

2

Mínimos Quadrados Móveis

Mínimos quadrados móveis (*moving least-squares*, em inglês) são utilizados neste trabalho para diversas finalidades: na redução do ruído dos pontos de amostragem, na geração de novos pontos de amostragem e na obtenção dos vértices da curva linear por partes obtida na reconstrução.

A técnica de mínimos quadrados móveis tem como objetivo aproximar um conjunto de pontos de amostragem através de funções polinomiais. A utilização de referências locais e a independência de malha são as maiores vantagens da técnica de mínimos quadrados móveis.

Quando queremos interpolar ou aproximar um conjunto de pontos esparsos, é praticamente impossível determinar uma referência global ou paramétrica que possa ser utilizada como base para a aproximação. Neste caso, o usual é utilizar uma malha que servirá de referência para o algoritmo de interpolação ou aproximação. A malha pode ser uma curva (ou superfície) linear por partes, onde cada segmento da malha serve como referência para uma partição da curva, e a aproximação obtida é a união dos 'pedaços de superfície definidos sobre cada um dos segmentos da malha de referência. O problema desta técnica é que, na maioria dos casos, a curva obtida depende da malha referência, ou seja, se duas malhas diferentes forem utilizadas como referência, a curva obtida com a primeira malha será diferente da curva obtida com a segunda malha.

A idéia principal do MLS é avaliar a projeção de um ponto r sobre uma variedade C (neste trabalho, a variedade C será uma curva) a partir de pontos de amostragem de C . Uma aproximação para C é obtida a partir da união das projeções de diversos pontos próximos ao conjunto de pontos de amostragem. Os parâmetros de entrada do algoritmo podem ser ajustados de modo que se obtenha uma interpolação dos pontos de amostragem. O processo de interpolação é utilizado quando não existe ruído nos pontos de amostragem, e a aproximação é utilizada quando há ruído (neste caso, não se deseja que a curva de aproximação passe pelos pontos de amostragem, mas próximo a eles). Note que a avaliação da projeção deve ser feita a partir dos pontos de amostragem, pois não sabemos de antemão quem é a curva de onde os pontos

foram amostrados.

A projeção de um ponto r sobre uma curva de amostragem C é realizada em duas etapas: primeiro, avalia-se uma referência local para a curva nas proximidades do ponto r . A seguir, estima-se uma função de mapeamento local nas proximidades de r . A função, na maioria dos casos, é um polinômio de grau baixo. Os procedimentos que calculam uma referência e uma função de mapeamento local para um determinado ponto no espaço baseiam-se nas idéias de mínimos quadrados móveis, técnica proposta por McLain em 1976 [13].

2.1

Mínimos Quadrados Móveis

Seja Π_m^d o espaço dos polinômios de grau m em $\mathbb{R}^d$, $\{x_i\}_{i \in I}$ um conjunto de pontos em $\mathbb{R}^d$, e $\{f(x_i)\}_{i \in I}$ o conjunto dos valores associados a esses pontos. A aproximação mínima quadrática flutuante de grau m para um ponto $x \in \mathbb{R}^d$ é o valor $\tilde{p}(x)$, onde $\tilde{p} \in \Pi_m^d$ minimiza, entre todos os valores de $p \in \Pi_m^d$, o erro mínimo quadrático com pesos, definido por:

$$\sum_{i \in I} (p(x_i) - f(x_i))^2 \theta(\|x - x_i\|) \quad (2-1)$$

onde

- $\|\cdot\|$ é a distância Euclideana em $\mathbb{R}^d$
- θ é uma função-peso não negativa

Como estamos buscando uma aproximação local para a função f , queremos que pontos de amostragem mais próximos de x possuam maior influência na minimização que pontos mais distantes. Esta localidade é obtida através da escolha da função-peso $\theta(s)$. O ideal é que a função-peso $\theta(s)$ seja uma função suave que decaia rapidamente quando s tende para o infinito, e tenha o seu máximo em $s = 0$. Se $\lim_{s \rightarrow 0} \theta(s) = \infty$, então a interpolação dos pontos de amostragem é alcançada.

