Solução Analítica para o Problema de Rotação dos Eixos de Representação dos Partidos no Radar Parlamentar

Saulo Trento

19 de agosto de 2012

Resumo

Uma função objetivo análoga à energia é proposta para o problema de rotação dos eixos de representação entre períodos consecutivos no Radar Parlamentar [1], e uma solução analítica é deduzida.

1 Introdução

Para dois períodos consecutivos de atividade de uma câmara legislativa, o Radar Parlamentar [1] faz duas análises de componentes principais (PCA) separadas, que permitem visualização em duas dimensões dos partidos representados por círculos (de tamanhos proporcionais ao número de cadeiras que ocupam). Nada garante que os dois resultados colocarão um mesmo partido em posições próximas no plano para os dois períodos, mas espera-se que as posições relativas entre os partidos mude pouco. Para melhor visualizar as mudanças nestas posições relativas é conveniente rotacionar o resultado de uma das análises em torno da origem, de tal forma que seja minimizada alguma medida das "distâncias" percorridas pelos partidos entre os períodos. A figura 1 ilustra esta idéia, em que o resultado da PCA de 2011 foi rotacionado de $\theta = 160°$.

2 Problema de otimização

Está claro que temos em mãos um problema de otimização, em que buscamos o ângulo θ que minimiza uma função objetivo $E(\theta)$ que dê uma medida das distâncias percorridas pelos partidos no plano.

Função Objetivo

É razoável esperar que partidos com mais parlamentares apresentem uma maior consistência de opiniões entre períodos, já que para que um partido com apenas um parlamentar altere radicalmente suas opiniões basta que o seu representante único o faça: assim associamos um peso M_k a cada partido k igual ao número de parlamentares que tem na casa.

Também é razoável penalizar mais as distâncias muito grandes, de forma que a função objetivo que usaremos será proporcional à soma das distâncias quadráticas, D_k^2, percorridas por cada partido k, ponderados pelos M_k.Temos portanto:

$$E(\theta) = \sum_{k \in \{partidos\}} M_k \cdot D_k^2 \tag{1}$$

Ou seja:

$$E(\theta) = \sum_{k \in \{partidos\}} M_k \cdot \left[\left(x_{k1} - x_{k2}(\theta)\right)^2 + \left(y_{k1} - y_{k2}(\theta)\right)^2 \right] \tag{2}$$

Onde os índices 1 e 2 referem-se aos dois períodos consecutivos, e o período 2 é o que está sendo rotacionado de θ. x e y referem-se às coordenadas cartesianas.

Note-se a analogia desta função objetivo com a definição de energia: para um conjunto de bolas k massa M_k que iriam de (x_{k1}, y_{k1}) a (x_{k2}, y_{k2}) em um tempo Δt (o tempo da animação), estamos buscando o ângulo θ que minimiza a energia envolvida neste movimento. Um enunciado anterior para o mesmo problema de otimização utilizava a distância (não quadrática), para uma analogia com a quantidade de movimento (e não com a energia), o que tem mais chance de resultar em partidos "caminhando longas distâncias".

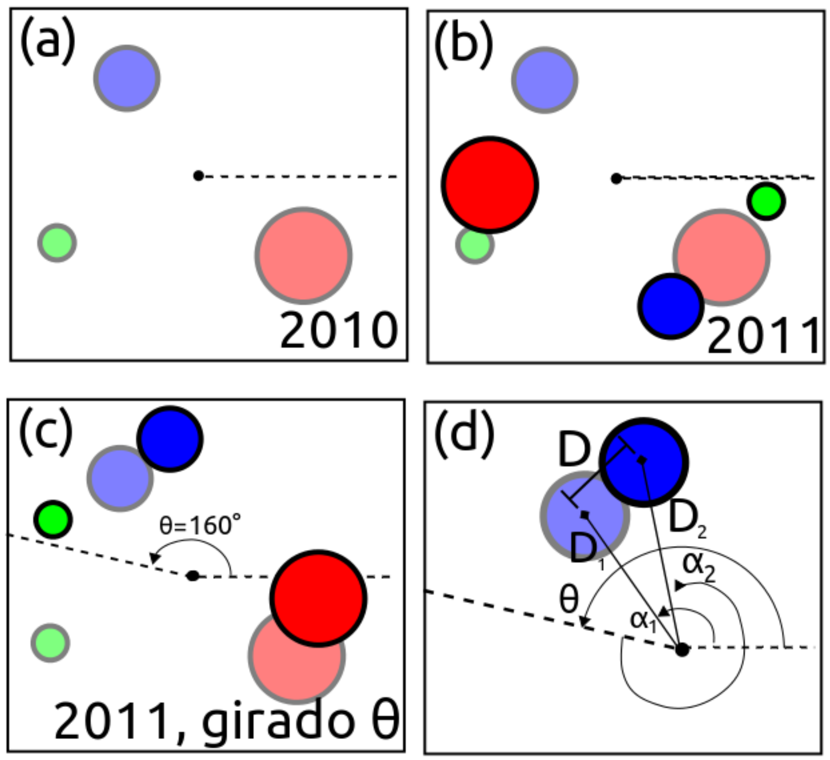

Figura 1: (a) Um resultado hipotético da análise de componentes principais para o ano de 2010 envolvendo três partidos. (b) Superposto ao primeiro, um resultado para o ano seguinte. (c) O mesmo resultado anterior, porém rotacionado de 160° para minimizar a movimentação dos partidos na animação entre um ano e outro. (d) Detalhe para um dos partidos.

