se</u> $\left|\nabla f\left(x^{(k)}\right)^T d_N^{(k)}\right| \leq \eta$ (com $\eta > 0\,(\approx 0)$)
  <u>então</u> $d_{SN}^{(k)} \leftarrow -\nabla f\left(x^{(k)}\right) \Rightarrow$ direção é ortogonal ao gradiente
  <u>senão</u>
    <u>se</u> $\nabla f\left(x^{(k)}\right)^T d_N^{(k)} > \eta$ (com $\eta > 0\,(\approx 0)$ então $d_{SN}^{(k)} \leftarrow -d_N^{(k)} \Rightarrow$ direção é de subida
    <u>senão</u> $d_{SN}^{(k)} \leftarrow d_N^{(k)} \Rightarrow$ direção é de descida
<u>senão</u>
  $d_{SN}^{(k)} \leftarrow -\nabla f\left(x^{(k)}\right)$

Algoritmo para o cálculo da <u>direção quasi-Newton</u>

Dado $x^{(k)}$
Calcular $d_{QN}^{(k)} \leftarrow -H^{(k)}\,\nabla f\left(x^{(k)}\right)$, sendo $H^{(k)}$ dada por:
<u>se</u> $k = 1$ <u>então</u>
  $H^{(k)} \leftarrow I$
<u>senão</u>

$$\begin{cases} s^{(k-1)} \leftarrow x^{(k)} - x^{(k-1)} \\ y^{(k-1)} \leftarrow \nabla f\left(x^{(k)}\right) - \nabla f\left(x^{(k-1)}\right) \\ \text{atualizar } H^{(k)} \text{ pela fórmula } \textbf{DFP} \text{ ou } \textbf{BFGS} \end{cases}$$

  <u>se</u> $\nabla f\left(x^{(k)}\right)^T d_{QN}^{(k)} \geq 0$ <u>então</u> $d_{QN}^{(k)} \leftarrow -\nabla f\left(x^{(k)}\right) \Rightarrow$ direção é de subida

---

Fórmula DFP: $H^{(k)} = H^{(k-1)} - \dfrac{H^{(k-1)} y^{(k-1)} {y^{(k-1)}}^T H^{(k-1)}}{{y^{(k-1)}}^T H^{(k-1)} y^{(k-1)}} + \dfrac{s^{(k-1)} {s^{(k-1)}}^T}{{s^{(k-1)}}^T y^{(k-1)}}$

Fórmula BFGS: $H^{(k)} = \left(I - \dfrac{s^{(k-1)} {y^{(k-1)}}^T}{{s^{(k-1)}}^T y^{(k-1)}}\right) H^{(k-1)} \left(I - \dfrac{y^{(k-1)} {s^{(k-1)}}^T}{{s^{(k-1)}}^T y^{(k-1)}}\right) + \dfrac{s^{(k-1)} {s^{(k-1)}}^T}{{s^{(k-1)}}^T y^{(k-1)}}$