A figura 2.1 mostra o processo de aproximação de funções através de mínimos quadrados móveis. Nos exemplos mostrados na figura 2.1, deseja-se aproximar o valor de $\log(10)$ a partir dos pontos de amostragem $x_i \in \{0.01, 1, 2, 3, 4, \dots, 19, 20\}$ e através de um polinômio de primeiro grau (uma reta). O objetivo do exemplo é mostrar como as funções-peso são importantes na determinação da localidade do MLS.

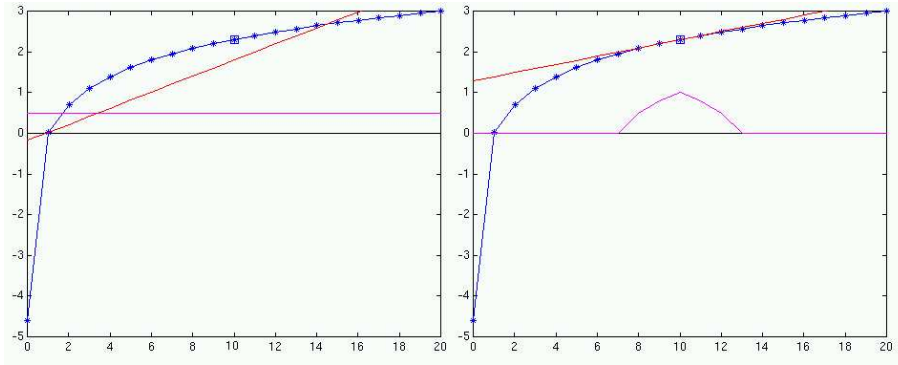


Figure 2.1: MLS com pesos

A curva em azul representa $f(x)$ (a curva que se quer aproximar). Os pontos azuis são os pontos de amostragem utilizados no processo. O ponto $x = 10$ (com um quadrado em volta) é o ponto para o qual deseja-se estimar o valor de $f(x)$. A função em rosa é a função-peso. Note que a função peso utilizada na figura da direita é uma função constante, e a função-peso da figura da esquerda é nula em todos os pontos, exceto para os valores de x entre 7 e 11.

A reta vermelha é o polinômio de primeiro grau $\tilde{p}(x)$ que minimiza o somatório das distâncias entre a função $\log(x)$ e o polinômio $p(x)$, balanceado pela função-peso θ , em cada valor de x_i . Observe que a reta de aproximação obtida na figura da direita aproxima melhor a função nas proximidades de $x = 10$ do que na figura da esquerda. De fato, na figura da direita, a reta obtida é bastante próxima à reta tangente. Isso aconteceu porque o processo de aproximação mostrado na figura da direita só utilizou os cinco pontos de amostragem mais próximos do ponto $x = 10$, sendo que a função-peso, neste caso, decai na medida em que o ponto de amostragem se distancia do ponto $x = 10$. Já no exemplo ilustrado pela figura da esquerda, todos os pontos de amostragem receberam o mesmo peso, resultando em uma distância maior entre a função original e o polinômio de aproximação no ponto $x = 10$.

2.2

Procedimento MLS Básico

Levin mostra em [12] como a técnica de aproximação de funções através de mínimos quadrados móveis pode ser aplicada na aproximação de superfícies de dimensão $d - 1$ a partir de pontos esparsos em $\mathbb{R}^d$. A diferença entre a aproximação de funções e a aproximação de superfícies é que, no primeiro caso, existe um domínio global para toda a função, o que não existe no segundo caso.