3 Resolução Analítica

Representando as coordenadas na forma polar, conforme a notação da figura 1(d), definimos D_{k1}, D_{k2}, α_{k1} e α_{k2} como segue:

$$\begin{aligned} x_{k1} &= D_{k1} \cdot \cos(\alpha_{k1}) \\ y_{k1} &= D_{k1} \cdot \sin(\alpha_{k1}) \\ x_{k2} &= D_{k2} \cdot \cos(\theta + \alpha_{k2}) \\ y_{k2} &= D_{k2} \cdot \sin(\theta + \alpha_{k2}) \end{aligned} \tag{3}$$

Pela lei dos cossenos temos que a distância quadrática D_k^2 percorrida pelo partido k é dada por

$$D_k^2 = D_{k1}^2 + D_{k2}^2 - 2D_{k1}D_{k2}\cos(\alpha_{k1} - \alpha_{k2} - \theta)$$

Fazendo $\alpha_k = \alpha_{k1} - \alpha_{k2}$:

$$D_k^2 = D_{k1}^2 + D_{k2}^2 - 2D_{k1}D_{k2}\cos(\alpha_k - \theta)$$

E a função objetivo $E(\theta)$ pode ser reescrita na forma:

$$E(\theta) = \sum_{k \in \{partidos\}} M_k \left[D_{k1}^2 + D_{k2}^2 - 2D_{k1}D_{k2}\cos(\alpha_k - \theta)\right] \tag{4}$$

Usando a fórmula do arco duplo, temos $\cos(\alpha_k - \theta) = \cos\alpha_k \cos\theta + \sin\alpha_k \sin\theta$. Estamos procurando um mínimo, então vamos derivar $E(\theta)$, para depois igualar a derivada a zero:

$$\frac{\mathrm{d}E(\theta)}{\mathrm{d}\theta} = \sum_k \left[(2M_k D_{k1} D_{k2} \cos\alpha_k)\sin\theta + (2M_k D_{k1} D_{k2} \sin\alpha_k)\cos\theta\right]$$

O seno e cosseno de θ não dependem de k, logo podem ser colocados em evidência antes da somatória:

$$\frac{\mathrm{d}E(\theta)}{\mathrm{d}\theta} = \sin\theta \sum_k 2M_k D_{k1} D_{k2} \cos\alpha_k + \cos\theta \sum_k 2M_k D_{k1} D_{k2} \sin\alpha_k$$

Igualando a derivada a zero, dividindo tudo por $\cos\theta$ (desde que $\theta \neq \pm\frac{\pi}{2}$), e isolando a incógnita:

$$\tan\theta = \frac{\sum_k M_k D_{k1} D_{k2} \sin\alpha_k}{\sum_k M_k D_{k1} D_{k2} \cos\alpha_k}$$

$$\boxed{\theta = \arctan \frac{\sum_k M_k D_{k1} D_{k2} \sin\alpha_k}{\sum_k M_k D_{k1} D_{k2} \cos\alpha_k} + C\pi} \tag{5}$$

onde $C = 0$ ou 1, um dos casos correspondendo ao mínimo e o outro ao máximo.

Para encontrar a solução basta calcular $E(\theta)$ para os θ que satisfazem (5) e verificar qual delas corresponde ao mínimo. Se durante o cálculo o denominador do argumento do arco-tangente for nulo, então o mínimo está em $\frac{\pi}{2}$ ou $\frac{3\pi}{2}$.

O mesmo procedimento deve ser repetido espelhando-se um dos eixos (por exemplo multiplicando todas as coordenadas x por -1), já que o sentido dos eixos que resulta da PCA é arbitrário. Assim o problema se resume à avaliação de $E(\theta)$ em 4 casos (espelhado e não-espelhado, cada um com dois valores de θ), e escolha do caso que resultar no mínimo.

4 Conclusão

A nova função objetivo, análoga à energia e não à quantidade de movimento, deve minimizar situações em que um partido "caminha longas distâncias" entre um período e o seguinte. Além disso, com este enunciado o problema de otimização tem agora solução analítica, permitindo mais precisão na solução para uma quantidade fixa de recursos computacionais.

Referências

[1] http://radarparlamentar.polignu.org