

**SERGIO AUGUSTO CUNHA**

**METODOLOGIA PARA A SUAVIZAÇÃO DE DADOS  
BIOMECÂNICOS POR FUNÇÃO NÃO PARAMÉTRICA  
PONDERADA LOCAL ROBUSTA**

**CAMPINAS**

**1998**

**SERGIO AUGUSTO CUNHA**

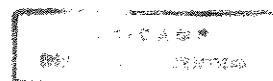
**METODOLOGIA PARA A SUAVIZAÇÃO DE DADOS  
BIOMECÂNICOS POR FUNÇÃO NÃO PARAMÉTRICA  
PONDERADA LOCAL ROBUSTA**

Tese de Doutorado apresentada à  
Faculdade de Educação Física da  
Universidade Estadual de Campinas.

**Orientador: Prof. Dr. Euclides Custódio de Lima Filho**

**CAMPINAS**

**1998**



UNIDADE	BC
N.º CHAMADA:	
V	Ex
TOMBO BC/	35096
PROC.	395/98
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	R\$ 11,00
DATA	16/09/98
N.º CPD	

CM-00115952-4

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA- FEF -  
UNICAMP**

C914m	<b>Cunha, Sérgio Augusto</b>
	<b>Metodologia para suavização de dados biomecânicos por função não paramétrica ponderada local robusta / Sérgio Augusto Cunha. -- Campinas, SP: [s.n.], 1998.</b>
	<b>Orientador: Euclides Custódio de Lima Filho</b>
	<b>Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Educação Física.</b>
<b>1. Biomecânica. 2. Movimento. 3. Cinemática. I. Lima Filho, Euclides Custódio de. II. Universidade Estadual de Campinas, Faculdade de Educação Física. III. Título.</b>	

Este exemplar corresponde à  
redação final da Tese defendida por  
Sergio Augusto Cunha e aprovada  
pela Comissão Julgadora em 14 de  
Maio de 1998.

Data: 10 de Junho de 1998

Assinatura: 

## BANCA EXAMINADORA

Prof. Dr. Euclides Custódio de Lima Filho

Euclides Lima Filho

Prof. Dr. Jocelyn Freitan Bennaton

Jocelyn

Prof. Dr. Ademir José Petenate

Ademir

Prof. Dr. Eduardo Kokubun

Eduardo Kokubun

Prof. Dr. René Brenzikofer

René Brenzikofer

## **AGRADECIMENTOS**

- \* Ao Euclides, grande amigo, incansável orientador e brilhante professor, que durante todo o curso demonstrou confiança no meu trabalho e grande paciência.
- \* À banca examinadora: René Brenzikofer, Jocelyn Bennaton, Eduardo Kokubun, Ademir Petenate, Ricardo Barros e Paulo Ruffino, pela leitura cuidadosa do texto e pelas preciosas sugestões.
- \* Aos meus amigos do Laboratório de Instrumentação para Biomecânica, pela colaboração constante no dia-a-dia de trabalho.
- \* Ao meu amigo Barreto, pelo auxílio na aprendizagem do S-Plus.
- \* Aos professores, alunos e funcionários da Unesp - Campus de Bauru e Campus de Rio Claro - pelo incentivo e compreensão.
- \* À Capes, pela concessão de uma bolsa de estudos parcial.
- \* À Unesp, pelo afastamento parcial para a realização do doutorado e pela bolsa de deslocamento concedida.
- \* À Fapesp e à Fundunesp, pelos recursos materiais utilizados neste trabalho.
- \* Ao Guarani F.C., pela permissão de realizar uma análise em seu campo.
- \* À Faculdade de Educação Física (FEF) da Unicamp, na figura de seu Diretor, Prof. Dr. Edison Duarte, pela disponibilidade das instalações para a realização da pesquisa.
- \* Aos funcionários da FEF pela presteza e cordialidade no atendimento às minhas solicitações, especialmente à Dulce, pela amizade desenvolvida durante todo este tempo.
- \* Aos meus pais Paulo e Heima (in memoriam) pela minha formação e por sempre terem acreditado muito em mim.

Dedico este trabalho à minha esposa Dena e às minhas filhas Juliana e Taís, que me incentivaram sempre com carinho e amor.

## SUMÁRIO

INTRODUÇÃO	01
OBJETIVOS	12
METODOLOGIA	13
EXEMPLOS DE APLICAÇÃO	30
APLICAÇÃO 1	31
APLICAÇÃO 2	35
APLICAÇÃO 3	40
APLICAÇÃO 4	44
CONCLUSÕES	47
REFERÊNCIAS BIBLIOGRÁFICAS	49



## LISTA DE FIGURAS

FIGURA 1 - Esquema simplificado de representação do estudo do movimento humano	01
FIGURA 2 - Gráficos de uma série de tempo e de uma família de variáveis aleatórias	05
FIGURA 3 - Gráfico do ajuste polinomial (grau 12) sobre um conjunto de dados simulados	10
FIGURA 4 – Gráfico do ajuste não paramétrico ( <i>loess</i> ) sobre um conjunto de dados simulados	10
FIGURA 5 - Gráfico da definição da janela a partir do ponto a ser ajustado ( $X_{50}, Y_{50}$ )	15
FIGURA 6 - Gráfico da função peso tricúbica definida dentro da janela escolhida	16
FIGURA 7 - Gráfico da regressão ponderada local usando polinômio de segundo grau	17
FIGURA 8 - Gráfico do ponto ajustado e do ponto medido	17
FIGURA 9 - Gráfico da definição da janela a partir do extremo inicial ( $X_1, Y_1$ )	18
FIGURA 10 - Gráfico da função peso tricúbica definida na extremidade inicial	19
FIGURA 11 - Gráfico da regressão ponderada local na extremidade inicial	19
FIGURA 12 - Gráfico do ponto ajustado e do ponto medido na extremidade inicial	20
FIGURA 13 – Gráfico das funções pesos definidas para a suavização dos pontos número 1, 6, 11, 16 e 21	21
FIGURA 14 - Gráfico da curva ajustada localmente sem e com o procedimento robusto	23
FIGURA 15 - Gráfico com o ajuste robusto e sem o ajuste robusto	23
FIGURA 16 - Gráfico da autocorrelação dos dados	25
FIGURA 17 - Gráficos do ajuste com $f = 0.8$ e dos resíduos com o mesmo ajuste	26
FIGURA 18 - Gráficos do ajuste com $f = 0.5$ e dos resíduos com o mesmo ajuste	27

FIGURA 19 - Gráficos do ajuste com $f = 0.6$ e dos resíduos com o mesmo ajuste	27
FIGURA 20 - Gráficos com os dados brutos e o ajuste, com os resíduos, com a primeira derivada e com a segunda derivada	29
FIGURA 21- Imagens de trechos de um ciclo de corrida, com o ângulo entre a perna e a vertical	33
FIGURA 22 - Gráficos da posição, dos resíduos, da velocidade e da aceleração dos dados de um ciclo da corrida	34
FIGURA 23 - Gráfico do espaço de fase do movimento de um ciclo da corrida	34
FIGURA 24 - Imagens da seqüência de um chute no futebol	38
FIGURA 25 - Gráficos da posição horizontal do pé, dos resíduos, da velocidade e da aceleração de um chute	39
FIGURA 26 - Gráfico das curvas representativas de cada chute e da média dos cinco chutes	39
FIGURA 27 - Imagens de uma seqüência de defesas durante o teste de goleiro	42
FIGURA 28 - Gráficos da posição horizontal, dos resíduos, da velocidade e da aceleração do deslocamento do goleiro	43
FIGURA 29 - Imagens de três posicionamentos do jogador durante a partida	45
FIGURA 30 - Gráfico da posição (vertical x horizontal) com ajuste e histograma das velocidade durante o jogo	46
FIGURA 31 - Gráfico do deslocamento do jogador durante o trecho analisado	46

## RESUMO

O movimento humano é um fenômeno complexo que pode ser estudado por diversas áreas. A Biomecânica atua junto a outras áreas do conhecimento para buscar uma melhor compreensão do ser humano em movimento. A análise cinemática muitas vezes gera dados discretos que necessitam do ajuste de uma função contínua para representá-los convenientemente. Se os dados estão na forma de série de tempo, a família de curvas assim caracterizadas, que possuem uma certa distribuição de probabilidades, é chamada de processo estocástico. Suavizar os dados de uma série temporal que possuem dependência estocástica entre os valores próximos consiste em deixar os dados indicarem qual a forma da curva ajustada. Os procedimentos não paramétricos, como, por exemplo, os ajustes por partes, permitem que o dado a ser suavizado sofra influência grande dos dados próximos e nenhuma dos dados mais afastados. A literatura registra várias funções para a suavização de dados discretos.

O objetivo deste trabalho é apresentar a metodologia para o ajuste de uma função não paramétrica ponderada local robusta denominada *loess* para a suavização de dados biomecânicos discretos a fim de descrever continuamente movimentos humanos.

A metodologia consiste em escolher o tamanho da janela (parâmetro  $f$ ) dentro da qual é realizada a regressão ponderada, a função peso a ser utilizada (neste caso a função tricúbica), o grau do polinômio ajustado naquela região (segundo grau para este trabalho) e comparar os valores dos resíduos com a estimação biquadrática para uma possível reponderação da regressão e exclusão dos *outliers* da mesma. Essa janela move-se de modo a suavizar cada ponto do vetor de dados,

cujo conjunto é interpolado por uma função contínua com derivadas contínuas até segunda ordem. Nos pontos próximos das extremidades, a janela assume uma forma assimétrica com os dados mais distantes participando da suavização dos pontos nas extremidades, mas a ponderação tricúbica resolve este problema.

São apresentados quatro exemplos de aplicação da função *loess* para suavizar dados biomecânicos: o movimento angular da perna esquerda de um garoto durante um ciclo da corrida, o movimento horizontal do tornozelo direito de um sujeito durante a realização de um chute no futebol com a bola parada, o deslocamento horizontal de um goleiro durante a realização de defesas em um teste específico de treinamento e a movimentação de um jogador profissional de futebol durante a realização de uma partida oficial. Em todos esses casos, coube ao pesquisador escolher o parâmetro de suavização ( $f$ ) que fornecesse uma suavização adequada, comprovada pela análise dos resíduos, pela análise da autocorrelação dos dados e pela análise do comportamento das derivadas.

A função *loess* mostrou-se muito apropriada na suavização de dados da cinemática do movimento humano, principalmente devido ao seu procedimento robusto para identificar os *outliers* e corrigir a ponderação dos dados.

**PALAVRAS-CHAVE:** Biomecânica

Ajuste de funções

Regressão ponderada local robusta

## ABSTRACT

Human movement is a complex phenomenon that must be studied by many areas. Biomechanics works together with other areas to search a better insight of the human movement. Many times, the kinematics analysis produce discrete data and one needs to adjust a continuous function to represent them with accuracy. If the data are in time series, and the curves family that is formed has a probability distribution, it is called stochastic process. To smooth the time series data that have a stochastic dependence between the nearest values, means to leave the data show their adjustment curve shape. The non parametric procedures, partial adjustments for instance, allow that the smoothed data to receive a big influence from the neighbors and none influence from the apart points. There are a lot of functions that can be found in the literature to smooth discrete data.

The aim of this work is to show a methodology for the robust local weighted non parametric function to smooth biomechanics discrete data to describe continuous human movements.

The methodology consists in choosing the window size (smoothing parameter  $f$ ) in witch the weighted regression will be held, the weight function that will be used (in this case, the tricubic weight), the polynomial degree adjusted in that region (second degree for this work) and compare the residuals values with the biquadratic estimation for a possible reweight in the regression to eliminate the outliers. This window moves itself so that it can smooth each point of the data vector, which ensemble will be interpolated by a continuous function that has a continuous second derivative. Near the edge points the window assumes an asymmetric form and there is a participation of the remote data, but the tricubic weight solves this problem.

