Otimização e algoritmos

Bla,Bla, and bla

IST

December 12, 2020

Outline

Part 3

Parte 1 - Menor região envolvente - convergência

Pretendemos resolver

$$\begin{aligned} \underset{\mathcal{D} \in \mathbf{R}^2}{\text{maximize}} \quad & \mathrm{Area}(\mathcal{D}) \\ \text{subject to} \quad & (\mathbf{p_0} + t^* \mathbf{v}) \in \mathcal{D} \\ & ||\mathbf{p_0} + t_k \mathbf{v} - \mathbf{c_k}|| \leq R_k, \quad k = 1, ..., K \ , \end{aligned} \tag{1}$$

- É não-convexo
- É aparentemente intratável

Parte 1 - Menor região envolvente - convergência

Foi aproximado por N soluções de

$$
\begin{aligned}
&\underset{(\mathbf{p_0},v)\in\mathbf{R}^2\times\mathbf{R}^2}{\text{minimize}} && \begin{bmatrix}\cos\phi & \sin\phi\end{bmatrix}(\mathbf{p_0}+t^*\mathbf{v})\\
&\text{subject to} && ||\mathbf{p_0}+t_k\mathbf{v}-\mathbf{c_k}|| \leq R_k, \quad k=1,...,K.\\[1em]
&\underset{(\mathbf{p_0},\mathbf{v})\in\mathbf{R}^2\times\mathbf{R}^2}{\text{minimize}} && \begin{bmatrix}-\cos\phi & -\sin\phi\end{bmatrix}(\mathbf{p_0}+t^*\mathbf{v})\\
&\text{subject to} && ||\mathbf{p_0}+t_k\mathbf{v}-\mathbf{c_k}|| \leq R_k, \quad k=1,...,K,\\[1em]
&\underset{(\mathbf{p_0},v)\in\mathbf{R}^2\times\mathbf{R}^2}{\text{minimize}} && \begin{bmatrix}-\sin\phi & \cos\phi\end{bmatrix}(\mathbf{p_0}+t^*\mathbf{v})\\
&\text{subject to} && ||\mathbf{p_0}+t_k\mathbf{v}-\mathbf{c_k}|| \leq R_k, \quad k=1,...,K,\\[1em]
&\underset{(\mathbf{p_0},\mathbf{v})\in\mathbf{R}^2\times\mathbf{R}^2}{\text{minimize}} && \begin{bmatrix}\sin\phi & -\cos\phi\end{bmatrix}(\mathbf{p_0}+t^*\mathbf{v})\\
&\text{subject to} && ||\mathbf{p_0}+t_k\mathbf{v}-\mathbf{c_k}|| \leq R_k, \quad k=1,...,K.
\end{aligned}
\tag{2}
$$

- Relaxação convexa

Parte 1 - Menor região envolvente - convergência

Theorem

Considere-se N retângulos, $\mathcal{R}_1, \ldots, \mathcal{R}_N$, cada um obtido da resolução de 4 problemas de otimização convexos (2), *para* $\phi \in \Phi$ *com*

$$\Phi = \left\{ \phi \in \mathbb{R} : \phi = \frac{\pi}{2}((n-1)/N) \; ; n = 1, \ldots, N \right\} .$$

Então a interseção dos N retângulos, converge para a solução do problema não convexo (1) *à medida que* $N \to \infty$*, i.e.*

$$\lim_{N \to \infty} \cap_{i=1}^{N} \mathcal{R}_i = \mathcal{D} \; ,$$

onde $\mathcal{D}$ *é a solução de* (1)*.*

Parte 3 - Introdução

Nesta parte pretende-se resolver o seguinte problema

$$\underset{\mathbf{y}\in\mathbf{R}^{Nk}}{\text{minimize}} \quad f(\mathbf{y}),$$

onde

$$f(\mathbf{y}) := \sum_{m=1}^{N} \sum_{n=m+1}^{N} (||\mathbf{y_m} - \mathbf{y_n}||_2 - D_{mn})^2 = \sum_{m=1}^{N} \sum_{n=m+1}^{N} f_{mn}(\mathbf{y})^2 \, ,$$

e

$$D_{mn} = ||\mathbf{x_m} - \mathbf{x_n}||_2 \, .$$

- Não é um problema de otimização convexo!
- Será resolvido usando o método Levenberg-Marquardt (LM)

