MCEN 3030

22 Feb 2024

HW & Quiz #4 Due Monday

Exam #1 Tuesday 2/27 in class

- accommodations email after class

Last time: (Linear) Least Squares

Today: Nonlinear Models

Last time: Linear models (and → more variables)

Data:

x_b	x_c	y
3.1	0.9	19.5
$\vdots$	$\vdots$	$\vdots$

$$\longrightarrow \quad \begin{aligned} y_1 &= a + b x_{b1} + c x_{c1} \\ y_2 &= a + b x_{b2} + c x_{c2} \\ y_3 &= a + b x_{b3} + c x_{c3} \\ &\cdots \end{aligned}$$

Call this $\underset{\substack{\uparrow \\ \text{called this} \\ y}}{Y} = Z \underset{\substack{\uparrow \\ \text{called this} \\ a}}{A}$

$$\longrightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & x_{b1} & x_{c1} \\ 1 & x_{b2} & x_{c2} \\ \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

This is an overdetermined system ⇒ no solution

Instead: Rearrange to $E = Y - ZA$, where E is vector of "residuals", e.g. $e_1 = y_1 - (a + b x_{b1} + c x_{c1})$

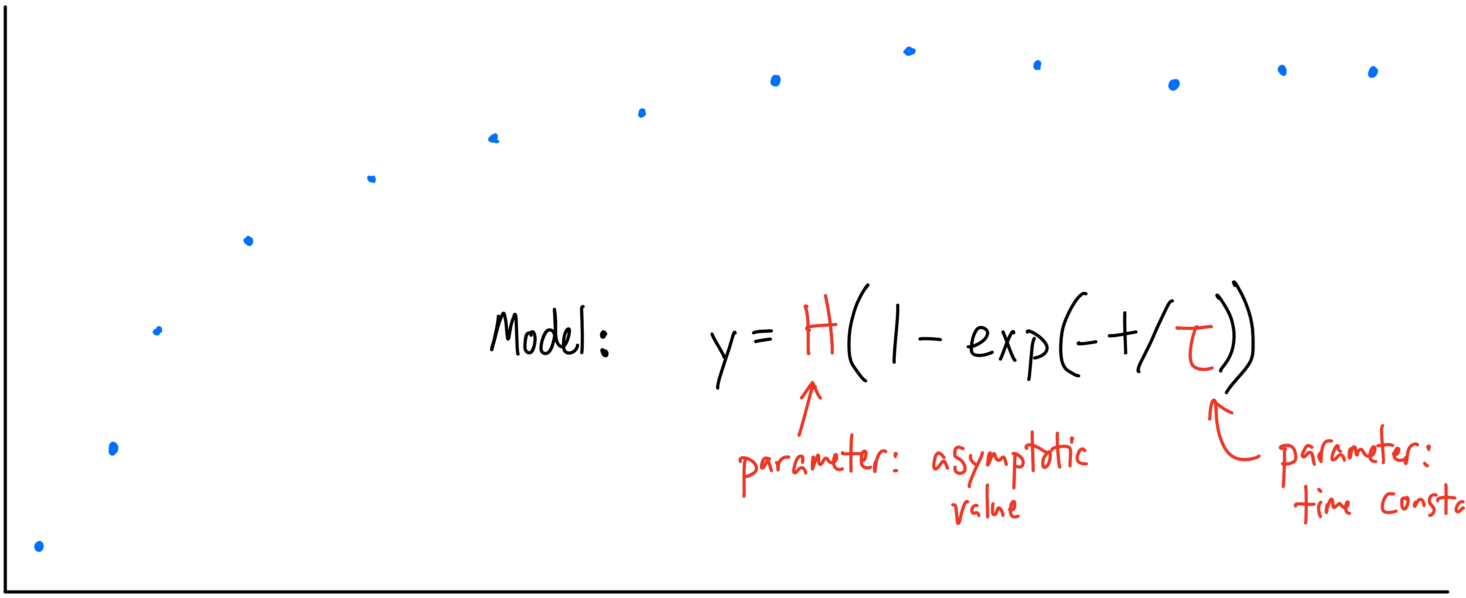
Summary of some linear algebra calculus:

$E^T E = S_r$ = the sum of the residuals (squared)

$$\frac{dS_r}{dA} = 0 \longrightarrow \boxed{A = (Z^T Z)^{-1} (Z^T Y)}$$

explicit equation for best fits $A \equiv \begin{bmatrix} a \\ b \\ c \\ \vdots \end{bmatrix}$

Least-Squares Regression w/ nonlinear models



Model: $y = \underset{\substack{\uparrow \\ \text{parameter: asymptotic} \\ \text{value}}}{H}\left(1-\exp\left(-t/\underset{\substack{\uparrow \\ \text{parameter:} \\ \text{time constant}}}{\tau}\right)\right)$

We can still define

Residual: $e_i \equiv y_i - \hat{y}(t_i) = y_i - H\left(1-\exp\left(-t_i/\tau\right)\right)$

