ANÁLISE DE VARIÂNCIA DE AJUSTAMENTO DE CURVAS POLINOMINAIS

Cleberto Luiz Copetti
Curso de Engenharia Civil. Centro de Tecnologia. UFSM. Santa Maria, RS.
Ana Maria A. Cappelletti
Departamento de Estatística. Centro de Ciências Naturais e Exatas. UFSM. Santa Maria, RS.

RESUMO

Este trabalho aplica a análise de melhoria ao ajustamento de curvas polinomiais, feito através do método dos mínimos quadrados para regressão.

Os pontos experimentais representativos da curva em estudo foram extraídos de um movimento de salto em extensão de um atleta.

Obteve-se o polinômio de quarto grau como o modelo que melhor ajusta-se à curva estudada.

Todos os cálculos foram realizados com o auxílio do computador *IBM-360* da Universidade Federal de Santa Maria.

SUMMARY

COPETTI,C.L. and CAPPELLETTI,A.M.A., 1980. Variance analysis of adjustment of polynomial curves. Ciência e Natura (2): 11-18.

This paper presents the betterment analysis to the adjustment of polynomial curves, made by the least square method to regression.

The experimental representative points of the studied curve were given by a moviment of an extension jump of an athleta.

One got the fourth degree polinomial as the model that better adjusts the curve studied.

All calculations were realized in *IBM-360* computer from Federal University of Santa Maria.

INTRODUÇÃO

Os dados experimentalmente, nem sempre dão uma idéia precisa do fenômeno em estudo.

Com o objetivo de melhor representar este fenômeno e aproximá-lo da realidade se faz o ajustamento dos dados amostrais, visando a determinação de um modelo teórico que melhor explique o fenômeno em estudo.

Pelo método dos mínimos quadrados, pode-se determinar a equação da reta, da parábola, ou da curva de grau 'k' que melhor se ajusta a um conjunto de pontos determinados experimentalmente. Com o emprego da análise de melhoria procura-se selecionar o grau do po linômio mais adequado aos dados em estudo.

A utilização de um computador neste tipo de trabalho torna-se praticamente imprescindível na medida que há o interesse primordial na precisão dos resultados, além de serem necessários cálculos repetitivos e trabalhosos.

FUNDAMENTAÇÃO TEÓRICA E MÉTODO

No método de interpolação de Newton, conforme HORNBECK(5), dado um conjunto de pontos não ajustados, tem-se, condições de encontrar uma equação que represente estes pontos e também de interpolar ou mesmo extrapolar pontos com base na referida equação.

Sob o ponto de vista do ajustamento, este método não confere a eficiência desejada, posto que a equação encontrada representa as coordenadas não devidamente ajustadas.

Há, então, a necessidade de se estudar outro método que melhor ajuste o conjunto de pontos observados.

COSTA NETO (2) refere que os pontos experimentais admitem uma relação funcional entre suas projeções X e Y, e vice versa, responsáveis pelo aspecto do diagrama, qual seja a tendência assumida pelo conjunto de pontos. Este relacionamento funcional corresponderia a linha existente na Figura 1, que seria a "linha de regressão".

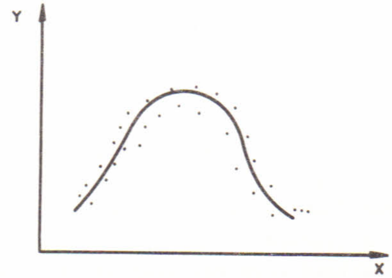

Figura 1. Trajetória do centro de gravidade.

O método dos mínimos quadrados, segundo PACITTI,(6), além de determinar uma equação que represente este conjunto de pontos, corrige-os deixando de lado aqueles que não estão enquadrados na tendência normal da curva.

No caso de não conhecer-se de antemão o modelo que rege o relacionamento de uma variável com outra, deve-se testar várias equações e utilizar aquela que melhor representa os pontos experimentais.

No caso da curva não obedecer apenas a uma lei, é necessário dividir a curva total em semicurvas, de maneira tal que cada semicurva tenha apenas um máximo ou um mínimo. Pode-se, neste caso, estabelecer os limites de cada semicurva como sendo os pontos de inflexão entre as semicurvas.

Muitas vezes, segundo COSTA NETO (2), o segmento retilíneo pode ser encarado como aproximação grosseira para uma função mais suave de um conjunto de pontos.

