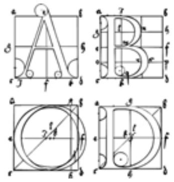
Curve Fitting: Fertilizer, Fonts, and Ferraris



Curve Fitting: Fertilizer, Fonts, and Ferraris

What's the difference between modeling and curve fitting, and what are polynomials used for, anyway?

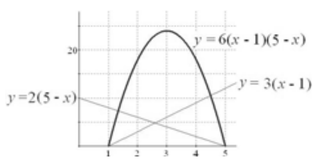
32nd AMATYC Annual Conference
November 3, 2006
Cincinnati, Ohio

Katherine Yoshiwara
Bruce Yoshiwara
Los Angeles Pierce College

Typical Quadratic Models

- Projectile and other motion problems from physics
- Problems involving area or the Pythagorean Theorem
- Revenue curves:
  Total revenue = (number of items)(price per item)

Some quadratic models arise as the product of two linear functions.



Revenue from theater tickets

A small theater troupe charges \$5 per ticket and sells 100 tickets at that price.

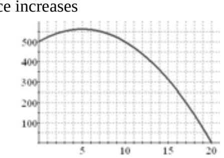
On subsequent nights they decide to increase the price by \$.25 at a time. For each \$.25 increase in price, they sell 10 fewer tickets.

Revenue from theater tickets

x = number of \$.25 price increases

p = price per ticket
$= 5.00 - .25x$

n = number of tickets
$= 100 + 10x$



Revenue = $p\,n$
$= (5.00 - .25x)(100 + 10x)$

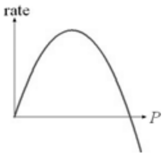
Rate of growth in a logistic model

$$dP/dt = kP\ (L - P)$$

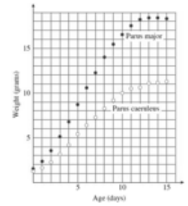
where P is the population at time t, and L is the carrying capacity.

Or, for classes before calculus,

$$r = kP\ (L - P)$$



Logistic growth



The figure shows the typical weight of two species of birds each day after hatching.

Compute the daily rate of growth for each species.

Logistic growth

Parus Major

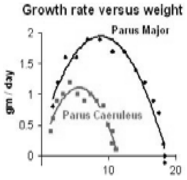
Day	0	2	4	6	8	10	12	14
Weight	1.6	3.6	6.8	10.6	14	16.6	18.2	18.4
Growth rate	0.8	1.6	1.9	1.7	1.4	0.9	0.2	-0.1

Parus Caeruleus

Day	0	2	4	6	8	10	12	14
Weight	1.3	2.3	4.2	6.4	8.3	10	10.7	11.2
Growth rate	0.4	0.9	1.2	0.9	0.9	0.5	0.4	0.1

Logistic growth (continued)

For each species, plot the rate of growth against weight in grams. What type of curve does the growth rate graph appear to be?



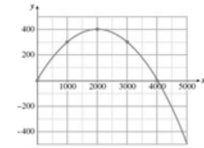
Maximum sustainable yield

Commercial fishermen rely on a steady supply of fish in their area. To avoid overfishing, they adjust their harvest to the size of the population. The equation

$$r = 0.0001x\ (4000 - x)$$

gives the annual rate of growth, in tons per year, of a fish population of biomass x tons.

Rate of growth of a fish population

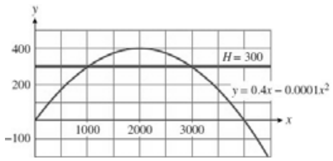


What is the significance of the intercepts in terms of the fish population?

What is the significance of the vertex?

For what values of x does the fish population decrease rather than increase?

Maximum sustainable yield



Suppose that 300 tons of fish are harvested each year. What sizes of biomass will remain stable from year to year?

Models for traffic flow

$$r = d\ s$$

traffic flow = (traffic density) (average speed)

cars/hour = (cars/mile) (miles/hr)

Greenshield's model for traffic flow

Assumes that the average speed s of cars on a highway is a linear function of traffic density

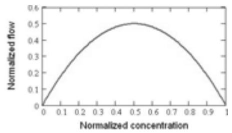
$$s = s_f\,(1 - d\ /\ d_j)$$

where s_f is the free-flow speed and d_j is the maximum (jam) density.

Greenshield's model for traffic flow

Traffic flow is a quadratic function of d, given by

$$r = ds = d \cdot s_f\,(1 - d\ /\ d_j)$$



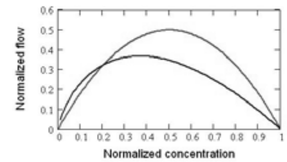
Greenshield's model for traffic flow

$$s = s_f\,(1 - d\ /\ d_j)$$

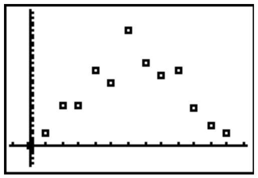
Greenberg's model for traffic flow

$$s = (s_f\,/2)\ \ln\,(d_j\,/d)$$

Two models for traffic flow



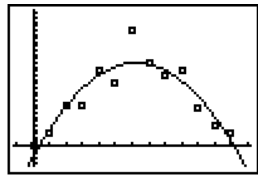
Mad cow disease



Annual deaths in the UK from vCJD (mad cow disease) from 1994 to 2006.