A essência do problema é a mesma: dado um ponto $r \in \mathbb{R}^d$ próximo ao conjunto de pontos de amostragem, deseja-se achar um polinômio $\tilde{p}_r(x) : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$, de grau m , que aproxime a superfície nas proximidades de um ponto r . Porém, ao contrário do caso de aproximação de funções, não existe uma referência global que sirva para a aproximação da superfície nas proximidades de qualquer ponto r próximo ao conjunto de pontos de amostragem. Então, antes de se estimar o polinômio, deve-se achar um sistema de referência para o mesmo. A seguir, é mostrado como se obter o hiperplano de referência, bem como obter o polinômio a partir do hiperplano. Como o presente trabalho tem como tema a reconstrução de curvas, será mostrado o caso de aproximação de curvas de dimensão 1 em planos de duas dimensões. Mas é importante lembrar que o procedimento MLS básico proposto por Levin é válido para qualquer dimensão.

Seja $\{x_i\}_{i \in I} \subset \mathbb{R}^2$ um conjunto de pontos amostrados de uma curva C , também em $\mathbb{R}^2$. O objetivo do procedimento MLS básico é definir um operador $\tilde{P}(p)$ sobre um ponto p qualquer de tal modo que se p é um ponto suficientemente próximo à curva C , então $\tilde{P}(p)$ está ainda mais próximo da curva. Este operador não define uma projeção, como será visto mais adiante, pois $\tilde{P}(p)$ não necessariamente coincide com $\tilde{P}(\tilde{P}(p))$. O procedimento MLS básico é realizado através das seguintes etapas:

1. Ache a reta H_r que aproxima a reta tangente à curva C nas proximidades de r .
2. Projete o ponto r na reta H_r , obtendo o ponto h . Este ponto será a origem do sistema ortonormal de coordenadas utilizado como referência para o resto do processo. O eixo de x do sistema de coordenadas será a reta obtida na etapa anterior, e o eixo de y será a reta ortogonal a H_r passando pela origem h .
3. Ache o polinômio $\tilde{p}(x)$ de grau m que aproxima a curva C nas proximidades de r , utilizando o sistema de coordenadas obtido na etapa anterior.
4. O ponto $\tilde{P}(p) = (0, \tilde{p}(0))$ aproxima o ponto p à curva C . Note que $\tilde{P}(p)$ está definido no sistema de coordenadas obtido anteriormente.

Um aspecto importante a se notar neste procedimento de aproximação é que, além de $\tilde{P}(p)$ (que é uma aproximação para a projeção de p sobre a curva C), este processo nos fornece uma aproximação para a curva nas proximidades de $\tilde{P}(p)$. Para obter esta aproximação, basta calcularmos pontos $(y, \tilde{p}(y))$, onde y é um valor real suficientemente próximo de zero.

Cálculo da reta tangente aproximada

A reta $H_r = \{x : \langle a, x \rangle - D = 0, x \in \mathbb{R}^2\}$, $a \in \mathbb{R}^2$, $\|a\| = 1$ que aproxima a reta tangente à curva C nas proximidades do ponto r é dada pelos valores de a e D que minimizam a seguinte expressão:

$$\sum_{i \in I} (\langle a, r_i \rangle - D)^2 \theta(\|r - r_i\|)$$

onde

- a é o vetor normal unitário à reta H_r
- D é a distância entre a reta H_r e a origem
- $\{r_i\}_{i \in I}$ é o conjunto de pontos de amostragem de C
- $\langle \cdot, \cdot \rangle$ é o produto interno padrão em $\mathbb{R}^2$
- $\langle a, r_i \rangle - D$ é a distância entre o ponto r_i e a reta H_r

Note que esta expressão é bastante similar à equação utilizada no método de mínimos quadrados móveis para aproximação de funções. Porém, como queremos achar uma reta, não precisamos de uma referência. Outro ponto a se observar é que vários valores de a e D podem minimizar a equação acima. Levin sugere que se escolha a reta que passa mais próxima do ponto r , ou seja, os valores de a e D para os quais $\|\langle a, r_i \rangle - D\|$ é mínimo.

Cálculo do polinômio

Seja q a projeção de p sobre H_r :

$$q = p - [\langle a, p \rangle - D]a$$

O sistema de coordenadas utilizado para o cálculo do polinômio $\tilde{p}$ tem origem em q , e seus eixos de x e y são, respectivamente, a reta H_r e a reta perpendicular a H_r que passa pelo ponto q .