Part 3 - Task 2

Definindo $\mathbf{y_{m-n}} := \mathbf{y_m} - \mathbf{y_n}$ podemos escrever

$$f_{mn}(\mathbf{y}) := ||\mathbf{y_{m-n}}|| - D_{mn}$$

e

$$D_{\mathbf{y}} f_{mn}(\mathbf{y}) = \frac{\left[\mathbf{0}_{1\times(m-1)k}\ \mathbf{y_{m-n}}^T\ \mathbf{0}_{1\times(n-m-1)k}\ \ -\mathbf{y_{m-n}}^T\ \mathbf{0}_{1\times(N-n)k}\right]}{||\mathbf{y_{m-n}}||}\ .$$

- Tanto $f_{mn}(\mathbf{y})$ como $D_{\mathbf{y}} f_{mn}(\mathbf{y})$ são escritos em função de cada $\mathbf{y_{m-n}}$ e da sua norma.

Part 3 - Task 2

A função de otimização é

$$f(\mathbf{y}) = \sum_{m=1}^{N} \sum_{n=m+1}^{N} f_{mn}(\mathbf{y})^2 ,$$

e o seu jacobiano é dado por

$$\begin{aligned} D_{\mathbf{y}} f(\mathbf{y}) &= \sum_{m=1}^{N} \sum_{n=m+1}^{N} D_u(u^2)\Big|_{u=f_{mn}(\mathbf{y})} D_{\mathbf{y}} f_{mn}(\mathbf{y}) \\ &= \sum_{m=1}^{N} \sum_{n=m+1}^{N} 2 f_{mn}(\mathbf{y}) D_{\mathbf{y}} f_{mn}(\mathbf{y}) . \end{aligned}$$

- Tanto $f(\mathbf{y})$ como $D_{\mathbf{y}} f(\mathbf{y})$ são escritos em função de cada $f_{mn}(\mathbf{y})$ e $D_{\mathbf{y}} f_{mn}(\mathbf{y})$.

Part 3 - Task 2

Para cada iteração do método LM é necessário calcular **A** e **b** dados por

$$\mathbf{A} := \begin{bmatrix} D_{\mathbf{y}} f_{1,1}(\mathbf{y}) \\ D_{\mathbf{y}} f_{1,2}(\mathbf{y}) \\ \vdots \\ D_{\mathbf{y}} f_{N-1,N}(\mathbf{y}) \\ \sqrt{\lambda}\mathbf{I}_{Nk\times Nk} \end{bmatrix} \quad \text{e} \quad \mathbf{b} := \begin{bmatrix} D_{\mathbf{y}} f_{1,1}(\mathbf{y})\mathbf{y} - f_{1,1}(\mathbf{y}) \\ D_{\mathbf{y}} f_{1,2}(\mathbf{y})\mathbf{y} - f_{1,2}(\mathbf{y}) \\ \vdots \\ D_{\mathbf{y}} f_{N-1,N}(\mathbf{y})\mathbf{y} - f_{N-1,N}(\mathbf{y}) \\ \sqrt{\lambda}\mathbf{y} \end{bmatrix} .$$

Note-se que

$$D_{\mathbf{y}} f_{mn}(\mathbf{y})\mathbf{y} - f_{mn}(\mathbf{y}) = \mathbf{y_{m-n}}(\mathbf{y_m} - \mathbf{y_n}) - ||\mathbf{y_{m-n}}|| + D_{mn} = D_{mn} ,$$

logo

$$\mathbf{b} = \begin{bmatrix} D_{1,1} & D_{1,2} & \dots & D_{N-1,N} & \sqrt{\lambda}\mathbf{y} \end{bmatrix} .$$

- Fração considerável das entradas de **b** é constante, pelo que pode ser calculada uma única vez.