Um polinômio de grau 'p' pode constituir-se em uma melhor aproximação aos dados.

Considerando-se um conjunto de 'n' pares de dados, tem-se para um modelo linear:

$$Y(x)=K_1+K_2x$$

e para um modelo polinomial de grau 'p', tem-se:

$$Y(x)=K_1+K_2x+K_3x^2+\ldots+K_mx^p$$

Com esta equação expressa a soma dos quadrados dos resíduos:

$$H=\sum_{i=1}^{n}\left[Y(xi)-Yi\right]^2$$

$$H=\sum_{i=1}^{n}\left[K_1+K_2xi+\ldots+K_jxi^{j-1}+\ldots+K_mxi^p-Yi\right]^2$$

Constrói-se as 'm' equações normais, determinando-se as 'm' derivadas parciais, e após iguala-se a zero. Isto para estimar-se os 'K_j'. Assim, tem-se:

$$\frac{\partial H}{K_1}=\sum_{i=1}^{n}2\left[K_1+K_2xi+\ldots+K_mxi^p-\hat{Y}i\right]=0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{\partial H}{K_j}=\sum_{i=1}^{n}2xi^{j-1}\left[K_1+K_2xi+\ldots+K_mxi^p-\hat{Y}i\right]=0$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{\partial H}{K_m}=\sum_{i=1}^{n}2xi^p\left[K_1+K_2xi+\ldots+K_mxi^p-\hat{Y}i\right]=0$$

Para a resolução do sistema de 'm' equações, isola-se as constantes 'K_j', com $j=1,\ldots m$.

Expressando-se em forma matricial, tem-se:

$$\begin{bmatrix} n & \Sigma xi & \Sigma xi^2 & \ldots & \Sigma xi^{j-1} & \ldots & \Sigma xi^p \\ \Sigma xi & \Sigma xi^2 & \Sigma xi^{j-1} & \ldots & \ldots & \ldots & \Sigma xi^{p-1} \\ . & .. & \ldots & \ldots & \ldots & \ldots & \ldots \\ \Sigma xi^p & \Sigma xi^{p-1} & \ldots & \ldots & \ldots & \ldots & \Sigma xi^{2p} \end{bmatrix} \begin{bmatrix} K_1 \\ K_2 \\ .. \\ K_m \end{bmatrix} \begin{bmatrix} \Sigma Yi \\ \Sigma xiyi \\ .. \\ \Sigma xi^p yi \end{bmatrix}$$

Emprega-se, neste artigo, o método de eliminação de Gauss para resolver as equações normais.

É válido salientar que se este sistema for resolvido por inversão de matrizes, a matriz das equações normais torna-se mal condicionada quando cresce o grau do polinômio, pois o determinante tende a zero com o aumentar do grau do polinômio.

Análise de Melhoria

A análise de melhoria é feita através da análise de variança aplicada à regressão, com objetivo de determinar o grau do polinômio que melhor se ajusta ao conjunto de dados.

Com a análise de melhoria busca-se equações que representem os pontos com significância maior em relação ao modelo anteriormente testado.

Segundo GRAYBILL (4), ao procurar-se a equação do polinômio que melhor se ajusta ao fenômeno em estudo, inicia-se com o modelo linear, isto é, com a reta de regressão. Após, testa-se se a parábola dá uma melhoria significativa em relação à reta. Em caso afirmativo, testa-se a equação de regressão cúbica em relação à parábola. E assim sucessivamente até que duas etapas sucessivas não apresentem melhoria significativa ao nível de significância adotado.

Testa-se para cada modelo a hipótese H0, de não haver melhoria significativa ao nível de significância adotado. Então, aceita-se a hipótese H0, se o 'F' calculado pela análise de variança for menor que o 'F' dado pelas tabelas F de Snedecor.

Monta-se uma tabela para cada modelo, a fim de calcular o valor de 'F'. A seguir compara-se com o valor de 'F' tabelado. Caso $F>F\alpha$, testa-se o próximo modelo; caso contrário aceita-se o modelo anterior.

Quando testa-se a melhoria de um polinômio de grau 'K' em relação à outro de grau 'K-1', é possível dividir-se a variação residual em torno da curva polinomial de grau 'K-1' em duas partes:

Uma explicada pela variação residual em torno da equação polinomial de grau 'K'; outra devida a melhoria de ajuste.