Sejam $\{x_i\}_{i \in I}$ as projeções ortogonais dos pontos $\{r_i\}_{i \in I}$ sobre a reta H_r , representadas no sistema de coordenadas definido acima, e seja f_i a distância entre o ponto f_i e a reta H_r . Observe que temos agora um conjunto de pontos de amostragem definidos no eixo de x do sistema de coordenadas, e um valor associado a cada um destes pontos de amostragem. O problema de achar um polinômio que aproxima a curva C nas proximidades de r é bastante similar ao problema de aproximação de funções através de mínimos quadrados móveis, visto anteriormente. O polinômio $\tilde{p}$ de grau m que aproxima a curva nas proximidades de p é o polinômio p que minimiza a expressão:

$$\sum_{i \in I} (p(x_i) - f_i)^2 \theta(\|r - r_i\|)$$

A diferença desta expressão para a expressão (2-1) é que, no caso de aproximação de funções, o peso atribuído a cada um dos pontos de amostragem é dado pela distância entre x e x_i . Já no caso de aproximação da curva, o peso atribuído a cada um dos pontos de amostragem é dado pela distância entre os pontos r e r_i , que “traduzidos” para o sistema de coordenadas local, podem ser escrito como os pontos $(0, f(0))$ e (x_i, f_i) .

O valor $\tilde{p}(0)$ aproxima a altura da curva C sobre a reta H_r na origem. Daí, o ponto $\tilde{r} = q + \tilde{p}(0)a$ aproxima o ponto p à curva C . Este procedimento pode ser escrito sob a forma de um operador como $\tilde{r} = \tilde{P}_m(r)$.

Observações sobre o procedimento MLS básico

Algumas observações devem ser feitas sobre o procedimento MLS básico:

1. O procedimento MLS básico não define uma projeção. Para que o operador $\tilde{P}_m(r)$ possa ser considerado uma projeção, ele deve satisfazer $\tilde{P}_m(\tilde{P}_m(r)) = \tilde{P}_m(r)$. Levin mostra em [12] que isso não é verdade. Um dos motivos pelo qual isto ocorre é que ao invés de utilizar a distância entre pontos sobre a reta H_r , este procedimento utiliza as distâncias entre os pontos de amostragem e o ponto r .
2. Seja $B(r, \delta)$ uma bola de centro r e raio δ , e seja $\tilde{P}_m(B(r, \delta))$ o conjunto das aplicações do operador $\tilde{P}_m$ a todos os pontos de $B(r, \delta)$. O conjunto $\tilde{P}_m(B(r, \delta))$ não é uma curva. Isto significa que se aplicarmos este operador aos pontos de amostragem não obteremos pontos exatamente sobre a curva, mas apenas uma nuvem de pontos “mais fina” que a nuvem de pontos inicial.

A figura 2.2 (obtida no artigo [12]) mostra o procedimento MLS básico. A figura da direita mostra a reta utilizada como referência local, e a figura da esquerda mostra as aproximações através de polinômios de segundo e terceiro graus. Nesta segunda figura, a imagem foi rotacionada de modo que a reta de referência local ficasse na horizontal. Em ambas, os pontos de amostragem foram conectados de modo a formar uma poligonal, mas isto não significa que exista alguma ordenação ou estrutura organizando os pontos de amostragem. A função peso utilizada foi $\theta(s) = e^{-s^2/h^2}$, onde h é a separação média entre os pontos de amostragem.

