Nonlinear Least Squares

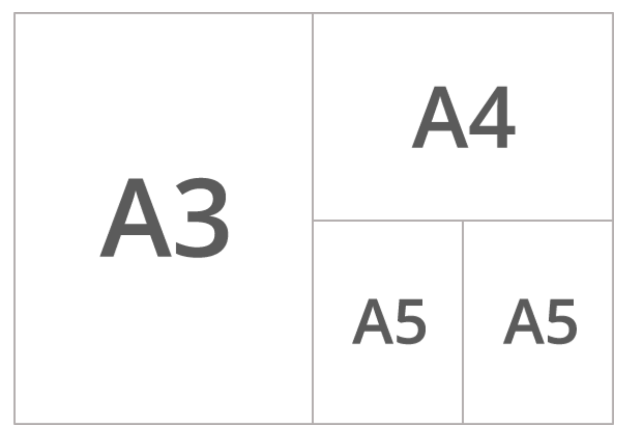
Mei-Chen Yeh

Announcements

- Final exam 2022 is available on moodle.

Announcements

- We will have the final on Dec. 22, starting at 14:20.
  - Programming problems
  - Written problems
- A cheat sheet (A5 two sides or A4 one side) is allowed.



Nonlinear Least Squares

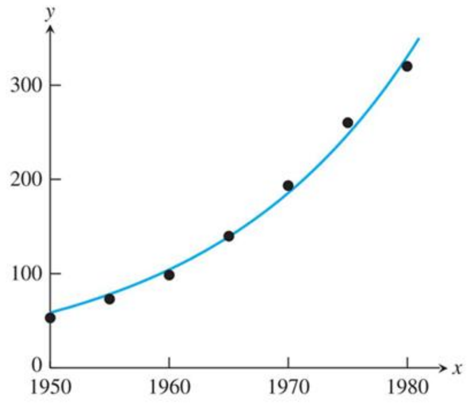
Least squares

- The least squares solution $\underline{\tilde{x}}$ of a linear system $A\underline{x} = \underline{b}$ minimizes the Euclidean norm of the residual $\left\|A\underline{\tilde{x}} - \underline{b}\right\|_2$.
- Two methods for finding $\underline{\tilde{x}}$
  – Normal equations
  – QR factorization
- But, we have cases in which neither method can be applied...

If the equations are *nonlinear*

- Example 1: exponential model

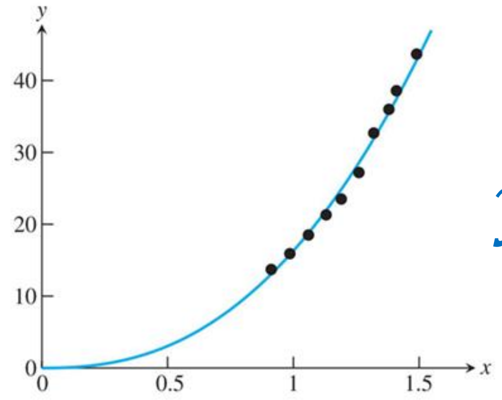
$$y = c_1 e^{c_2 x}$$



$$y = 54.03e^{0.06152x}$$

- Example 2: power law model

$$y = c_1 x^{c_2}$$



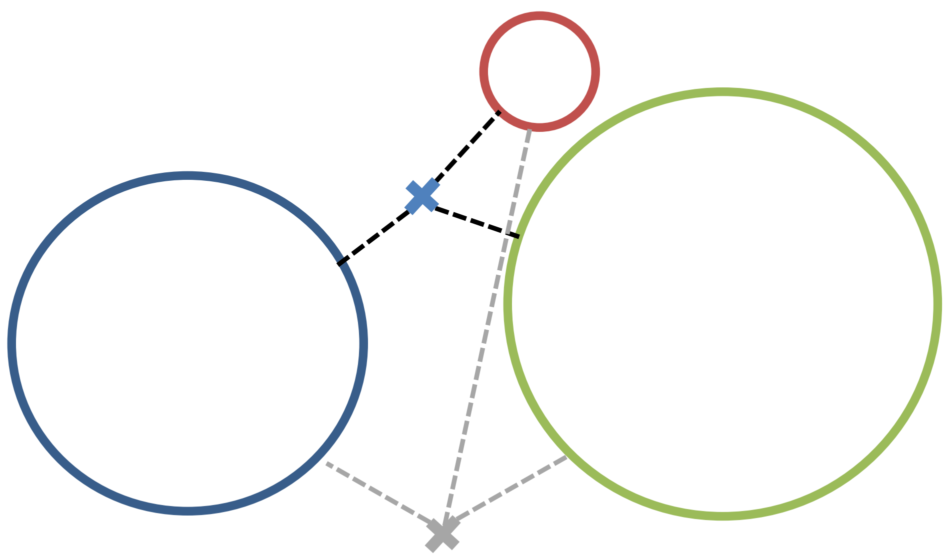
$$y = 16.3x^{2.42}$$

Example

$$2x + y = 1$$
$$x^2 + y^2 = 1$$

Example

- Given three circles (centers x_i, y_i, and radii R_i), find the point for which *the sum of the squared distances to the three circles is minimized*.