Geometricamente, conforme a Figura 2, tem-se:

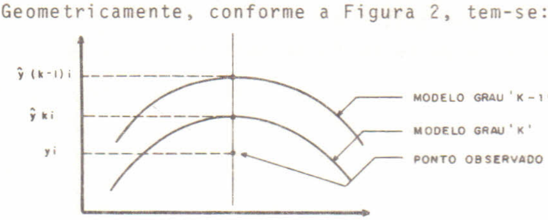

Figura 2. Variação residual.

Então, pode-se equacionar da seguinte maneira:

$$\underbrace{\sum_{i=1}^{n} (Yi-\hat{Y}(K-1)i)^2}_{1} = \underbrace{\sum_{i=1}^{n} (\hat{Y}Ki - \hat{Y}(K-1)i)^2}_{2} + \underbrace{\sum_{i=1}^{n} (Yi - \hat{Y}Ki)^2}_{3}$$

Sendo:

1. Variação residual em torno do polinômio de grau 'K-1'

2. Melhoria de ajuste

3. Variação residual em torno do polinômio de grau 'K'

Utilização da Soma de Erros Quadrados como Avaliação

Utilizou-se a soma de erros quadrados na medida da oscilação dos pontos em torno da 'linha de regressão'.

A soma de erros quadrados pode ser equacionada como:

$$S^2 = \sum_{i=1}^{n} \left[Y(i) - \hat{Y}(i)\right]^2$$

Onde:

$Y(i)$ é o valor de y observado

$\hat{Y}(i)$ é o valor de y estimado

MATERIAL

Os dados experimentais desta pesquisa foram retirados de ÁVILA *et alii* (1), onde a coleta de dados foi feita utilizando o método da cronociclografia.

A curva em estudo é representativa da trajetória do centro de gravidade do corpo de um atleta durante a execução de um salto em extensão.

Para o ajustamento dos pontos utilizou-se o método dos mínimos quadrados para polinômios. O programa computacional de ajuste dos dados fundamentou-se em GOTTFRIED (3).

Foi montado, também, um programa para calcular a análise de melhoria para cada grau de polinômio testado.

Os cálculos foram realizados por meio do computador IBM-360 existente no Núcleo de Processamento de Dados da Universidade Federal de Santa Maria. A falta de um traçador de gráficos contínuos (*plotter*) não permitiu, entretanto, melhor representação da correção da curva.

RESULTADOS E DISCUSSÃO

Com base nas somas de erros quadrados extraídas do programa de ajustamento, montou-se a Tabela I, utilizada no cálculo da análise de melhoria. O valor crítico de 'F' retira-se das tabelas de distribuição 'F' de Snedecor.

TABELA I. SOMA DE ERROS QUADRADOS, SEGUNDO O GRAU DO POLINÔMIO.

GRAU DO POLINÔMIO	SOMA DE ERROS QUADRADOS
1	4,179000
2	0,038800
3	0,011010
4	0,003759
5	0,003575
6	0,000750
7	0,000749

Como pode-se notar nas Tabelas II, III, IV e V o polinômio que melhor ajusta-se ao conjunto de dados é o polinômio de quarto grau, já que na análise de quinto grau não houve melhoria ao nível de significância adotado.

TABELA II. ANÁLISE DE MELHORIA DO POLINÔMIO DE 2º GRAU EM RELAÇÃO AO POLINÔMIO DE 1º GRAU.

FONTES DE VARIAÇÃO	G.L.	S.Q.	Q.M.	F	Fα
Melhoria de Ajuste	1	4,140200	4,140200	5335	3,33
Resíduo para Polinômio 2º grau	50	0,038800	0,000776		
Resíduo para Polinômio 1º grau	51	4,179000			

TABELA III. ANÁLISE DE MELHORIA DO POLINÔMIO DE 3º GRAU EM RELAÇÃO AO POLINÔMIO DE 2º GRAU.

FONTES DE VARIAÇÃO	G.L.	S.Q.	Q.M.	F	Fα
Melhoria de Ajuste	1	0,027790	0,027790	123,51	3,27
Resíduo para Polinômio 3º grau	49	0,011010	0,000225		
Resíduo para Polinômio 2º grau	50	0,038800			

TABELA IV. ANÁLISE DE MELHORIA DO POLINÔMIO DE 4º GRAU EM RELAÇÃO AO POLINÔMIO DE 3º GRAU.