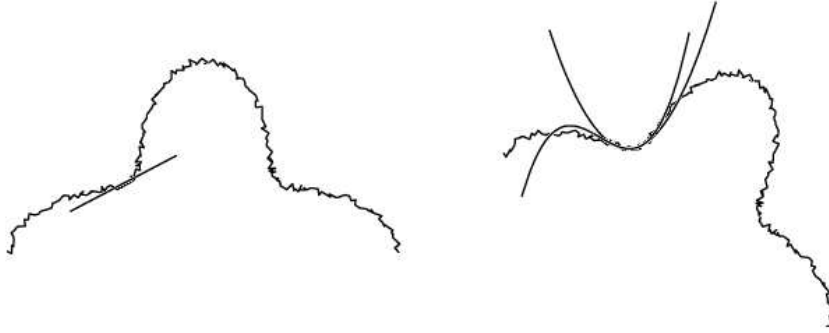


Figure 2.2: Procedimento MLS básico

2.3

Procedimento de Projeção MLS

Para corrigir os problemas apresentados acima, Levin[12] modificou o procedimento MLS básico, de modo que o novo procedimento correspondesse a um operador de projeção, e que a união dos pontos projetados fosse de fato uma curva. O procedimento de projeção MLS de um ponto r é bastante similar ao procedimento básico: primeiro, calcula-se a reta que servirá como referência local. A seguir, calcula-se o polinômio que aproxima a curva das proximidades do ponto r . A diferença entre ambos está na distância utilizada na função-peso, como veremos a seguir.

Cálculo da reta tangente aproximada

Nesta etapa, devemos encontrar, além da reta $H_r = \{x : \langle a, x \rangle - D = 0, x \in \mathbb{R}^2\}$, $a \in \mathbb{R}^2$, $\|a\| = 1$, um ponto q sobre H , de modo que $(r - q)\|a$ (ou seja, $q = r + ta$, para algum $t \in \mathbb{R}$), e que os valores de q , a e D que minimizem a seguinte expressão:

$$\sum_{i \in I} (\langle a, r_i \rangle - D)^2 \theta(\|r_i - q\|); \quad (2-2)$$

que, reescrita de modo a substituir q por $r + ta$, resulta em:

$$\sum_{i \in I} \langle a, r_i - r - ta \rangle^2 \theta(\|r_i - r - ta\|).$$

Note que como $D = \langle a, q \rangle$ e $q = r + ta$, então $(\langle a, r_i \rangle - D) = \langle a, r_i - r - ta \rangle$. Vários pares $\{a, t\}$ são mínimos da expressão acima. O par escolhido é o que apresente t de menor módulo. Observe que o t de menor módulo será aquele

onde o ponto $q = r + ta$ está mais próximo de r . Esta é uma heurística proposta em [12].

Cálculo do polinômio

Seja $\{x_i\}_{i \in I}$ o conjunto das projeções dos pontos $\{r_i\}_{i \in I}$ sobre a reta H , e seja $f_i = \langle a, r_i \rangle - D, i \in I$ as alturas dos pontos $\{r_i\}_{i \in I}$ sobre a reta H . Podemos aproximar a curva nas proximidades do ponto p através do polinômio $\tilde{p} \in \Pi_m^d$ que, dentre os polinômios $p \in \Pi_m^d$, minimiza a seguinte expressão:

$$\sum_{i \in I} (p(x_i) - f_i)^2 \theta(\|r_i - q\|)$$

A projeção MLS de r sobre a curva é dada por

$$\begin{aligned} \tilde{P}_m(r) &\equiv q + \tilde{p}(0)a = \\ &= r + ta + \tilde{p}(0)a = \\ &= r + (t + \tilde{p}(0))a \end{aligned}$$

Note que, novamente, a função peso não é mais determinada pela distância entre os pontos r e r_i , mas pela projeção de r sobre a reta de referência e os pontos de amostragem.

Propriedades da projeção MLS

Seja um ponto $r \in \mathbb{R}^2$, e seja o par $\{a, t\}$ que minimiza a expressão

$$\sum_{i \in I} \langle a, r_i - r - ta \rangle^2 \theta(\|r_i - r - ta\|)$$

Como visto anteriormente, o par $\{a, t\}$ define uma reta H , que é uma aproximação para a reta tangente à curva nas proximidades de r . A projeção q de r sobre H é dada por $q = r + ta$.