We show four examples of application of *loess* function to smooth biomechanics data: an angular movement of a boy's left shank during a cycle of a race, the horizontal movement of the right ankle of a subject during a soccer kick session with stopped ball, the goalkeeper horizontal displacement during a specific training test of defenses, and the professional soccer player movement during an official game. In all of these cases, the researcher chose the smooth parameter ( $f$ ) that provide an adequate smoothing, that were proved by the residuals analysis, by the data autocorrelation analysis and by the derivatives behaviors analysis.

The *loess* function revealed itself very appropriate in kinematics data smoothing of the human movement, mainly depict of the robust procedure to identify outliers and to correct data weight.

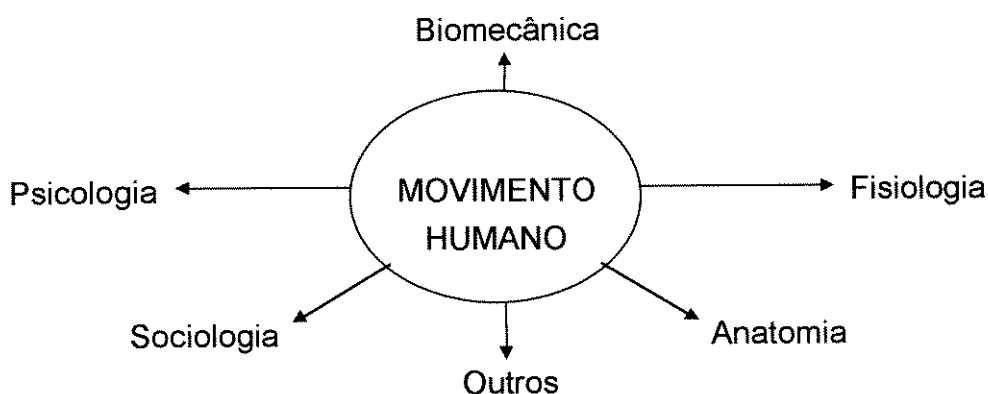
**KEY WORDS:** Biomechanics

Adjustment of functions

Robust local weighted regresion

## INTRODUÇÃO

Para compreender o movimento humano, é importante o seu aprofundamento envolvendo diferentes aspectos, que estão relacionados às mais diversas áreas de estudo e podem trazer uma contribuição significativa ao estudo da motricidade humana. Na área da Educação Física, o estudo da motricidade é um fator fundamental para a elaboração de programas de atividades motoras adequados para cada faixa etária. Desta forma, pode-se considerar o movimento humano como o centro de um esquema que envolve as áreas que compõem seu universo de estudo (Figura 1).



**FIGURA 1**  
**Esquema simplificado de representação do estudo do movimento humano**

Dentre os vários aspectos sob os quais se pode abordar o movimento humano, a Biomecânica é a área que estuda os movimentos do ponto de vista das leis da Mecânica. É a partir dos conceitos desta área que a Biomecânica atua no sentido de auxiliar outras áreas na busca da compreensão deste fenômeno tão complexo.

Um dos ramos da Biomecânica é a análise cinemática de movimentos. Para realizá-la, é preciso registrar o movimento desejado para poder arquivá-lo de modo a executar os procedimentos metodológicos necessários para obter, por exemplo, posição, velocidade e aceleração de um certo segmento, ou as variáveis angulares de cada articulação. A determinação desses conceitos pode auxiliar na inferência de quais grupos musculares estão agindo. Outro ramo é a Dinâmica, que permite estudar, por exemplo, quais são as sobrecargas em cada articulação.

As informações provenientes de todas essas áreas são importantes para que o profissional da Educação Física e dos Esportes possa planejar suas atividades com um amplo conhecimento do que ocorre sob o ponto de vista biomecânico. Hoje em dia, pode-se dizer que a Biomecânica está diretamente ligada a áreas como Aprendizagem Motora, Controle Motor, Fisiologia do Exercício, Educação Física Adaptada, Práticas Desportivas (individuais e coletivas), Fisioterapia, Ortopedia e Bioengenharia.

A partir do registro do movimento que se deseja investigar, há, basicamente, dois tipos de análises: a qualitativa e a quantitativa. No entanto, estes dois tipos não podem ser dissociados, já que, na prática, devem ocorrer concomitantemente. No entanto, quando se procura detectar alterações sensíveis no movimento, é fundamental a utilização da análise quantitativa para mensurar as variações, exigindo um instrumental e uma metodologia que permitam captar as alterações procuradas.

Com o avanço dos recursos tecnológicos, têm sido realizadas cada vez mais coletas de dados em Biomecânica. O surgimento dos computadores após a segunda guerra mundial possibilitou o desenvolvimento de técnicas de análise numérica

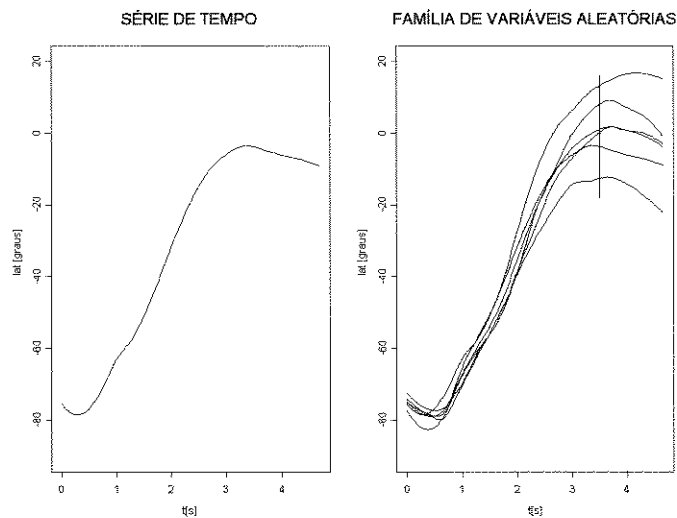


exigidas para a solução de uma série de problemas, como, por exemplo, a suavização ou filtragem de variáveis obtidas pela mensuração do movimento humano. Como este é altamente complexo, modelá-lo matematicamente não é uma tarefa trivial. Sempre que se coletam dados em um experimento com o uso de instrumentos, estes dados vêm contaminados com ruídos. Separar o que é sinal do que é ruído é uma etapa importante (WINTER *et al*, 1974), devendo ser seguida pelo estudo de que tipo de flutuação é inerente à natureza humana e qual é relativo ao instrumento.

Os dados biomecânicos são muitas vezes registrados por sistemas de medições que produzem variáveis discretas. A variável discreta é aquela cujos pontos não preenchem todo o intervalo, e a contínua aquela que ocupa todo o intervalo. Sabe-se que o movimento humano é notoriamente contínuo. Portanto, a partir dos dados discretos que geralmente são obtidos numa análise biomecânica busca-se uma representação do fenômeno investigado. Pode-se representá-lo utilizando uma curva que se aproxime dos pontos mensurados, possibilitando, assim, obter informações em qualquer instante desejado, não apenas naqueles em que os dados foram registrados. Encontrar a curva mais adequada ao conjunto de pontos obtidos no processo de medição é procurar dentre a família de funções matemáticas contínuas aquela que represente melhor o movimento estudado. Além disto, de posse da variável contínua, é possível obter, por exemplo, as derivadas primeira e segunda para chegar à velocidade e à aceleração, no caso da análise cinemática.

Sabe-se que o ser humano movimenta-se com uma variabilidade inerente, que deve ser levada em conta ao estruturar a metodologia de análise do movimento.

O presente estudo, por exemplo, baseia-se na análise quantitativa de variáveis cinemáticas, a partir do ajuste de funções matemáticas representativas do movimento escolhido, considerando a flutuação dos dados medidos e o grau de dependência dos mesmos. A escolha de um modelo matemático para simular o fenômeno físico investigado pode ser feita através de um processo determinístico, utilizando métodos numéricos para definir a curva ajustada. No presente estudo, entretanto, analisam-se os dados através de um processo estocástico, também denominado de função aleatória. Um processo estocástico é uma família de variáveis aleatórias com uma certa distribuição de probabilidades (MORETTIN, 1979). Cada uma das trajetórias (ou curvas) de um processo estocástico pode ser uma série temporal, se a mesma estiver definida no domínio do tempo. Em análise do movimento, muitas vezes os dados colhidos são séries de tempo. Como o fenômeno do movimento humano é variável, ele é representado por uma nuvem de curvas características do movimento em questão, uma para cada medição realizada. Fixando-se um determinado instante de tempo, tem-se uma flutuação gerada pelas várias trajetórias, ou seja, tem-se uma variável aleatória com uma distribuição de probabilidades para cada instante (Figura 2). Desta forma, pode-se investigar o movimento humano levando em conta tais flutuações. Para isso, devem ser aplicados procedimentos estatísticos (esperança e variância) em processos estocásticos.



**FIGURA 2**

**Gráficos mostrando uma série de tempo (esquerda) e uma família de variáveis aleatórias com uma certa distribuição de probabilidades (direita)**

Antes de escolher qual a curva que será utilizada, deve-se ressaltar que há uma diferença quando se fala em interpolação ou aproximação. No primeiro caso, a curva passa exatamente por todos os elementos do conjunto de dados, não havendo resíduos (distância entre o ponto medido e curva ajustada naquele ponto). Contudo, quando os dados são obtidos a partir de medidas experimentais, esta técnica não é adequada, pois os resultados obtidos contêm uma incerteza característica deste processo (CUNHA, 1993a). Sendo assim, para dados biomecânicos, é muitas vezes preferível fazer uma aproximação, ou seja, ajustar uma curva segundo algum critério de otimização deste processo. Por isso a aproximação também pode ser chamada de ajuste. Neste caso, escolhe-se dentre uma determinada classe de funções aquela que melhor se ajusta aos dados experimentais e analisa-se a qualidade do ajuste efetuado através, por exemplo, dos resíduos (CUNHA, 1993b). Um dos métodos mais utilizados para essa otimização é o método dos quadrados mínimos, que consiste em minimizar a soma dos quadrados dos resíduos.

Para realizar um ajuste dito paramétrico, deve-se ter informações a respeito do modelo que representa o fenômeno que se quer investigar. Se, por exemplo, existe a informação antecipada sobre o comportamento de um movimento, basta resolver o sistema de equações para achar os parâmetros do modelo e, assim, obter a curva ajustada, seja de primeiro grau (reta) ou polinômios de graus maiores. Dentre os ajustes paramétricos, um dos mais importantes é o ajuste por funções polinomiais (LANCASTER & SALKAUSKAS, 1986), cujo melhor exemplo é a regressão linear. A escolha adequada do modelo é de fundamental importância para que a curva ajustada represente adequadamente o fenômeno investigado.

Entretanto, pouco se sabe a respeito da modelagem para o movimento humano e, na maior parte das vezes, é muito difícil prever qual o modelo que se adapta a uma determinada situação específica, o que dificulta o uso dos ajustes paramétricos. Além disto, como os estudos da cinemática do movimento humano são, geralmente, séries de tempo, nas quais existe uma dependência estocástica entre os dados próximos, é necessário utilizar outra forma de ajuste para resolver o problema da discretização dos dados em tais casos. Assim, é preferível utilizar os ajustes não paramétricos, nos quais é o próprio conjunto de dados que vai indicar qual a forma da curva ajustada. Quando não se conhece a relação entre a variável resposta, ou variável dependente ( $Y$ ), e a variável preditora, ou variável independente ( $X$ ), deve-se deixar que os dados mostrem qual a melhor forma funcional que os relacionem (HASTIE, TIBSHIRANI, 1990). Este tipo de ajuste também é chamado de *smoothing* (suavização) porque produz uma estimativa da resposta menos variável que a resposta original (STATISTICAL SCIENCES, 1993). Os *smoothers* (suavizadores) são capazes tanto de captar a tendência de  $Y$  sobre  $X$

como também de estimar a dependência da média de  $Y$  sobre o preditor  $X$  (ou preditores  $X_i$ ) (HASTIE, TIBSHIRANI, 1990). Alguns exemplos de funções não paramétricas são: *loess* ou *lowess* (*robust locally weighted regression and smoothing scatterplots* ou regressão ponderada local robusta e suavização em diagrama de dispersão) (CLEVELAND, 1979; CLEVELAND 1985 ; HASTIE & TIBSHIRANI, 1990), *kernel* (HÄRDLE, 1990 ; WAND & JONES, 1995) e *spline* (CRAVEN & WAHBA, 1979; SILVERMAN, 1985), entre outras. As duas primeiras são métodos de ajuste ponderado local, no qual a regressão é realizada em uma vizinhança de cada ponto a ser suavizado, com pesos maiores para dados mais próximos desse ponto e peso zero para dados fora da vizinhança determinada. Já os *splines* tem como característica principal a escolha de "nós", em que se define uma combinação linear de polinômios básicos para realizar a suavização (SILVERMAN, 1985). Todos esses casos baseiam-se em operadores por convolução (MÜLLER, 1987).