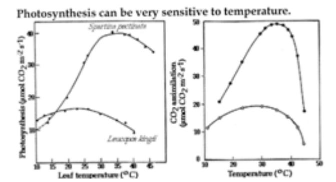
http://www.cjd.ed.ac.uk/vcjdqjun06.htm

Mad cow disease



$y = -0.57x^2 + 7.14x - 2.26$

Photosynthesis

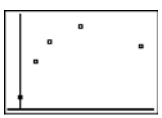


Is photosynthesis a quadratic function of temperature?

http://biology.uwsp.edu/faculty/esingsaa/reference/lecture5.5/lftemp.htm

Soy bean yield as a function of fertilizer

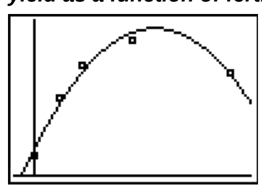
N rates (kg/ha)	Yield (ton/ha)
0	2.12
15	2.46
30	2.65
60	2.8
120	2.6



Is crop yield quadratic in fertilizer rate?

http://www.arc-avrdc.org/pdf_files/Tuxnewen(8-N).pdf

Crop yield as a function of fertilizer



$y = 2.158 + 0.019x - 0.000132x^2$

But are these models or just examples of curve-fitting?

Two types of models

Mechanistic models provide insight into the chemical, biological, or physical process thought to govern the phenomenon under study. The parameters derived are estimates of real system properties.

Empirical models simply describe the general shape of the data. The parameters do not necessarily correspond to a chemical or physical process. Empirical models may have little or no predictive value.

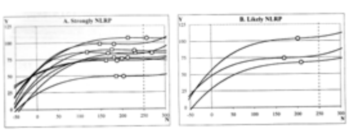
Choosing a model: A quote from GraphPad Software

“Choosing a model is a scientific decision. You should base your choice on your understanding of chemistry or physiology (or genetics, etc.). The choice should not be based solely on the shape of the graph.”

“Some programs...automatically fit data to hundreds or thousands of equations and then present you with the equation(s) that fit the data best... You will not be able to interpret the best-fit values of the variables, and the results are unlikely to be useful for data analysis”

(*Fitting Models to Biological Data Using Linear and Nonlinear Regression*, Motulsky & Christopoulos, GraphPad Software, 2003)

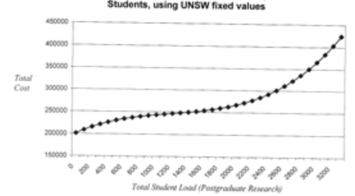
Crop yield as a cubic function of fertilizer rate



http://ag.arizona.edu/AREC/pubs/researchpapers/2005-02beattieetal.pdf

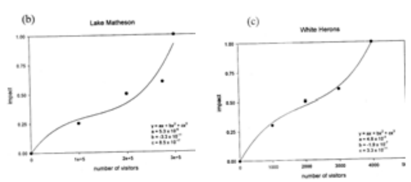
Cost function for higher education in Australia

Figure 1. Total Cost Curve for Postgraduate by Research Students, using UNSW fixed values



http://www.melbourneinstitute.com

Visitor impact at tourist sites in New Zealand



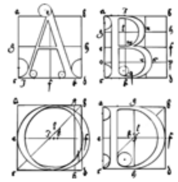
http://www.landcareresearch.co.nz/research/sustain_business/tourism/documents/tourist_flow_data.pdf

Polynomial curve fitting

Although higher-degree polynomials typically do not provide meaningful models, they are useful for approximating continuous curves.

Polynomials are easy to evaluate, their graphs are completely smooth, and their derivatives and integrals are again polynomials.

Font design

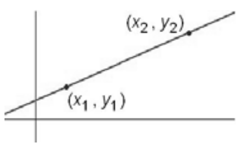


Lagrange interpolation

Given any n + 1 points in the plane with distinct x-coordinates, there is a polynomial of degree at most n whose graph passes through those points.

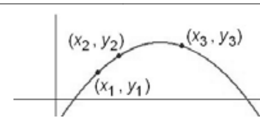
Lagrange polynomial through 2 points

$$y=\frac{x-x_2}{x_1-x_2}\cdot y_1+\frac{x-x_1}{x_2-x_1}\cdot y_2$$



Lagrange polynomial through 3 points

$$y=\frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)}\cdot y_1+\frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)}\cdot y_2+\frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}\cdot y_3$$



Lagrange polynomial through n+1 points

$$\sum_{0\le i\le n} y_i \prod_{j\ne i} \frac{x-x_j}{x_i-x_j}$$

(Excel demo)

Osculating polynomials

An *osculating polynomial* agrees with the function and all its derivatives up to order m at n points in a given interval.