Part 3 - Task 2

Em cada iteração do método LM é necessário calcular $\mathbf{f}(\mathbf{y})$,$||D_\mathbf{y} f(\mathbf{y})||$, $\mathbf{A}$, e $\mathbf{b}$ $\rightarrow$ desenvolvida a função:

$$[f(\mathbf{y}), ||D_\mathbf{y} f(\mathbf{y})||, \mathbf{A}, \mathbf{b}] = \mathit{objectiveF(y)}\,.$$

- Calculadas apenas parte de $\mathbf{A}$ e $\mathbf{b}$ independentes de λ;
- $f(\mathbf{y}), ||D_\mathbf{y} f(\mathbf{y})||, \mathbf{A}$, e $\mathbf{b}$ calculados simultaneamente;
- As quantidades $\mathbf{f}(\mathbf{y})$,$||D_\mathbf{y} f(\mathbf{y})||$, e $\mathbf{A}$ são calculadas iterativamente para todos os pares

  $$(m, n) : m \in \{1, \ldots, N-1\}, n \in \{m+1, \ldots, N\}\,,$$

  prefazendo um total de $N(N/2-1)$ iterações;
- $\mathbf{b}$ é calculado apenas na primeira instância desta função.

Part 3 - Task 2

Para a i-th iteração (m, n):

- $\mathbf{y_{m-n}} \leftarrow \mathbf{y_m} - \mathbf{y_n}$
- $||\mathbf{y_{m-n}}|| \leftarrow \sqrt{\mathbf{y_{m-n}}^T \mathbf{y_{m-n}}}$
- $f_{mn}(\mathbf{y}) \leftarrow ||\mathbf{y_{m-n}}|| - \mathbf{b}(i)$
- $D_{\mathbf{y}} f_{mn}(\mathbf{y}) \leftarrow \frac{\left[\mathbf{0}_{1\times(m-1)k} \; \mathbf{y_{m-n}}^T \; \mathbf{0}_{1\times(n-m-1)k} \; -\mathbf{y_{m-n}}^T \; \mathbf{0}_{1\times(N-n)k}\right]}{||\mathbf{y_{m-n}}||}$
- $f(\mathbf{y}) \leftarrow f(\mathbf{y}) + f_{mn}(\mathbf{y})^2$
- $D_{\mathbf{y}} f(\mathbf{y}) \leftarrow D_{\mathbf{y}} f(\mathbf{y}) + 2 f_{mn}(\mathbf{y}) D_{\mathbf{y}} f_{mn}(\mathbf{y})$
- $\mathrm{row}_i(\mathbf{A}) \leftarrow D_{\mathbf{y}} f_{mn}(\mathbf{y})$

Melhoramento de 2 ordens de grandeza no esforço computacional em relação a uma primeira implementação ingénua.

Part 3 - Task 3

Implementado o algoritmo LM de forma genérica:

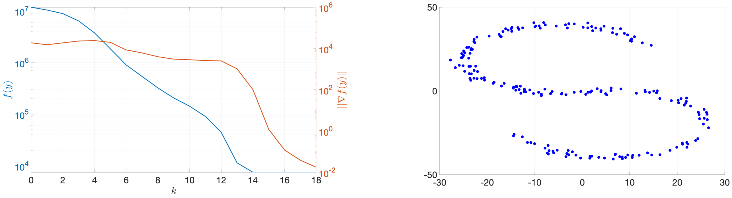

Figure 1: $k = 2$.

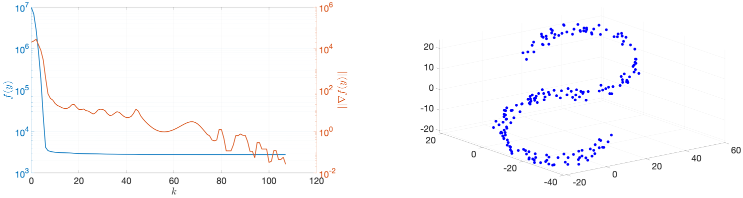

Figure 2: $k = 3$.