A suavização tem como idéia básica a determinação de médias ponderadas entre os elementos do vetor de dados contidos em uma determinada vizinhança, ou seja, em uma região composta por esses elementos. A escolha da vizinhança, também chamada de janela, janelamento, ou parâmetro de suavização, é responsável pela determinação da relação entre a tendência e a variância. Chama-se de tendência a diferença entre a curva gerada (teórica) e o valor médio das curvas suavizadas após amostragens repetidas (HASTIE & TIBSHIRANI, 1990). E variância é a medida da dispersão das curvas em torno da esperança, fornecendo a idéia da variabilidade do fenômeno. Portanto, uma janela grande significa uma suavização forte dos dados, produzindo uma variância pequena, mas com aumento

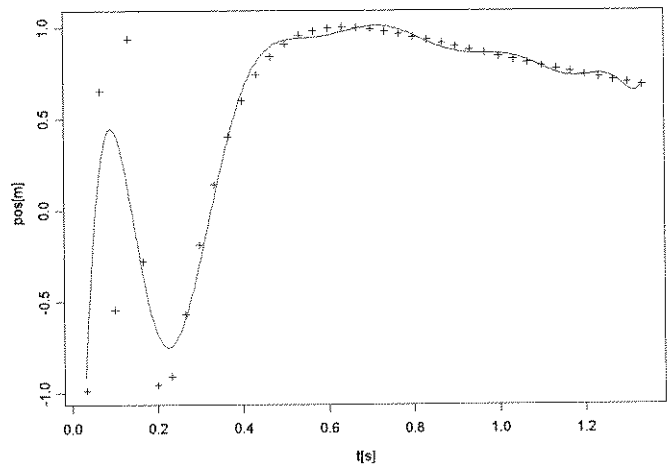
da tendência. A escolha de uma janela pequena gera, ao contrário, pouca suavização, com grande variância e diminuição da tendência (HASTIE & TIBSHIRANI, 1990).

Muitos trabalhos em Biomecânica encontrados na literatura mostram estudos avaliando diferentes tipos de funções para suavizar os dados cinemáticos e dinâmicos. Zernicke *et al* (1976) compararam o ajuste polinomial com o ajuste por função *spline* em dados de um chute no futebol, concluindo que o uso do ajuste por *spline* era promissor e apresentava melhores resultados do que o ajuste com polinômio de quinto grau para o caso estudado. McLaughlin *et al*, 1977 utilizaram os *splines* cúbicos para o ajuste de dados cinemáticos de um salto vertical comparados com os dados dinâmicos obtidos por uma plataforma de força, com o ajuste polinomial e com os valores da derivada segunda obtidos por diferença finita, concluindo que um dos principais problemas dos *splines* está nos pontos iniciais e finais. Pezzack *et al*, (1977) compararam os dados obtidos analogicamente por um acelerômetro com os dados cinemáticos de aceleração obtidos por diferença finita, polinômio de Chebyshev e filtro digital seguido por técnica de diferença finita, concluindo que o filtro digital produz os dados de aceleração mais próximos dos dados analógicos. Wood & Jennings (1979) utilizaram os dados obtidos por Pezzack *et al* (1977) para fazer uma comparação entre os *splines* cúbicos e quádruplos, chegando à conclusão de que estes últimos se mostraram mais apropriados pelo fato de apresentarem, nos dados finais da segunda derivada, maior proximidade com os dados brutos. No entanto, não foram encontrados na literatura exemplos de utilização de ajustes semelhantes aos realizados através do método *loess* proposto por Cleveland (1979). Esse método é uma combinação desenvolvida a partir de um

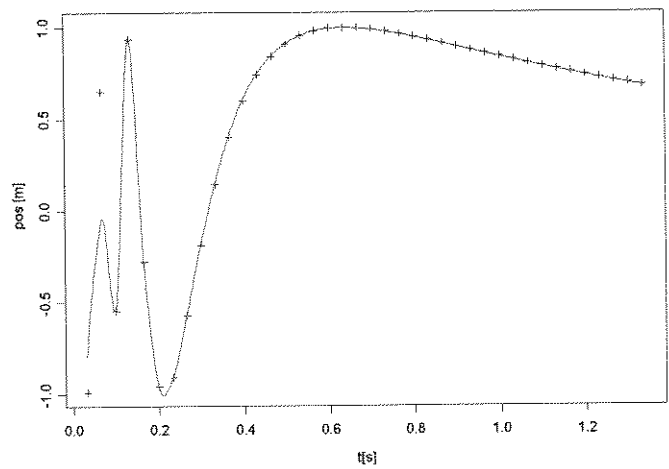
ajuste polinomial local, usado por muito tempo para suavizar séries temporais com valores igualmente espaçados (MACAULEY *apud* CLEVELAND, 1979), acrescida de um procedimento robusto para evitar que os valores discrepantes perturbem os pontos ajustados (MOSTELLER & TUKEY, 1977). Valor discrepante (*outlier*) é um ponto raro cujo valor é muito maior ou muito menor do que o dos outros pontos na vizinhança (CLEVELAND, 1985). Associam-se pesos a cada ponto contido na vizinhança escolhida e a função peso é aplicada ao vetor de dados por convolução. Escolhe-se o grau do polinômio que será utilizado e realiza-se a regressão para determinar o ponto ajustado, repetindo-se esse procedimento para todos os pontos do conjunto de dados. Finalmente, se há valores extremos nos resíduos, recalcula-se a função peso a partir de uma técnica de estimação robusta (MOSTELLER & TUKEY, 1977).

Deve-se levar em conta, também, que podem haver padrões distintos ao longo do conjunto de coordenadas. Desta forma, se uma curva for ajustada sobre todo o vetor, o comportamento de uma parte estará influenciando diretamente sobre todas as outras. Por exemplo, se no início o movimento for realizado com alta frequência e da metade para o final for estável, a função ajustada da mesma forma do começo até o fim vai ter de considerar todo o conjunto como se existisse relação entre as partes. Para dados biomecânicos, o ideal é fazer o ajuste por partes, ajustando cada subintervalo conforme suas características, para que não sofra influências dos outros subconjuntos. Para exemplificar, as Figuras 3 e 4 apresentam dados simulados (com frequências distintas) ajustados parametricamente por um polinômio de grau doze e por uma regressão não paramétrica local, respectivamente. Fica evidente nos gráficos que o ajuste paramétrico não consegue

reproduzir adequadamente o conjunto de dados, exatamente por sua característica global. Já com os dados suavizados por uma regressão não paramétrica ponderada local (*loess*) consegue-se um ajuste que acompanha as variações de modo satisfatório.



**FIGURA 3**  
**Ajuste polinomial (grau 12) sobre um conjunto de dados simulados**



**FIGURA 4**  
**Ajuste não paramétrico (*loess*) sobre um conjunto de dados simulados**



Portanto, os ajustes não paramétricos são uma ferramenta poderosa para a representação contínua do movimento humano. O presente estudo aborda o método denominado *loess* (ou *lowess*) aplicado em dados cinemáticos de movimentos desportivos diversos, a fim de ampliar o espectro de funções que podem ser utilizadas para suavizar dados biomecânicos.

## OBJETIVOS

Objetivo geral:

ajustar uma função não paramétrica ponderada local robusta para suavizar conjuntos de dados biomecânicos discretos e descrever continuamente movimentos humanos.

Objetivos específicos:

- descrever os procedimentos de um ajuste não paramétrico denominado regressão ponderada local robusta (*loess*);
- apresentar quatro exemplos de aplicação da função ajustada utilizando dados cinemáticos de movimentos humanos.

## METODOLOGIA

Neste capítulo serão detalhados os procedimentos principais para o ajuste da função *loess*. Será utilizado um conjunto de dados simulados para exemplificar os passos seguidos na descrição da metodologia. Esse conjunto de dados está na forma de uma série de tempo, pois grande parte dos dados em Biomecânica (e também nos exemplos de aplicação que serão apresentados no capítulo seguinte) são colhidos como série temporal, muitas vezes com a variável preditora equispçada. Nessa simulação, utilizou-se a função seno (definida entre 0 e  $2\pi$  e com 100 pontos discretos definidos nesse intervalo e igualmente espaçados) contaminada por um ruído que continha aleatoriamente 20% dos seus componentes vindos da distribuição gaussiana com média zero e variância igual a um ( $N\sim(0;1)$ ) e 80% vindos da distribuição gaussiana com média zero e variância igual a um décimo ( $N\sim(0;0.1)$ ). A escolha desse exemplo teve a intenção de mostrar como se aplica a metodologia proposta em um vetor de dados que contém alguma flutuação, como acontece normalmente com dados biomecânicos que carregam uma incerteza devido ao processo de medição e à variabilidade dos sujeitos que executam a ação.

Inicialmente, apresenta-se a formulação matemática do problema. Dado o conjunto de dados  $(x_i, y_i)$ , pode-se considerar que, para cada valor da variável estudada, temos:

$$y_i = g(x_i) + \varepsilon_i, i=1, \dots, n$$

onde  $g$  é a função que relaciona  $x$  com  $y$  (no caso estudado,  $g$  pertence à classe das funções que têm derivada contínua até ordem 2),  $\varepsilon$  é a incerteza associada à medição e  $n$  é o tamanho do vetor de dados.

O valor ajustado é dado por:

$$\hat{y}_i = \hat{g}(x_i),$$

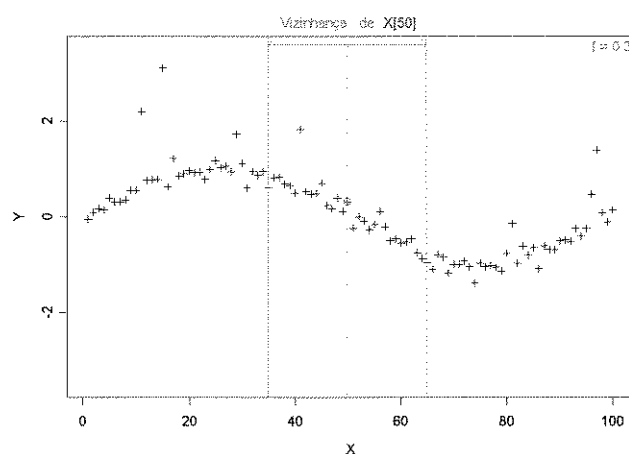
e é a função *loess* que procura a  $\hat{g}$ , que é a regra do ajuste, não explicitada analiticamente.

Como se trata de um ajuste local, apenas um subvetor por vez participa da regressão. É fundamental, então, a escolha da janela (**f**) dentro da qual ocorrerá tal procedimento, ou seja, a determinação da porcentagem do vetor de dados que será usada para cada ajuste parcial. Associam-se pesos (**W**) a cada ponto contido na vizinhança de  $x_i$  para suavizar  $y_i$ . A seguir, escolhe-se o grau do polinômio (**d**) que será ajustado por quadrados mínimos em cada subintervalo gerado. Desta forma, obtém-se o valor suavizado ( $\hat{y}_i$ ) do ponto  $(x_i, y_i)$ . Esse procedimento repete-se para todos os pontos pertencentes ao conjunto de dados, com pequenas modificações quando se trata de pontos próximo às extremidades, como será explicado adiante. Se houver valores discrepantes na análise dos resíduos, executa-se uma reponderação dos dados utilizando-se um procedimento robusto que evita que esses *outliers* distorçam a característica do conjunto de pontos.