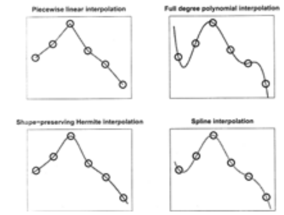
Hermite polynomials are osculating polynomials of order $m = 1$, that is, they agree with the function and its first derivative at each point.

Piecewise interpolation

Many of the most effective interpolation techniques use piecewise cubic Hermite polynomials.

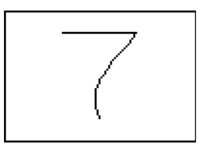
There is a trade-off between smoothness and local monotonicity or shape-preservation.

Piecewise polynomials fits



http://www.mathworks.com/moler/interp/pdf

Parametric approximations



To approximate more general curves, we can use parametric equations.

Bezier curves

Bezier curves are the most frequently used interpolating curves in computer graphics.

They were developed in the 1960s by Paul de Casteljau, an engineer at Citroen, and independently by Pierre Bezier at Renault.

Linear Bezier curves

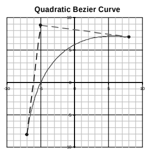
The linear Bezier curve through two points P_0 and P_1 is defined by

$$P(t) = (1 - t)\, P_0 + t\, P_1, \qquad 0 \le t \le 1$$

It is just the line segment joining P_0 and P_1.

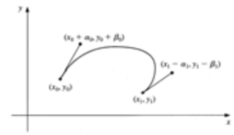
Quadratic Bezier curves

A quadratic Bezier curve is defined by two endpoints, P_0 and P_2 , and a control point P_1.



$$P(t) = (1 - t)^2 P_0 + 2t\,(1 - t)\,P_1 + t^2 P_2$$

Hermite curve with 2 control points

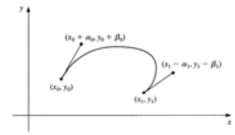


We want x (and y) to be cubic in t

$$x(t) = at^3 + bt^2 + ct + d$$

satisfying $x(0) = x_0$, $x(1) = x_1$, $x'(0) = \alpha_0$, $x'(1) = \alpha_1$

Hermite curve with 2 control points



$$x = [2(x_0 - x_1) + (\alpha_0 + \alpha_1)]\,t^3 + [3(x_1 - x_0) - (\alpha_1 + 2\alpha_0)]\,t^2 + \alpha_0\,t + x_0$$

$$y = [2(y_0 - y_1) + (\beta_0 + \beta_1)]\,t^3 + [3(y_1 - y_0) - (\beta_1 + 2\beta_0)]\,t^2 + \beta_0\,t + x_0$$

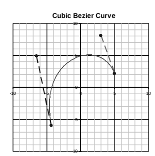
Curves with 2 control points

$$x = at^3 + bt^2 + ct + d$$

Hermite	Bezier
$a = 2(x_0 - x_1) + (\alpha_0 - \alpha_1)$	$a = 2(x_0 - x_1) + \mathbf{3}(\alpha_0 - \alpha_1)$
$b = 3(x_1 - x_0) - (\alpha_1 + 2\alpha_0)$	$b = 3(x_1 - x_0) - \mathbf{3}(\alpha_1 + 2\alpha_0)$
$c = \alpha_0$	$c = \alpha_0$
$d = x_0$	$d = x_0$

Cubic Bezier curves

A cubic Bezier curve can be defined by two endpoints, P_0 and P_3 , and control points P_1. and P_2 as follows.



$$P(t) = (1 - t)^3 P_0 + 3t\,(1 - t)^2\,P_1 + 3t^2\,(1 - t)P_2 + t^3 P_3$$

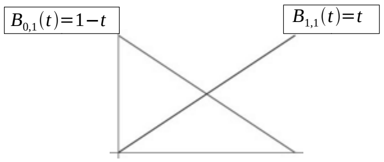
Bernstein polynomials

The Bernstein polynomials of degree n are defined by

$$\boxed{B_{i,n}(t) = \binom{n}{i} t^i (1-t)^{n-i}}$$

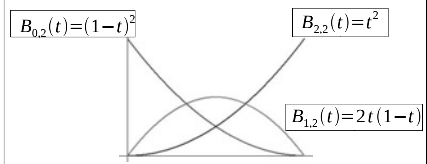
Bernstein polynomials of degree 1

The Bernstein polynomials of degree 1 are



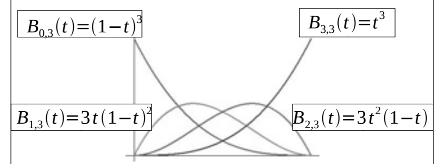
Bernstein polynomials of degree 2

The Bernstein polynomials of degree 2 are



Bernstein polynomials of degree 3

The Bernstein polynomials of degree 3 are



Bernstein polynomials

- Form a basis for polynomials of degree *n*.
- Form a partition of unity, that is, the sum of the Bernstein polynomials of degree *n* is 1.
- When a Bezier polynomial is expressed in terms of the Bernstein basis, the coefficients of the basis elements are just the points P_0 through P_n.

Download presentation and handouts

www.piercecollege.edu/faculty/yoshibw/Talks/amatyc06/