Sum of Residuals (squared):

$$S_r = \sum\left(\underset{\substack{\uparrow \\ \text{known}}}{y_i} - \overset{\text{TBD}}{H}\left(1-\exp\left(-\underset{\substack{\uparrow \\ \text{known}}}{t_i}/\underset{\substack{\uparrow \\ \text{TBD}}}{\tau}\right)\right)\right)^2$$

$$\frac{\partial S_r}{\partial H} = 0 = \sum 2\left(y_i - \boxed{H}\left(1-\exp\left(\frac{-t_i}{\boxed{\tau}}\right)\right)\right)(-1)\left(1-\exp\left(\frac{-t_i}{\boxed{\tau}}\right)\right)$$

$$\frac{\partial S_r}{\partial \tau} = 0 = \sum 2\left(y_i - \boxed{H}\left(1-\exp\left(\frac{-t_i}{\boxed{\tau}}\right)\right)\right)\left(\boxed{H}\exp\left(\frac{-t_i}{\boxed{\tau}}\right)\left(\frac{t_i}{\boxed{\tau^2}}\right)\right)$$

→ We are unable to frame this as a linear system

2D Taylor series (for $f(\underset{\substack{\uparrow \\ \text{independent} \\ \text{variable}}}{t}; \underbrace{H, \tau}_{\text{parameters}})$)

$$f(H,\tau) \approx f(H_0,\tau_0) + \left.\frac{\partial f}{\partial H}\right|_{(H_0,\tau_0)} \cdot \Delta H + \left.\frac{\partial f}{\partial \tau}\right|_{(H_0,\tau_0)} \cdot \Delta \tau$$

$$+ \left(\text{higher-order terms}\right)$$

→ We have made $f(t; H, \tau)$ "look linear" near to some reference values (H_0, τ_0)

Idea: • Guess values for (H, τ): (H_0, τ_0)

- Use this guess to write a <u>linear</u> approximation for $f(t; H, \tau)$.
  - Importantly, the coefficients in this linearization depend on (H_0, τ_0), but are "known"
- Linears system → borrow ideas from last lecture
- Iterate, updating our guess until the parameter values converge.

Let's see how we can use this for a 1-parameter model, and then extrapolate. $f(t;\tau)$

Residual: $e_i \equiv y_i - f(t_i;\tau)$

$$= y_i - \left(f(t_i;\tau_j) + \frac{\partial f}{\partial \tau}\bigg|_{\tau_j} \cdot \Delta\tau_j \right)$$

New interpretation "How much e_i is changed, approximately, if $\tau_j \to \tau_j + \Delta\tau_j$"

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} y_1 - f(t_1;\tau_j) \\ y_2 - f(t_2;\tau_j) \\ \vdots \end{bmatrix} - \begin{bmatrix} \frac{\partial f}{\partial \tau}\Big|_{(t_1,\tau_j)} \\ \frac{\partial f}{\partial \tau}\Big|_{(t_2,\tau_j)} \\ \vdots \end{bmatrix} \Delta\tau_j$$

$$E_j = D_j - Z_j (\Delta A)_j$$

"difference" vector

a bit diff than in linear version, but still each row is from independent variable (and now parameters), and N data points $\Rightarrow$ N rows

not the best fit, but how to adjust A to make it better

Some linear algebra math, similar to before

$$\underbrace{E_{j+1}^{T} E_{j+1}}_{S_{r,j+1}} = \left(D_j - Z_j \, \Delta A_j\right)^{T}\left(D_j - Z_j \, \Delta A_j\right)$$

via $\frac{dS_{r,j+1}}{d(\Delta A)} = 0 \longrightarrow \boxed{\Delta A_j = \left(Z_j^{T} Z_j\right)^{-1} \cdot \left(Z_j^{T} D_j\right)}$

This is the iterative scheme we use to adjust the parameter values until we are happy with the convergence:

$$\left| \frac{\tau_{j+1} - \tau_j}{\tau_{j+1}} \right| < \text{error}$$

How is it different for multi-parameter models?

Residual: $e_i = y_i - \left(f(t_i; \tau_j, H_j, \ldots) + \frac{\partial f}{\partial \tau} \Delta \tau_j + \frac{\partial f}{\partial H} \Delta H_j + \ldots \right)$

$\downarrow$

$$\underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \end{bmatrix}}_{E_j} = \underbrace{\begin{bmatrix} y_1 - f(t_1; \tau_j, H_j, \ldots) \\ y_2 - f(t_2; \tau_j, H_j, \ldots) \\ \vdots \end{bmatrix}}_{D_j} - \underbrace{\begin{bmatrix} \left.\frac{\partial f}{\partial \tau}\right|_{t_1} & \left.\frac{\partial f}{\partial H}\right|_{t_1} & \cdots \\ \left.\frac{\partial f}{\partial \tau}\right|_{t_2} & \left.\frac{\partial f}{\partial H}\right|_{t_2} & \cdots \\ \vdots & \vdots & \end{bmatrix}}_{Z_j} \underbrace{\begin{bmatrix} \Delta \tau \\ \Delta H \\ \vdots \end{bmatrix}}_{\Delta A_j}$$

$\longrightarrow$ Same strategy: $\boxed{\Delta A_j = (Z_j^T Z_j)^{-1} (Z_j^T D_j)}$

update $\tau_{j+1} = \tau_j + \Delta \tau_j$

$H_{j+1} = H_j + \Delta H_j$

$\vdots$

until parameters don't change much