Para tanto, define-se uma vizinhança (**f**) a partir de cada ponto  $x_i$  a ser ajustado. Fazem parte dessa vizinhança todos os pontos do conjunto de dados que estão posicionados a uma certa distância de  $x_i$ , tendo como ponto mais distante aquele  $x_k$  medido pela norma infinita ( $||.||_\infty$ ), que é o valor máximo dos elementos do vetor (CUNHA, 1993a). Esta subsequência pode ser definida como segue:

$$\Delta(x_i) = ||x_i - x_k||_\infty \quad \text{e} \quad f = 2 \Delta(x_i) n^{-1}$$

O pesquisador fornece o valor do parâmetro  $f$  para poder determinar onde serão colocados os limites da janela. Convém salientar que, quando se trata de série de tempo equiespaçada, esta janela é bilateral, ou seja, está localizada a  $f/2$  pontos à direita como também a  $f/2$  pontos à esquerda de  $x_i$ . Os pontos limitantes da janela não participam da regressão pois recebem peso igual a zero, mas todos os outros pontos que estão dentro do janelamento participam com pesos distintos, em função da distância em que se encontram do ponto a ser suavizado (Figura 5).



**FIGURA 5**  
**Definição da janela ( $f=0.3$ ) a partir do ponto a ser ajustado**  
**(neste exemplo, a coordenada  $(X_{50}, Y_{50})$ , em vermelho)**

Quanto menor for a janela escolhida, menor será a suavização e, portanto, menores também os valores dos resíduos, e quanto maiores forem os valores do parâmetro, mais suavizada será a curva e maior o resíduo.

Como se trata de um ajuste ponderado, deve-se escolher a função peso na qual os pontos mais próximos de  $x_i$  recebem peso maior do que os mais afastados (Figura 6). Isso significa que os pontos mais próximos daquele a ser ajustado carregam mais informação a respeito do que está ocorrendo naquele instante do que os pontos

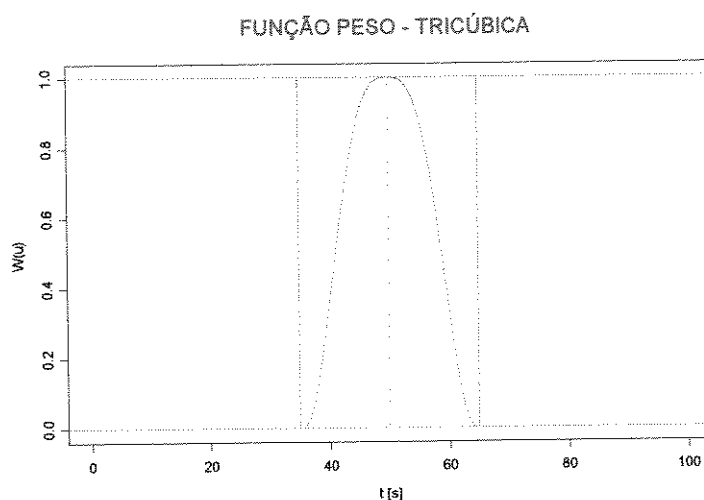
mais afastados. Define-se  $u = |x_i - x_j| / \Delta(x_i)$ , sendo  $j$  os indexadores dos  $x$  que pertencem à vizinhança  $\Delta(x_i)$ .

Uma das funções mais utilizadas é a tricúbica:

$$W(u) = W\left(\frac{|x_i - x_j|}{\Delta(x_i)}\right)$$

$$W(u) = (1 - |u|^3)^3, \text{ para } |u| < 1 \text{ e}$$

$$W(u) = 0, \text{ para } |u| \geq 1$$

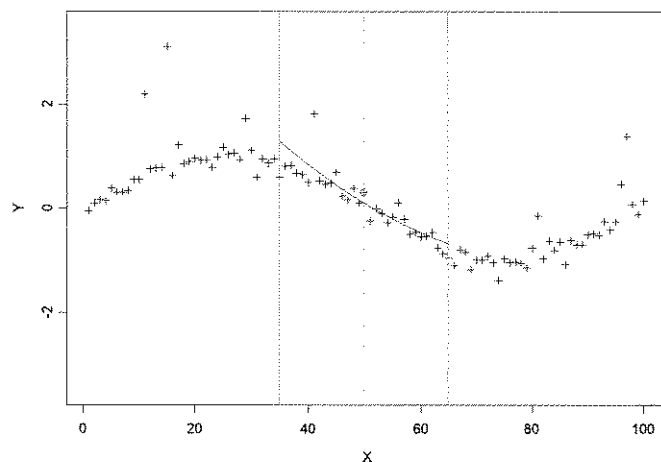


**FIGURA 6**

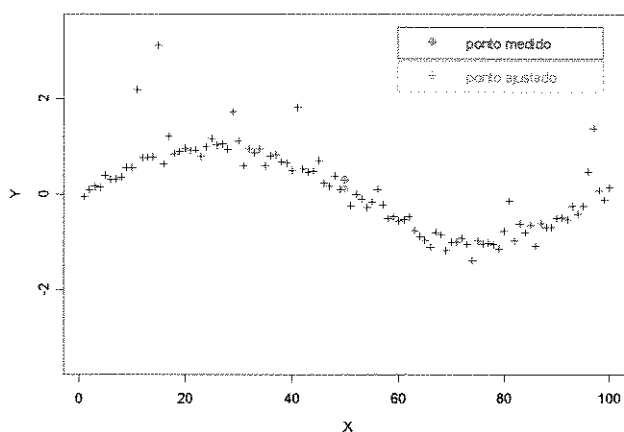
**Gráfico da função peso tricúbica definida dentro da janela escolhida ( $f=0.3$ )**

A seguir, escolhe-se o grau do polinômio ( $d$ ) que será ajustado por quadrados mínimos em cada subintervalo gerado. No caso específico deste trabalho, utilizou-se o polinômio de segundo grau, porque é interessante para efeitos computacionais escolher o menor grau possível para a realização da regressão, mas fundamentalmente a escolha cabe ao investigador.

Por fim, obtém-se a função ajustada por quadrados mínimos ponderados  $\hat{y}$  sobre uma base de polinômios de grau  $d$  em cada subintervalo (Figuras 7 e 8).



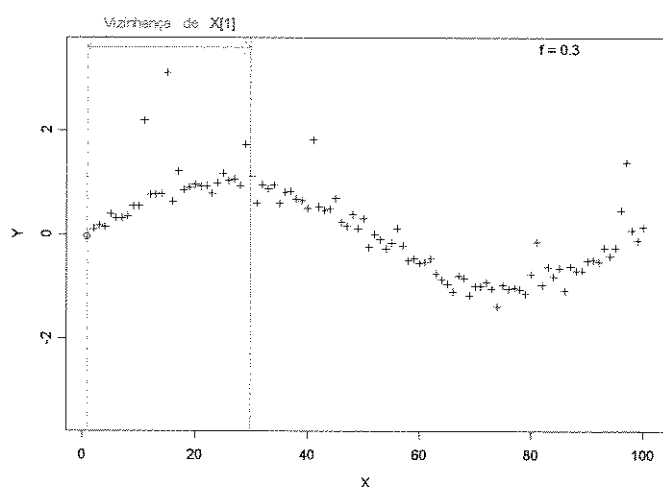
**FIGURA 7**  
**Regressão ponderada local usando polinômio de segundo grau**



**FIGURA 8**  
**Ponto ajustado (em azul) colocado junto com o ponto medido (em vermelho)**

Este procedimento é repetido para ajustar todos os pontos do vetor de dados. Nas extremidades, a função utiliza um janelamento unilateral, ou seja, só são incluídos na janela os pontos situados à direita do ponto a ser suavizado quando este se encontra no início do movimento, ou à esquerda quando se trata do ponto situado no final (Figura 9). A janela assimétrica também acontece quando os pontos

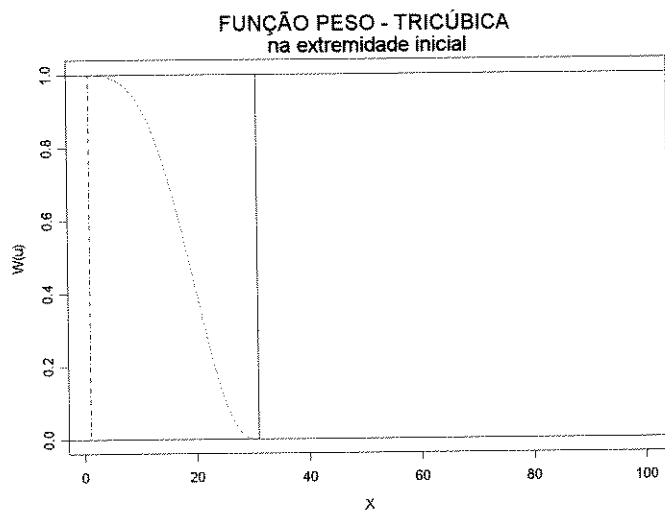
estão posicionados perto das extremidades e quando não se tem o mesmo número de pontos em ambos os lados. No caso dos extremos, poderia haver uma influência maior dos pontos mais distantes, conflitando com a dependência estocástica que as séries temporais possuem. No entanto, a ponderação tricúbica resolve esse problema, pois atribui pesos bem pequenos para os dados que estão muito distantes (HASTIE & TIBISHIRANI, 1990), conforme mostra a Figura 10.



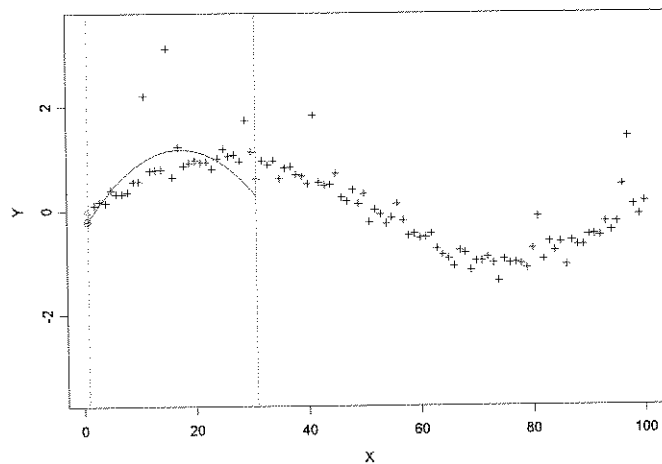
**FIGURA 9**  
**Definição da janela ( $f=0.3$ ) a partir do extremo inicial ( $X_1, Y_1$ )**

Os mesmos procedimentos de regressão relatados anteriormente são utilizados para determinar o ponto ajustado (Figuras 11 e 12).