Seja um ponto r' qualquer próximo de r , $r' \neq r, r' \in H^\perp$, onde $H^\perp$ é uma reta perpendicular a H que passa por r . Podemos escrever r' em função de r como $r' = r + sa$. O par $\{a', t'\} = \{a, t - s\}$ minimiza $\sum_{i \in I} \langle a', r_i - r' - t'a' \rangle^2 \theta(\|r_i - r' - t'a'\|)$ [12].

A partir desta afirmação, podemos dizer que tanto r quanto r' possuem o mesmo sistema de referência local. Logo, o polinômio $\tilde{p}$ que aproxima localmente a curva nas proximidades da origem do sistema de referência será o mesmo tanto para r quanto para r' (note que a função de minimização utilizada

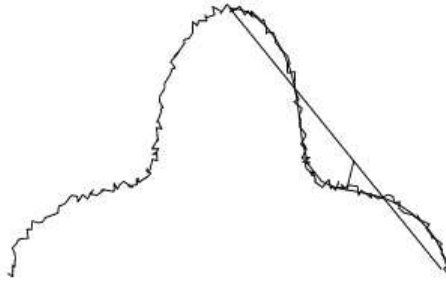


Figure 2.3: Projeção MLS

para obter o polinômio $\tilde{p}$ necessita apenas da projeção de r , que é a mesma projeção de r'). Logo, temos que

$$\begin{aligned} P_m(r') &= P_m(r + sa) = \\ &= r + sa + ((t - s) + \tilde{p}(0))a = \\ &= r + sa - sa + (t + \tilde{p}(0))a = \\ &= P_m(r) \end{aligned}$$

Fazendo $s = t - \tilde{p}(0)$, de modo que $r' = r + sa = r + (t - \tilde{p}(0))a = \tilde{P}_m(r)$, temos que, pela demonstração acima, $\tilde{P}_m(P_m(r)) = \tilde{P}_m(r)$, concluindo que o operador $\tilde{P}_m$ é um operador de projeção.

A figura 2.3 [12] mostra a suavização de uma curva através da projeção MLS. Nesta figura é mostrado o mesmo conjunto de pontos de amostragem da figura 2.2, além de um segmento de reta L próximo ao conjunto dos pontos de amostragem. A curva suave que pode ser observada na figura foi obtida a partir da projeção de 30 pontos de L . Para efeitos de visualização do processo de projeção, um dos pontos em L foi ligado à sua projeção.

2.4 Implementação

A implementação da projeção MLS de um ponto r sobre uma curva C pode ser descrita em duas etapas distintas: a primeira etapa consiste em achar a reta H de referência local, e a segunda etapa consiste em achar o polinômio $\tilde{p}$.

Como foi visto anteriormente, achar a reta H de referência local significa achar um vetor unitário a normal à reta H e um escalar t que representa a distância entre o ponto r e a reta H . Para achar a e t simultaneamente, precisamos de um esquema iterativo de minimização não-linear que decaia para o mínimo local mais próximo. Note que podem existir vários mínimos

locais, e o mínimo local escolhido deve ser aquele de menor t , de acordo com a heurística proposta em [12].

Em [20], várias observações são feitas de modo a permitir uma simplificação para o problema. A primeira observação diz respeito a um valor inicial para t e a , utilizados no método de minimização. Em [20], é sugerido que, inicialmente, t seja fixado em zero, de modo que se obtenha um vetor normal a que minimize a expressão (2-2). Este é um problema linear, já que os pesos associados a cada ponto de amostragem podem ser fixados para um valor θ_i . A nova expressão de minimização pode ser escrita como:

$$\sum_{i \in I} (\langle a, r_i \rangle)^2 \theta_i \quad (2-3)$$

Seja $B = \{b_{jk}\}$, $B \in \mathbb{R}^{2 \times 2}$ a matriz de covariância com pesos definida por:

$$b_{jk} = \sum_{i \in I} \theta_i (p_{ij} - r_j)(p_{ik} - r_k) \quad (2-4)$$

A solução para a expressão (2-3) é autovetor associado ao menor autovalor da matriz B .