FONTES DE VARIAÇÃO	G.L.	S.Q.	Q.M.	F	Fα
Melhoria de Ajuste	1	0,007251	0,007251	92,96	3,20
Resíduo para Polinômio 4º grau	48	0,003759	0,000078		
Resíduo para Polinômio 3º grau	49	0,011010			

TABELA V. ANÁLISE DE MELHORIA DO POLINÔMIO DE 5º GRAU EM RELAÇÃO AO POLINÔMIO DE 4º GRAU.

FONTES DE VARIAÇÃO	G.L.	S.Q.	Q.M.	F.	Fα
Melhoria de Ajuste	1	0,000184	0,000184	2,42	3,13
Resíduo para Polinômio 5º grau	47	0,003575	0,000076		
Resíduo para Polinômio 4º grau	48	0,003759			

A Figura 3 é representativa da curva ajustada como polinômio do 4º grau da trajetória do centro de gravidade do corpo, na execução de um salto em extensão.

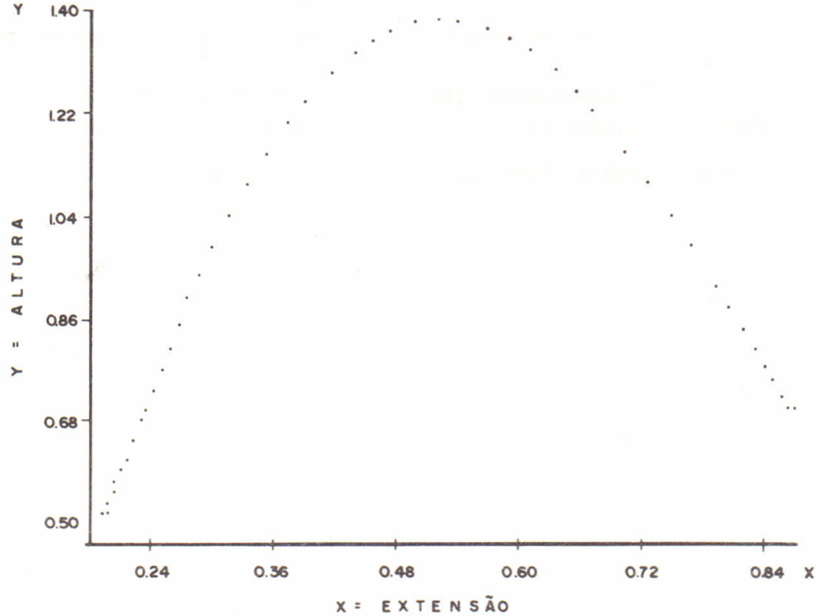

Figura 3. Curva ajustada 4º grau - trajetória do centro de gravidade. Salto e extensão.

CONCLUSÃO

Pelo método dos mínimos quadrados e análise de melhoria, o polinômio do quarto grau foi o que melhor se ajustou aos dados referentes ao movimento em estudo, relativo ao salto em extensão de um atleta.

BIBLIOGRAFIA CITADA

1. ÁVILA,A.O.V.; BAECKER,I.M.; KREBS,R.J.; SAMPEDRO,R. Trabalho de Conclusão do Curso de Especialização em Biomecânica. *Revista do Centro de Educação Física e Desportos*. Santa Maria, 1979, 40p.
2. COSTA NETO,P.O. Estatística. São Paulo, Ed. Edgar Blucher Ltda, 1977, 268 p.
3. GOTTFRIED,B.S. *Programming with Fortran*. New York, Ed. Quantum Publisher Inc., 1972, 272 p.
4. GRAYBILL,F.A. *An Introduction to Linear Model*. New York, Ed. Mc Graw Hill, 1961, 463 p.
5. HORNBECK,R.W. *Numerical Methods*. New York, Ed. Quantum Publisher

Inc., 1975, 310 p.

6. PACITTI,T. & ATKINSON,C.P. *Programação e Métodos Computacionais*. Rio de Janeiro, Ed. Livros Técnicos e Científicos, 1977,386 p.
7. SOUZA DIAS,D. *Programação Fortran para Estudantes de Ciências e Engenharia*. Rio de Janeiro, Ed. Livros Técnicos e Científicos, 1977, 258 p.
8. SPIEGEL,M.P. *Estatística*. São Paulo, Ed. Mc Graw Hill, 1976, 506p.
9. YAMANE,T. *Estatística*. México, Ed. Harper & Row, 1967, 573 p.

Recebido em novembro, 1980; aceito em dezembro, 1980.