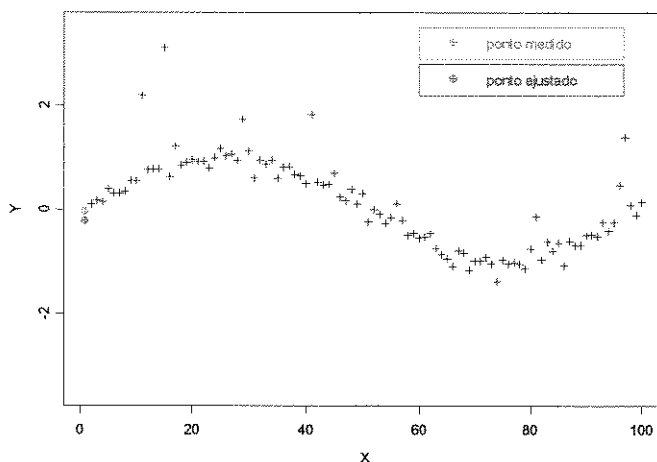




**FIGURA 10**  
Gráfico da função peso tricúbica definida na extremidade inicial

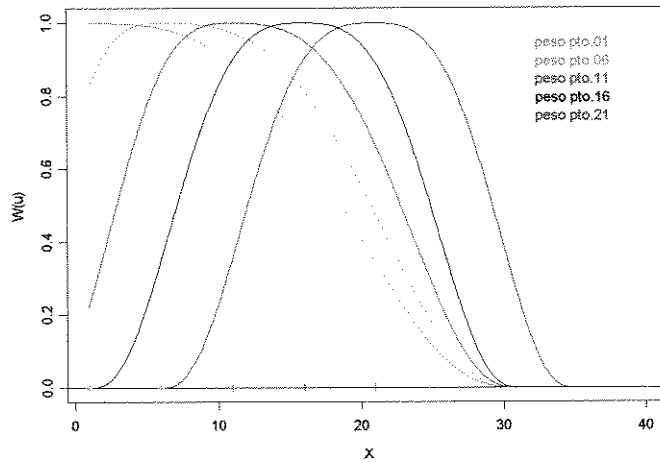


**FIGURA 11**  
Regressão ponderada local na extremidade inicial



**FIGURA 12**  
**Ponto ajustado (azul) e ponto medido (vermelho) na extremidade inicial**

Tanto no início como no final do vetor, a ponderação dos dados é feita de forma assimétrica. Isto ocorre sempre que o ponto a ser suavizado não está definindo uma janela que contém igual número de pontos à direita e à esquerda dele. Pode-se observar na Figura 13 que, conforme os pontos a serem ajustados afastam-se da extremidade, a função peso (cujos limites laterais coincidem com os limites da janela) vai se deslocando até tornar-se simétrica. Para o caso exemplificado aqui, essa simetria surge a partir do ponto número dezesseis, mantendo-se até o ponto número oitenta e cinco, já que  $f$  igual a 0.3 para um vetor de tamanho 100 indica que os limites da janela devem estar localizados sobre os 15º ponto mais afastado de cada lado, o que resulta, portanto, em 29 pontos participando do processo de regressão com pesos diferentes de zero.



**FIGURA 13**

**Funções pesos definidas para a suavização dos pontos números:  
1 (azul), 6 (verde), 11 (vermelho), 16 (preto) e 21 (magenta)**

Para obter a função contínua, é preciso prever os valores em cada subintervalo gerado por dois pontos distintos consecutivos. Para isso, faz-se a interpolação dos dados suavizados, escolhendo a melhor curva em uma família de curvas contínuas com derivadas também contínuas até segunda ordem. Pode-se, ainda, suavizar novamente os valores interpolados, para evitar que as derivadas contenham altas frequências devido ao processo de interpolação.

Define-se o resíduo como:

$$e_i = y_i - \hat{y}_i$$

A parte robusta deste procedimento pode ser implementada a partir do cálculo da mediana dos valores absolutos dos resíduos ( $s$ ), como segue:

$$s = \text{med}(|e_i|)$$

$$\delta_k = B(e_k/6s)$$

$$v = e_k/6s$$

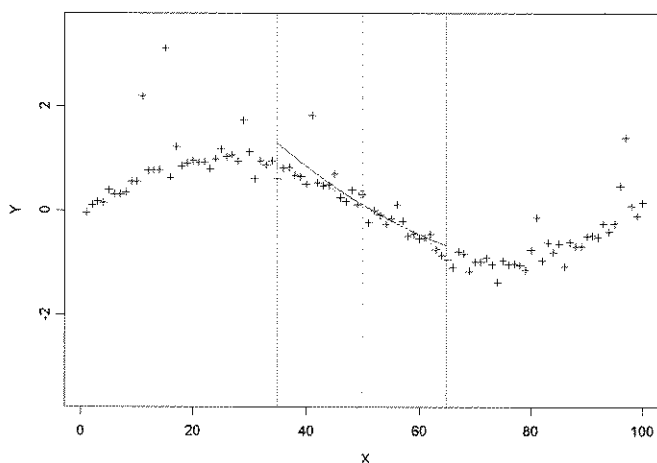
$$B(v) = (1 - |v|^2)^2, \text{ para } |v| < 1 \text{ e}$$

$$B(v) = 0, \text{ para } |v| \geq 1$$

sendo **B** a estimação biquadrática que será empregada para calcular a nova função peso (**W'**) a ser utilizada na reponderação dos dados, como segue:

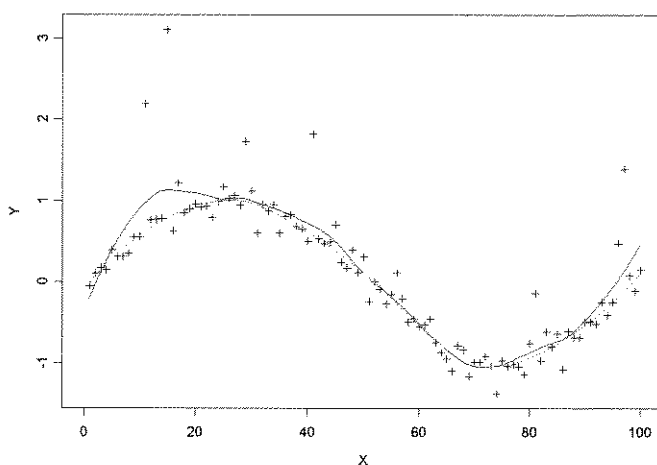
$$W' = \delta_k \cdot W(x_k)$$

Esta iteração é realizada *t* vezes, até que não haja mais *outliers* nos resíduos. A literatura sobre o assunto indica que os pontos que produzem os *outliers* são eliminados da regressão com no máximo duas iterações (CLEVELAND, 1985). A estimação biquadrática **B** é uma função que apresenta melhores resultados que qualquer outra técnica robusta (MOSTELLER & TUKEY, 1977). A principal idéia desse procedimento é comparar os valores absolutos dos resíduos ( $|e_i|$ ) com a variabilidade do ajuste ( $6s$ ), que em dados com comportamento aproximadamente gaussiano, corresponde a quatro vezes o desvio-padrão dos dados. Se essa relação ( $e_k/6s$ ) for maior ou igual a um, significa que o resíduo apresenta realmente um valor discrepante e, portanto, o ponto que produziu tal resíduo não deve ser levado em conta no ajuste. Nesse caso, o peso desse ponto passa a valer zero. Por outro lado, se essa relação for menor do que um, repondera-se o ajuste dando pesos maiores para os pontos cujos resíduos inicialmente apresentaram valores menores. A Figura 14 mostra um ajuste sem o procedimento robusto (curva em vermelho) e um ajuste executado com o procedimento robusto (curva em azul), realçando a diferença que a reponderação com a estimação biquadrática produz por excluir os valores discrepantes (*outliers*) da regressão.



**FIGURA 14**  
**Curva ajustada sem o procedimento robusto (vermelha)**  
**e com o procedimento robusto (azul)**

É possível escolher entre o ajuste com ou sem o procedimento robusto, dependendo do movimento que está sendo analisado. A Figura 15 mostra um gráfico no qual aparecem as duas situações, com a curva em vermelho representando o ajuste normal e a curva em azul o ajuste com o procedimento robusto.



**FIGURA 15**  
**Gráfico com o ajuste robusto (azul) e sem o ajuste robusto (vermelho)**

Dentre todos os itens descritos, o grau do polinômio (**d**), a função peso utilizada (**W**) e o número de iterações para eliminar os *outliers* (**t**) não precisam ser alterados em função do movimento a ser analisado. Já o parâmetro de suavização (**f**) deve ser escolhido de acordo com a característica dos dados obtidos. Essa escolha é feita subjetivamente e sua otimização compete ao investigador. Um critério que pode auxiliar nessa opção é a análise da autocorrelação dos dados. A autocorrelação fornece uma idéia do grau de dependência entre cada elemento do vetor dos dados. Assim, pode-se executar uma análise prévia para obter mais um indicativo da melhor escolha do parâmetro **f**. A autocorrelação é definida a partir do conceito de autocovariância.

Sendo  $\gamma(\mathbf{k}) = \text{cov}\{\mathbf{Y}(\mathbf{t}), \mathbf{Y}(\mathbf{t}-\mathbf{k})\}$  a autocovariância de passo **k** entre as variáveis aleatórias **Y(t)** e **Y(t-k)**, e  $\gamma(\mathbf{0}) = \text{var}(\mathbf{Y}(\mathbf{t}))$  a variância de **Y(t)**, define-se a função de autocorrelação de passo **k** como:

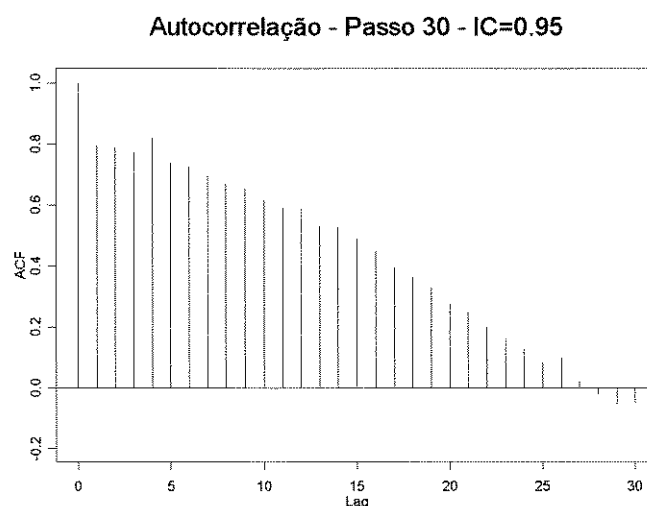
$$\rho(\mathbf{k}) = \gamma(\mathbf{k}) / \gamma(\mathbf{0})$$

Esta função, também chamada de correlação serial por ser aplicada a uma série de tempo, tem as seguintes propriedades:

- 1- simétrica, ou seja,  $\rho(\mathbf{k}) = \rho(-\mathbf{k})$ ;
- 2- definida no intervalo  $[-1, 1]$ ;
- 3- se **Y(t)** e **Y(t-k)** são independentes, então  $\rho(\mathbf{k}) = 0$ .

É importante salientar que o fato da autocorrelação ser igual a zero não garante a independência das variáveis. No caso em estudo, as variáveis **Y(t)** e **Y(t-k)** serão independentes se o par for gaussiano.

A Figura 16 mostra a autocorrelação da variável apresentada como exemplo. Nota-se que os valores dos 22 pontos mais próximos estão fora do intervalo de confiança igual a 95% (limitado pelas linhas azuis), o que significa que, com exceção dessas coordenadas mais próximas, todas as outras assumem uma autocorrelação igual a zero. Isto sugere um ajuste com  $f$  aproximadamente igual a 0.5 (correspondente a 22 pontos de cada lado do ponto a ser suavizado mais o próprio ponto, totalizando 45 pontos participando da regressão). Entretanto, é possível escolher um valor menor para  $f$  em função da análise dos resíduos ou da observação do comportamento das derivadas, já que a autocorrelação é apenas um indicativo do grau de informação que os pontos vizinhos carregam a respeito do ponto a ser ajustado, servindo para auxiliar na escolha do parâmetro de suavização.