O passo seguinte é, fixado um vetor normal, encontrar um valor de t que minimize a expressão (2-2). Note que, como observado anteriormente, podem existir vários mínimos locais para t . Assim, deveríamos achar todos estes mínimos locais e escolher aquele de menor t . O número de mínimos locais é limitado unicamente pelo número de pontos de amostragem [20], o que pode impossibilitar a resolução do problema, já que o número de pontos de amostragem pode ser muito grande. Porém, Alexa et al. afirmam em [20], que para valores de t distantes de no máximo $h/2$ da curva, só existe um mínimo local para t .

Assim, o valor de t pode ser encontrado a partir de um método de minimização simples, como o método de Powell [14]. A derivada parcial da expressão (2-2) em relação a t é:

$$b_{jk} = 2 \sum_{i \in I} \langle a, r_i - r - ta \rangle \left(1 + \frac{\langle a, r_i - r - ta \rangle^2}{h^2} \right) e^{\|r_i - r - ta\|/h^2} \quad (2-5)$$

Agora devemos voltar ao passo de minimização em relação ao vetor normal. Porém, agora temos que $t \neq 0$, e por isso, o problema não é mais linear (observe que o ponto $q = r + ta$ muda à medida em que o vetor a muda, e por isso, o peso associado a cada ponto de amostragem também muda com a alteração de a). Como t está fixado, o espaço de busca para um vetor normal

pode ser visualizado como a união dos planos tangentes à esfera de raio t e centro em r . Contudo, sabe-se que na prática o vetor normal a muda muito pouco nesta nova minimização, e por isso, Alexa et al. sugerem que este espaço de busca seja aproximado para a reta tangente de normal a (encontrado no primeiro passo do processo), e com distância t de r . Assim, deve-se buscar um novo ponto q nesta reta que minimize a expressão (2-2), e achar novos valores para t e a de modo que $q = r + ta$.

O problema de encontrar a reta H pode ser, então, resumido nos seguintes passos:

1. Fixe $t = 0$, e ache um vetor normal a que minimize a expressão (2-2).
2. A partir do vetor normal a fixado, encontre um valor para t que minimize a expressão (2-2).
3. Encontre um ponto q na reta H que minimize a expressão (2-2).
4. A partir do ponto q , obtenha novos valores de a e t , de modo que $q = r + ta$. Com estes valores, repita os passos 2 e 3 até que nenhum dos parâmetros a e t mude mais que um limiar ϵ prefixado.

Note que esta aproximação foi feita supondo-se que a distância do ponto r para a curva C seja menor que $h/2$. Nos casos em que essa distância for maior, devemos aproximar o ponto r da curva antes de realizar a projeção MLS. Esta aproximação pode ser feita conforme proposto em [19], onde uma aproximação para vetor normal à curva nas proximidades do ponto r é inicialmente encontrado, utilizando-se a mesma matriz de covariância B descrita na expressão (2-4). O vetor normal a é o autovetor associado ao menor autovalor de B . Em seguida, o ponto r é aproximado à curva na direção de a , movendo-se o ponto r pelo vetor ta , onde t minimiza a seguinte expressão:

$$\sum_{i \in I} \|r_i - r - ta\|^2 \quad (2-6)$$

Com este novo valor de r mais próximo da curva C , já é possível calcular a reta de referência para a projeção MLS através do procedimento descrito acima.

Uma vez obtida a reta de referência, o passo seguinte é encontrar o polinômio $\tilde{p}$. Este problema é um problema padrão de mínimos quadrados, pois a função peso associada a cada ponto de amostragem é previamente conhecida. Alexa et al. afirmam em [20] que polinômios de grau 3 ou 4 são os que melhor se ajustam à curva nas proximidades do ponto r , já que polinômios de graus mais altos tende a oscilar mais. Outra vantagem dos polinômios de grau baixo

é o menor custo computacional para encontrá-los, já que menos coeficientes devem ser calculados.