Part 3 - Task 4

Não é dada inicialização:

- Resolver o problema com LM para vários $\mathbf{y}_0$ aleatórios
- Cada LM pode ser calculado em paralelo
- Escolhe-se a melhor solução

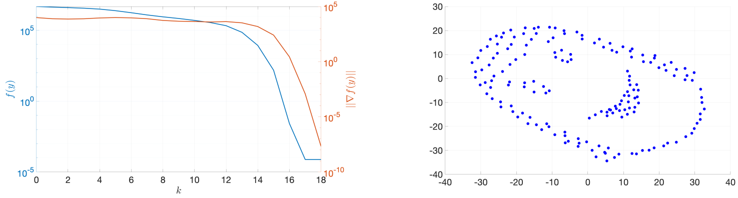

Figure 3: Melhor solução em 24 inicializações distintas para $k = 2$.

Part 3 - Task 4

Verificou-se que:

- Para todas as 24 inicializações obtém-se o mesmo valor de f
- As verificam $f(\mathbf{y}_{sol}) = 7.6430 \times 10^{-5}$
- A função de otimização é não-negativa
- Se se reduzir ϵ, $f(\mathbf{y}_{sol})$ aproxima-se de 0
- Na prática, pode ser considerado um mínimo global ($f(\mathbf{y}_{sol}) << D_{ij}$)

Part 3 - Task 4 - Unicidade da solução

Seja

$$\bar{\mathbf{y}} = \mathrm{col}(\mathcal{T}\mathbf{y_1} + \mathbf{w}, \ldots, \mathcal{T}\mathbf{y_N} + \mathbf{w}) \, ,$$

onde $\mathcal{T} \in \mathbb{R}^{k \times k}$ é uma matriz de rotação e $\mathbf{w} \in \mathbb{R}^k$. É evidente que

$$||\bar{\mathbf{y}}_\mathbf{m} - \bar{\mathbf{y}}_\mathbf{n}|| = ||\mathbf{y_m} - \mathbf{y_n}||$$

para qualquer par $(m, n) : m \in \{1, \ldots, N\}, n \in \{1, \ldots, N\}$. Logo

$$f(\bar{\mathbf{y}}) = f(\mathbf{y}) \, .$$

- Se for encontrada uma solução, então existe uma infinidade de soluções.
- Logo, a solução encontrada não é única.

Part 3 - Task 4 - Unicidade da solução

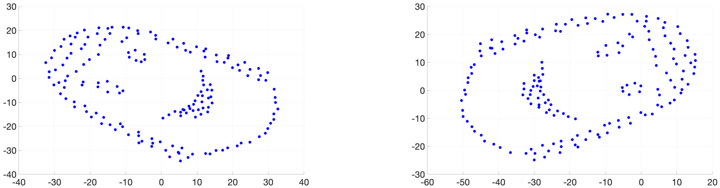

Figure 4: Duas melhores soluções em 24 inicializações distintas para $k = 2$.

- Uma solução pode ser obtida da outra por via de uma rotação e translação de cada $\mathbf{y_i}$, $i \in \{1, \ldots, N\}$.

Part 3 - Task 4 - Unicidade da solução

- Será que todas as soluções são da forma

$$\mathbf{\bar{y}} = \mathrm{col}(\mathcal{T}\mathbf{y_1} + \mathbf{w}, \ldots, \mathcal{T}\mathbf{y_N} + \mathbf{w})\ ?$$

- Não! A reflexão de cada $\mathbf{y_i}$, $i \in \{1, \ldots, N\}$ em relação ao eixo da primeira coordenada origina um valor igual da função de custo.

Part 3 - Task 4 - cMDS

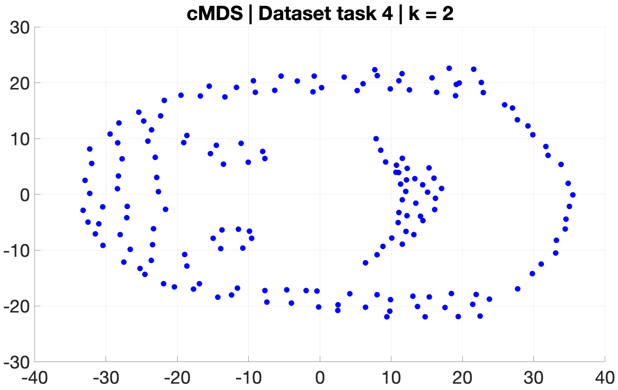

Figure 5: Solução método cMDS.

- A solução corresponde a uma translação e rotação das duas soluções anteriores.