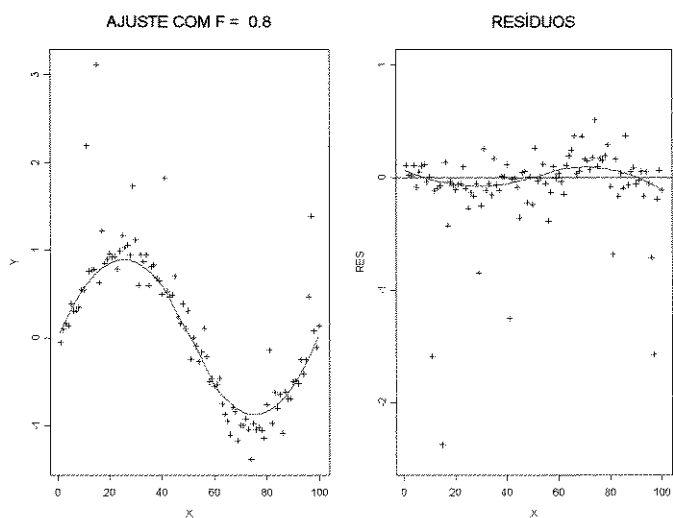


**FIGURA 16**  
**Gráfico da autocorrelação dos dados**

Uma outra forma de avaliar se a escolha de  $f$  foi acertada é observar o gráfico dos resíduos em função da variável preditora, com o ajuste da função `loess`

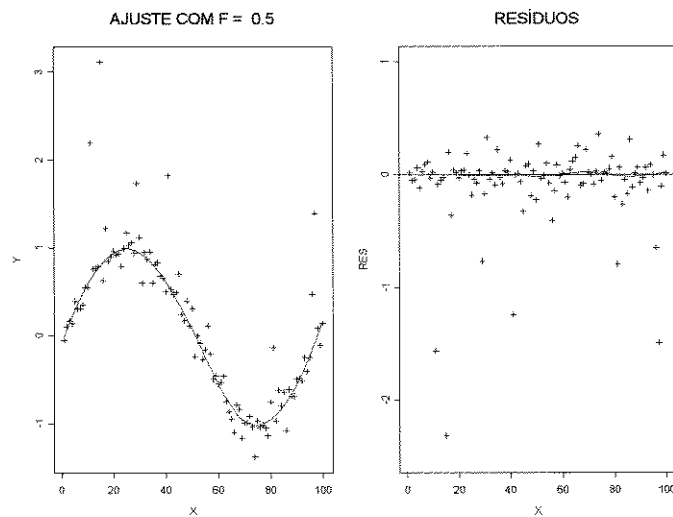
executado sobre eles. A escolha de um valor muito grande para  $f$  no ajuste dos dados brutos faz com que haja uma supersuavização, resultando em um corte que inclui parte do sinal característico do fenômeno investigado. Isso aparece no gráfico dos resíduos ajustados como uma curva que mostra haver alguma dependência dos resíduos em relação a  $X$  (CLEVELAND, 1985). A Figura 17 mostra o exemplo ajustado com  $f$  igual a 0.8 (80% dos dados fazendo parte da janela) e o gráfico dos resíduos indicando que a suavização foi muito acentuada.

Diminuindo o tamanho da janela para 0.5 (Figura 18), pode-se observar que há uma linearidade da curva do ajuste sobre os resíduos, indicando que houve uma escolha do parâmetro que possibilitou uma suavização adequada. Para saber se esse valor escolhido não foi pequeno demais, aumenta-se o seu valor ( $f=0.6$ ) até que comece a aparecer a curva no gráfico dos resíduos (Figura 19) que denuncia uma suavização excessiva. Atribui-se, então, o valor imediatamente abaixo para  $f$  como o mais indicado para aquele conjunto de dados.



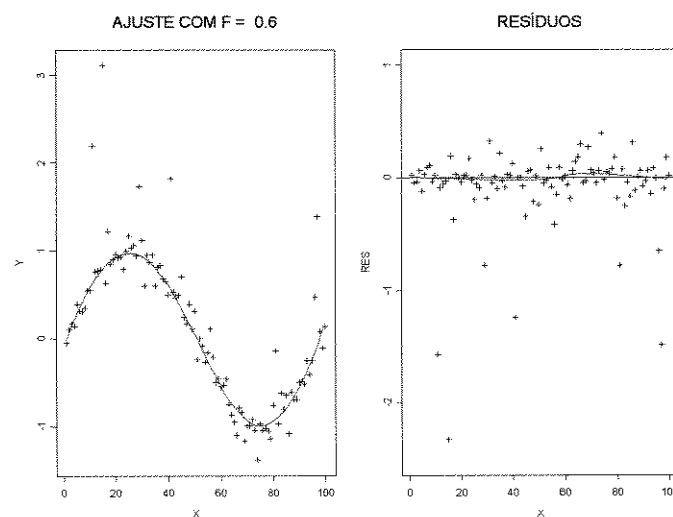
**FIGURA 17**  
**Gráfico do ajuste com  $f=0.8$  e dos resíduos com o mesmo ajuste**





**FIGURA 18**

**Gráficos do ajuste com  $f=0.5$  e dos resíduos com o mesmo ajuste**



**FIGURA 19**

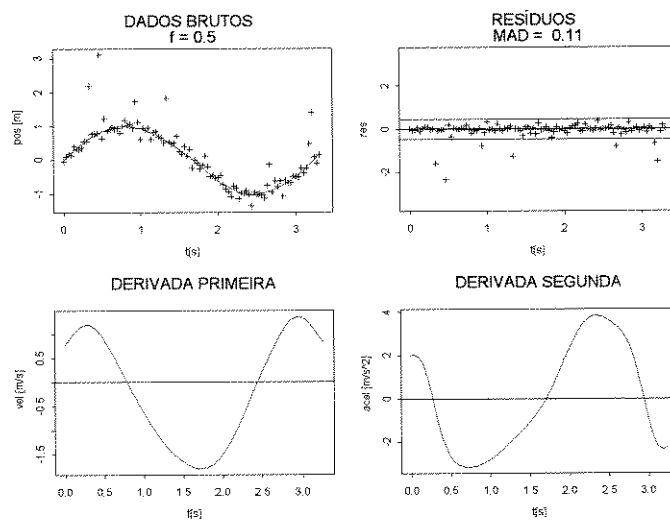
**Gráficos do ajuste com  $f=0.6$  e dos resíduos com o mesmo ajuste**

A partir da curva ajustada obtêm-se as derivadas primeira e segunda, que também podem auxiliar na determinação da qualidade do ajuste. Quando há informações sobre o comportamento das derivadas, pode-se usá-las também para

auxiliar na determinação do grau de suavização dos dados originais a fim de produzir as curvas derivadas de acordo com essas características. No exemplo da função seno contaminada com ruídos gaussianos, deseja-se excluir esses ruídos e retirar o sinal senoidal original. Como as derivadas primeira e segunda de uma função seno têm formas conhecidas, essas informações podem ser utilizadas para definir o valor do parâmetro de suavização. A Figura 20 mostra um ajuste com  $f$  igual a 0.5 no qual as derivadas apresentam o comportamento esperado.

Para o conjunto de dados exemplificado aqui, o parâmetro de suavização utilizado ( $f=0.5$ ) indica que a suavização foi apropriada. Portanto, pode-se usar essa metodologia para suavizar dados biomecânicos tanto cinemáticos quanto dinâmicos, escolhendo adequadamente o parâmetro  $f$  de acordo com a característica dos dados. É importante lembrar que a decisão final dessa escolha é do pesquisador, pois a observação do gráfico dos resíduos e das derivadas vai ser fundamental para decidir sobre a qualidade do ajuste.

Serão apresentados a seguir alguns exemplos de aplicação da função *loess* em dados biomecânicos do movimento humano, especificamente em dados cinemáticos obtidos a partir de quatro situações desportivas distintas.



**FIGURA 20**  
**Gráficos com os dados brutos, resíduos, primeira e segunda derivadas**

## EXEMPLOS DE APLICAÇÃO

Neste capítulo são apresentados quatro exemplos de aplicação do ajuste da função *loess* para dados cinemáticos em análise biomecânica do movimento humano. Para cada um deles, há um breve resumo de como os dados foram obtidos e, em seguida, são apresentados os resultados dos ajustes através de gráficos. Há também uma pequena discussão sobre o processo utilizado para cada caso. A primeira aplicação é sobre a posição angular dos membros inferiores durante uma corrida com velocidade moderada. Os outros três exemplos são de situações do futebol: uma sobre a posição dos segmentos inferiores no chute com bola parada, outro sobre a posição do goleiro durante um teste específico para treinamento e o terceiro sobre a posição de jogadores em campo durante a realização de uma partida. Cada um deles representa uma situação específica, na qual as características de cada movimento devem ser levadas em conta para que se possa obter um ajuste com boa qualidade. Em todos foram usados os procedimentos metodológicos descritos anteriormente. Nos três primeiros casos, são apresentados os gráficos de velocidade e aceleração, devido ao interesse que as situações apresentam para essas variáveis. Todos os exemplos de aplicação foram obtidos a partir de registro em vídeo e a conseqüente conversão digital para análise em *softwares* específicos (BARROS, 1997 ; STATISTICAL SCIENCES, 1993). Portanto, os dados apresentados estão na forma de série temporal, com frequência igual a 30 Hz, e a variável tempo igualmente espaçada. As variáveis dependentes são as posições dos segmentos em questão ou os ângulos dos mesmos em relação à vertical.

## APLICAÇÃO 1

O primeiro caso corresponde a uma corrida executada por um garoto com 11 anos de idade durante um trecho de aproximadamente 10 metros, durante a qual mediu-se o ângulo da perna esquerda em relação à vertical, num intervalo de tempo igual a 0.8 segundos, correspondente a um ciclo da passada (CUNHA, 1993b). Esse movimento foi registrado por uma única câmera de vídeo que se movia para cobrir o deslocamento do sujeito durante todo o percurso em questão, utilizando-se as correções de perspectiva e escala para compensar a movimentação da câmera. A Figura 21 mostra três instantes desse movimento e o ângulo ( $\alpha$ ) medido entre o segmento perna e a vertical.

Para a suavização dos dados brutos, foi usado um valor de  $f$  igual a 0.25, demonstrando, através do gráfico dos resíduos, que o parâmetro escolhido conseguiu separar o sinal do ruído, pois o gráfico apresenta uma linearidade no ajuste dos resíduos, indicando que não sobrou sinal no gráfico dos mesmos (Figura 22). Além disso, as derivadas apresentaram comportamentos que justificam a escolha de  $f$ . Esse mesmo gráfico mostra os valores discrepantes (*outliers*) dos resíduos, que são aqueles que estão posicionados acima ou abaixo das linhas horizontais determinadas pela mediana do desvio absoluto (*median absolute deviation* ou *mad*) multiplicado por seis.

Como a corrida é considerada um movimento periódico, o ajuste não paramétrico apresenta a vantagem de levar em conta possíveis variações na amplitude e na fase da variável angular devido a fatores como fadiga ou mudança no ritmo do movimento. Se fosse realizado um ajuste polinomial, por exemplo, este

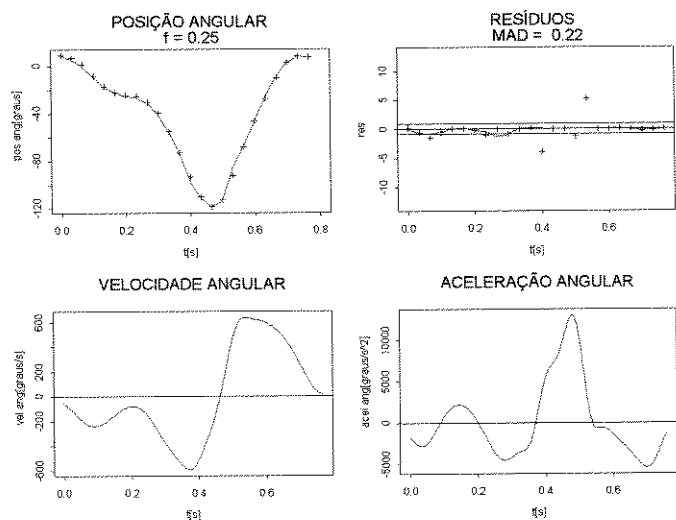
trataria todo o conjunto da mesma forma, de modo que essas alterações poderiam passar despercebidas.

O ângulo do segmento é utilizado largamente em estudos biomecânicos para a determinação de padrões motores. Pode-se, ainda, apresentar o gráfico da velocidade angular em função da posição angular, conhecido como espaço de fase (NICOLIS & PRIGOGINE), conforme ilustra a Figura 23.

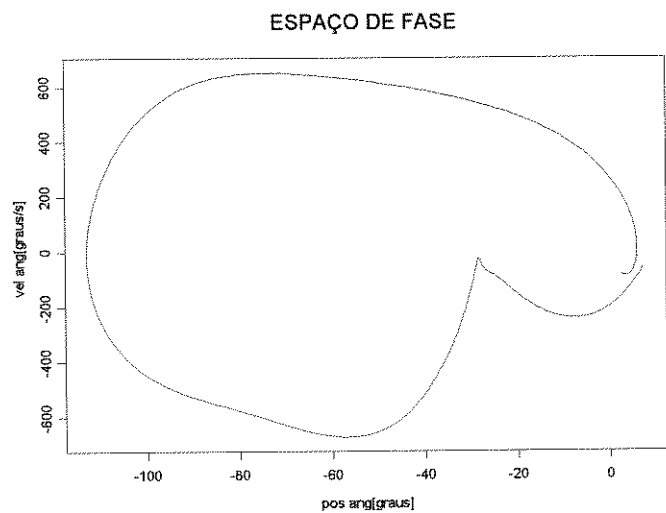
Esse tipo de representação gráfica é muito sensível às variações causadas pelos diversos tipos de ajuste que podem ser realizados. Para garantir que o espaço de fase realmente descreva o sistema envolvido, é muito importante investigar cuidadosamente o ajuste realizado, bem como a determinação de suas derivadas.



**FIGURA 21**  
**Trechos de um ciclo da corrida, com o ângulo entre a perna e a vertical**



**FIGURA 22**  
**Gráficos da posição, resíduos, velocidade e aceleração dos dados de um trecho de corrida**



**FIGURA 23**  
**Espaço de fase do movimento da corrida**



## APLICAÇÃO 2

Nessa aplicação, são usados os dados referentes a uma série de chutes realizados por um sujeito com 37 anos de idade. Esta série é constituída de cinco chutes com a bola parada, a uma distância de onze metros do gol (com as dimensões: 3 m de largura por 2 m de altura), tentando acertar a bola dentro de um arco pendurado no vértice superior esquerdo da trave. Este protocolo foi idealizado para garantir uma padronização do tipo de chute a ser realizado, forçando o sujeito a executar um chute colocado, sem muita força. O movimento foi registrado por duas câmeras de vídeo (S-VHS) posicionadas de modo a focalizar o pé de chute durante toda a sua trajetória. A análise iniciou-se no instante do contato do pé de apoio (pé esquerdo) com o solo chegando até o momento em que o coxa direita iniciou a extensão em direção ao contato do pé com o solo. Escolheu-se o maléolo lateral direito para a medição, representando a articulação do tornozelo durante a ação do chute. Esse ponto foi mensurado em ambas as câmeras, sendo feita a sua reconstrução tridimensional a partir das duas imagens sincronizadas (BARROS, 1997). A Figura 24 mostra o instante em que o sujeito coloca o pé de apoio no solo, o instante de contato do pé de chute com a bola e o instante que foi considerado o final do ciclo do chute.

O sistema de referências utilizado foi escolhido com o eixo  $x$  definido na direção em que a bola foi chutada, o eixo  $y$  coincidindo com a vertical (orientada para cima) e o eixo  $z$  ortogonal aos dois anteriores, com a sua direção e sentido definidos a partir do produto vetorial de  $x$  por  $y$ . São apresentados os pontos da coordenada horizontal ( $x$ ) em função do tempo, a fim de investigar em que instante

da ação inicia-se a frenagem na extremidade do segmento inferior. Esta investigação é muito importante na determinação de padrões motores do chute, que podem ser usados tanto para efeito de aprendizagem como para treinamento de atletas de alto nível. Além disto, havendo um conjunto de curvas representando o movimento do mesmo sujeito, pode-se analisar a sua variabilidade a partir do estudo das distâncias entre as curvas e dos conceitos de esperança e variância. Com isto, é possível classificar o nível de habilidade de sujeitos realizando o chute (ou outros fundamentos desportivos), levando em conta a posição de determinado segmento corporal em função do tempo e as suas variabilidades, analisadas como um processo estocástico.

A Figura 25 mostra os gráficos da posição horizontal do tornozelo direito em função do tempo suavizado com  $f$  igual a 0.4, dos resíduos, da velocidade e da aceleração. A linha vertical em  $t = 0.13$  s indica o instante do contato do pé com a bola.

Observa-se que a velocidade horizontal do tornozelo direito atinge o seu valor máximo imediatamente antes do contato do pé com a bola. No momento do contato, já está havendo uma desaceleração, indicando que este mecanismo de frenagem serve para que haja um encaixe adequado do pé com a bola. A escolha adequada do parâmetro de suavização garante a visualização deste fenômeno, que deve ser levado em conta durante trabalhos de ensino e treinamento em futebol.

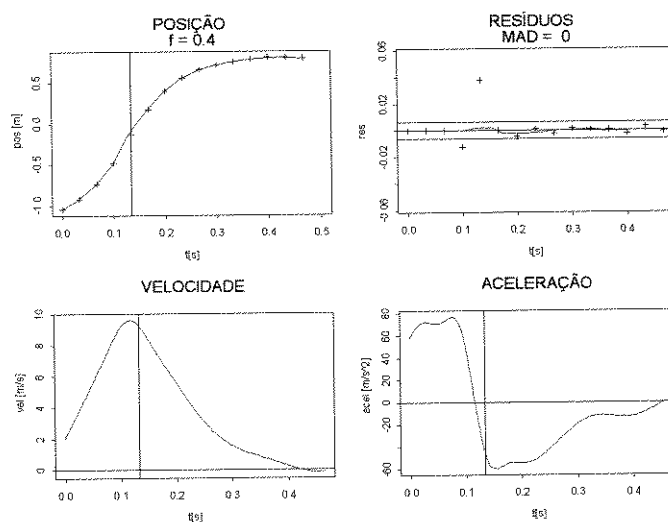
Para analisar a variabilidade dos chutes, colocam-se em um mesmo gráfico as cinco curvas (correspondentes a cada chute) ajustadas com o mesmo  $f$ , juntamente com a média (esperança) das séries temporais (Figura 26).

Nota-se qualitativamente que, principalmente no final do movimento, existe

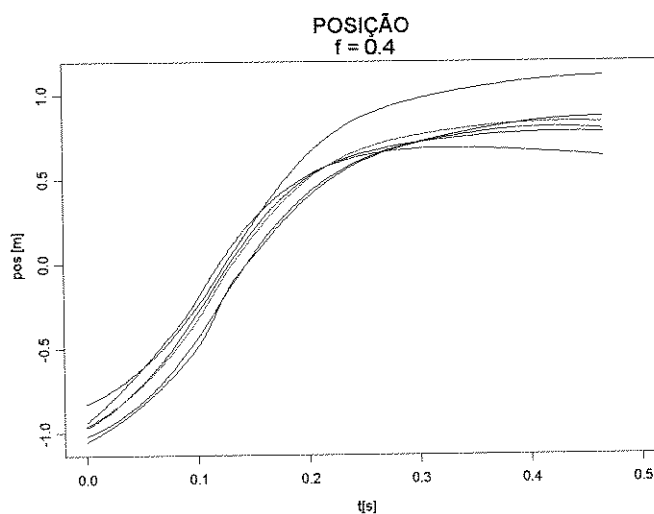
uma variação acentuada entre os cinco chutes. Isto significa que, mesmo com um único sujeito executando o mesmo tipo de chute, há uma variabilidade em função da tentativa de aprimorar a qualidade do fundamento, além da variabilidade inerente à natureza do movimento humano. Essa flutuação pode ser quantificada, medindo-se a variância dentro do espaço de funções constituído.



**FIGURA 24**  
**Seqüência de imagens de um chute no futebol**



**FIGURA 25**  
 Gráficos da posição horizontal, resíduos, velocidade e aceleração dos dados de um chute no futebol



**FIGURA 26**  
 Curvas representativas de cada um dos cinco chutes (preto) e a média (esperança) dessas séries temporais (vermelho)

### APLICAÇÃO 3

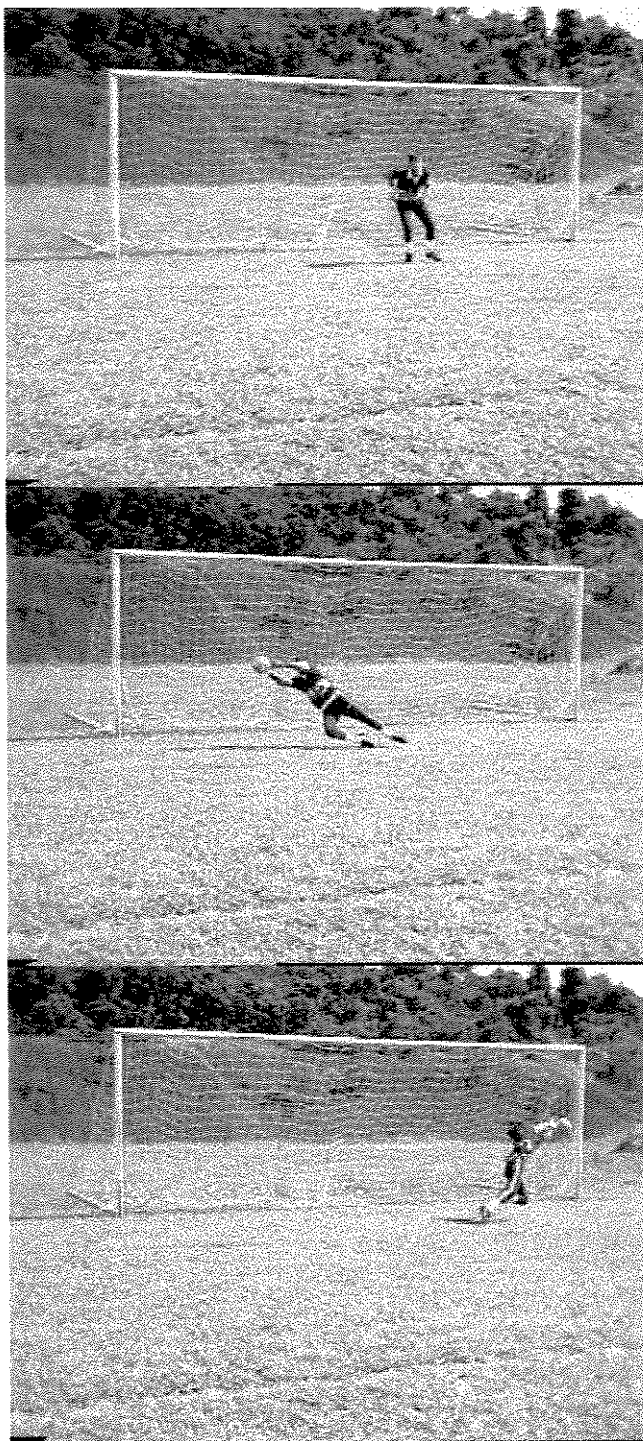
Esta aplicação consistiu de um teste para treinamento de goleiros com o intuito de verificar a relação entre variáveis biomecânicas e fisiológicas. O goleiro deveria tentar a defesa em todos os chutes realizados, com a bola sendo chutada o mais perto possível do canto inferior da trave, a uma distância de onze metros do gol, alternando os lados, com intervalo de 2.0 s entre uma bola e outra. A série só era interrompida quando o goleiro estivesse em situação de fadiga. Após a série, houve uma coleta de sangue para a análise da concentração de lactato, além da realização de outros dois tipos de testes para a determinação das variáveis fisiológicas. A movimentação do goleiro foi filmada para determinar as variáveis cinemáticas do seu centro de massa aproximado e, a partir dos dados de velocidade e aceleração, estimar os valores da potência necessária para a realização do teste, comparando-os com os valores de potência obtidos em protocolos fisiológicos. Esta aplicação atém-se aos dados cinemáticos obtidos pelo emprego da metodologia apresentada.

Para esta análise, também foram utilizadas duas câmeras, de modo que a imagem de cada uma delas cobrisse a maior área possível em que o goleiro movimentava-se durante o teste. Os procedimentos de digitalização foram seguidos pela medição da posição estimada do centro de massa do goleiro para cada quadro gerado. Esta medição iniciou-se no instante em que o goleiro retira o pé do solo para iniciar a corrida até o instante em que o mesmo termina a sétima defesa consecutiva e interrompe o teste por cansaço. Após a sincronização das câmeras e a reconstrução tridimensional do ponto estimado do centro de massa, obteve-se a

posição do sujeito testado (20 anos de idade) em função do tempo. A Figura 27 mostra a imagem de uma das câmeras quando o goleiro vai iniciar a corrida para a primeira defesa, a imagem do goleiro realizando a primeira defesa e a imagem do goleiro realizando a segunda defesa.

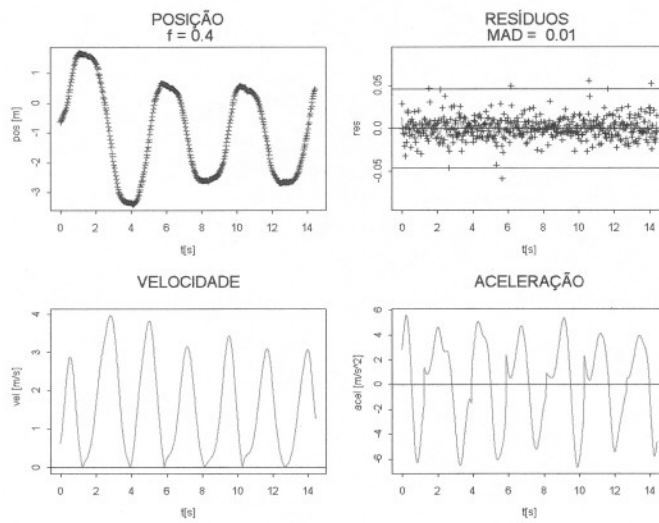
O ajuste foi executado com um valor de  $f$  igual a 0.04, com o gráfico dos resíduos e as variáveis velocidade e aceleração indicando que essa escolha foi apropriada (Figura 28).

A partir dos valores calculados das velocidades e das acelerações, podem ser calculados os valores da potência estimada para comparar com os dados fisiológicos, buscando uma relação entre a Biomecânica e a Fisiologia para testes desse tipo.



**FIGURA 27**  
**Imagens de uma seqüência de defesas durante o teste de goleiro**





**FIGURA 28**  
**Gráficos da posição horizontal, resíduos, velocidade e aceleração dos dados do deslocamento do goleiro durante teste específico**

## APLICAÇÃO 4

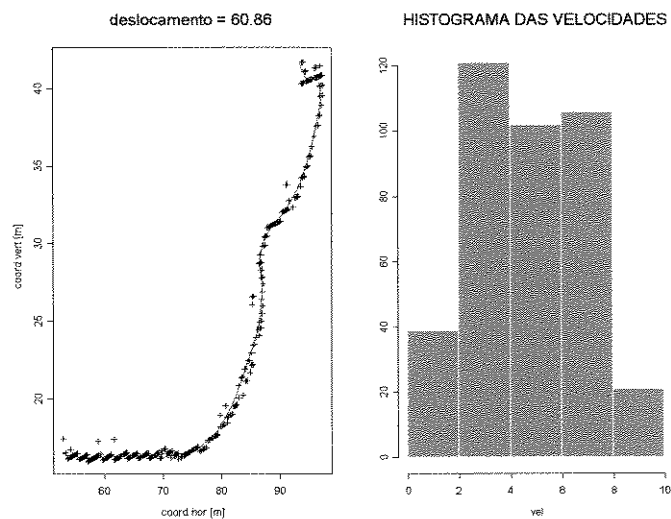
O último exemplo de aplicação é a análise do deslocamento do jogador de futebol durante a realização de uma partida. A metodologia consistiu na filmagem em vídeo (S-VHS) utilizando uma câmera para focalizar cada metade do campo, conversão analógico-digital de um trecho do jogo, medição automática da posição de cada jogador em campo, correção dos efeitos de perspectiva e escala e determinação da distância percorrida durante o período analisado (os primeiros 13 segundos do jogo). O sujeito analisado foi um jogador pertencente à equipe profissional do Guarani F.C. de Campinas, durante a realização de um jogo da primeira divisão do Campeonato Brasileiro de Futebol. A Figura 29 mostra imagens em três situações: no momento em que se inicia a análise (quando o jogador está no meio de campo), num instante intermediário, com o jogador dirigindo-se para a área adversária e no instante final, quando o jogador encontra-se próximo da linha da grande área.

A partir dos dados da posição do jogador medidos automaticamente pelo computador, já convertidos para a escala real do campo, executou-se a suavização separadamente, tanto para o comprimento como para a largura em função do tempo ( $x(t)$  e  $y(t)$  respectivamente). A partir dos dados suavizados, calculou-se a distância percorrida pelo jogador no plano do campo, através da norma quadrática dos deslocamentos horizontais e verticais. O parâmetro  $f$  usado foi igual a 0.1, demonstrando ter sido escolhido com um valor apropriado, referendado pelos gráficos apresentados na Figura 30, que também mostra um histograma das velocidades alcançadas pelo jogador no trajeto percorrido. O estudo das velocidades

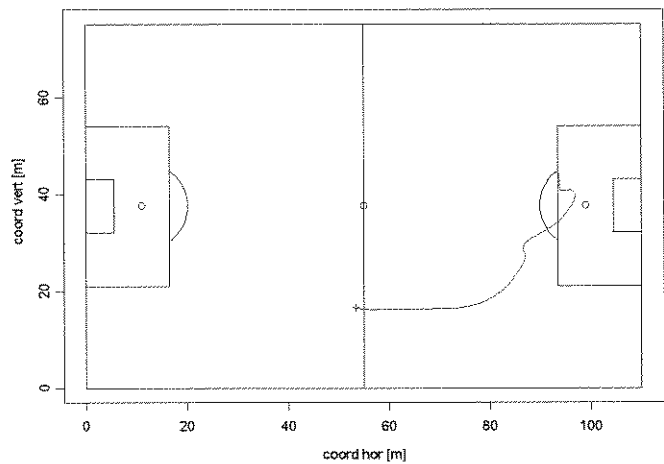


**FIGURA 29**  
**Imagens de três posicionamentos do jogador durante a partida**

que o jogador alcança durante uma partida são dados de fundamental importância na preparação do atleta. O resultado do deslocamento (60.86 m) condiz com o deslocamento real do jogador em campo, como pode ser observado na Figura 31.



**FIGURA 30**  
**Gráficos do ajuste (curva vermelha – esquerda)**  
**e histograma das velocidades (direita)**



**FIGURA 31**  
**Gráfico do deslocamento do jogador (vermelho) durante o trecho analisado**

## CONCLUSÕES

A função não paramétrica ponderada local robusta, denominada *loess*, demonstrou ser adequada para ajustar curvas contínuas e com derivadas contínuas até segunda ordem em dados cinemáticos do movimento humano, mais especificamente em situações desportivas. A escolha do parâmetro de suavização ( $f$ ) é de fundamental importância para a qualidade do ajuste, determinando o grau de suavidade que a situação experimental exige.

A implantação do procedimento robusto para evitar que os valores extremos (*outliers*) atrapalhem o procedimento de regressão local torna esta metodologia capaz de atuar nos mais diversos conjuntos de dados, quer identificando, quer corrigindo estes valores através da reponderação. A metodologia permite ainda uma interação com outras técnicas de aproximação, possibilitando o ajuste de várias funções sobre os valores suavizados pelo método robusto.

Essa função é, portanto, muito versátil, e mostrou-se capaz de ajustar curvas em dados biomecânicos com precisão e acurácia demonstradas pela análise dos resíduos, pela autocorrelação dos dados ou pela determinação das derivadas quanto há um conhecimento prévio do comportamento que tais derivadas devem assumir na situação específica.

Essa metodologia pode ser empregada em um grande número de movimentos a fim de identificar características próprias dos mesmos. Através dela, pode-se investigar a possibilidade de separar sujeitos pelo nível de habilidade na prática de determinado movimento. A análise de movimentos é uma ferramenta muito importante para o aprofundamento do conhecimento sobre alunos e atletas,

tão necessária a professores e técnicos.

Outra característica muito interessante desta metodologia é a facilidade na escolha dos parâmetros e a possibilidade de controle do sucesso dessa escolha através da observação visual de gráficos (TUKEY, 1977), tornando-a atrativa para a difusão de sua utilização em meios científicos e nas aplicações práticas.

Com tais características, esta metodologia torna-se uma contribuição importante em estudos de processos estocásticos em Biomecânica, especificamente com séries de tempo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BARROS,R.M.L. *Concepção e implementação de um sistema para análise cinemática de movimentos humanos.* Tese de doutorado - Faculdade de Educação Física - UNICAMP - Campinas, 1997.
- CLEVELAND,W.S. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association.* v.74, n.368, p.829-836, 1979.
- CLEVELAND,W.S. *The elements of graphing data.* Monterey:Wadsworth, 1985.
- CRAVEN,P. ; WAHBA,G. Smoothing noisy data with spline functions. *Numerische Mathematik*, n.31, p.377-403, 1979.
- CUNHA,C. *Métodos numéricos para as engenharias e ciências aplicadas.* Campinas:Editora da UNICAMP, 1993a.
- CUNHA,S.A. *Metodologia para análise cinemática da corrida.* Dissertação de mestrado - Faculdade de Educação Física - UNICAMP - Campinas, 1993b.
- GREEN,P.J. ; SILVERMAN,B.W. *Nonparametric regression and generalized linear models.* London:Chapman&Hall, 1994.
- HÄRDLE,W. *Smoothing techniques with implementation in S.* New York:Springer-Verlag, 1990.
- HASTIE,T.J. ; TIBSHIRANI,R.J. *Generalized additive models.* London:Chapman&Hall, 1990.
- LANCASTER,P ; SALKAUSKAS,K. *Curve and surface fitting.* London:Academic Press, 1986.
- MCLAUGHLIN,T.M. ; DILLMAN,C.J. ; LARDNER,T.J. Biomechanical analysis with cubic spline functions. *Research Quartely*, 48(3):569-582, 1977.

- MORETTIN,P.A. *Análise harmônica de processos estocásticos.* Rio de Janeiro:IMPA, 1979.
- MOSTELLER,F. ; TUKEY,J.W. *Data analysis and regression.* Massachusets:Addison-Wesley, 1977.
- MÜLLER,H.G. Nonparametric regression analysis of longitudinal data. *Lecture Notes in Statistics*, 46:15-25, 1987.
- NICOLIS,G. ; PRIGOGINE,I. *Exploring complexity.* New York:W.H.Freeman and Cia., 1989.
- PEZZACK,J.C. ; NORMAN,R.W. ; WINTER,D.A. An assessment of derivative determining techniques used for motion analysis. *Journal of Biomechanics*, 10:377-382, 1977.
- SILVERMAN,B.W. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society*, v.47, n.1, p.1-21, 1985.
- STATISTICAL Science, Statistical analysis in S-PLUS. Seattle:Mathsoft, 1993.
- WAND,M.P. ; JONES,M.C. *Kernel smoothing.* London:Chapman&Hall, 1995.
- WINTER,D.A. ; SIDWALL,H.G. ; HOBSON,D.A. Measurement and reduction of noise in kinematics of locomotion. *Journal of Biomechanics*, 7:157-159, 1974.
- WOOD,G.A. ; JENNINGS,L.S. On the use of spline functions for data smoothing. *Journal of Biomechanics*, 12:477-479, 1979.
- TUKEY,J.W. *Exploratory data analysis.* Massachusets:Addison-Wesley, 1977.
- ZERNICKE,R.F. ; CALDWELL,G. ; ROBERTS,E.M. Fitting biomechanical data with cubic spline functions. *Research Quartely*, 47(1):9-18, 1976.