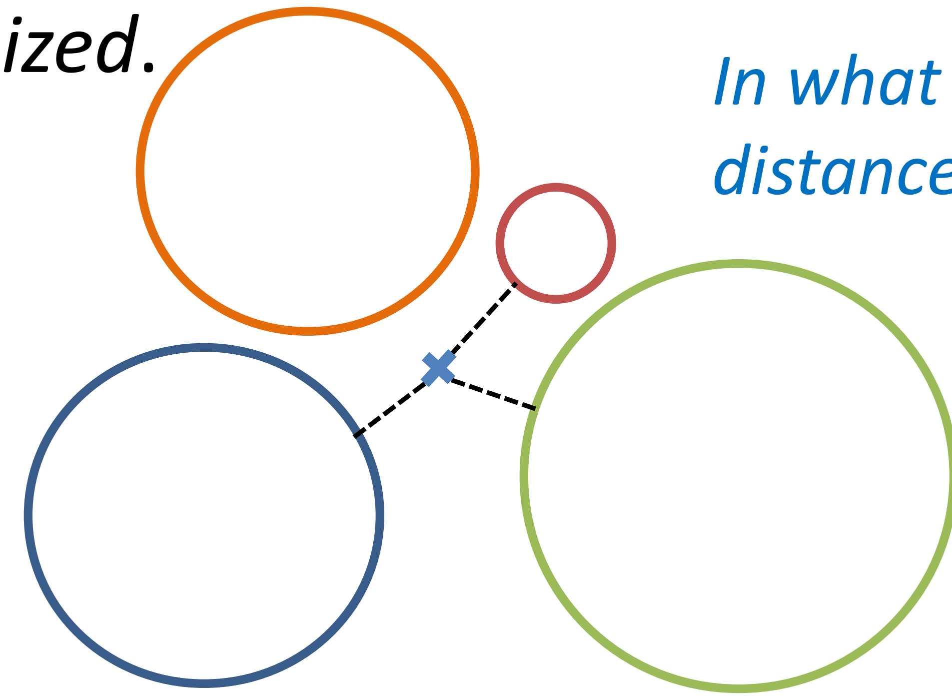


Example

four

- Given ~~three~~ circles (centers x_i, y_i, and radii R_i), find the point for which *the sum of the squared distances to the three circles is minimized*.

In what situation the distance is zero?



Today

- *Gauss-Newton method* for solving nonlinear least squares problems

Multivariate Newton's method + Normal equations

HW#7 HW#10

Recall the method for solving nonlinear systems

- Multivariate Newton's method (HW#7)
  - An iterative method

$$\underline{x}_0 = \text{initial vector}$$
$$\underline{x}_{k+1} = \underline{x}_k - \left(D_F\left(\underline{x}_k\right)\right)^{-1} F(\underline{x}_k) \text{ for } k = 0,1,2,\ldots$$

$$F\left(\underline{x}_k\right) = ?$$

$$D_F\left(\underline{x}_k\right) = ?$$

Review: $F(\underline{x})_{n \times 1}$ and $D_F(\underline{x})_{n \times n}$

- Suppose we have 3 unknowns, 3 nonlinear equations ($m = n$):

$$f_1 = (u, v, w) = 0$$
$$f_2 = (u, v, w) = 0$$
$$f_3 = (u, v, w) = 0$$

- Define the vector-valued function:

$$F(\underline{x}) = F(u, v, w) = (f_1, f_2, f_3)$$

where

$$\underline{x} = (u, v, w).$$

Review: $F(\underline{x})_{n \times 1}$ and $D_F(\underline{x})_{n \times n}$

- 3 variables: *u*, *v*, *w*
- 3 equations: f_1, f_2, f_3

$$D_F(\underline{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} & \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} & \frac{\partial f_2}{\partial w} \\ \frac{\partial f_3}{\partial u} & \frac{\partial f_3}{\partial v} & \frac{\partial f_3}{\partial w} \end{bmatrix}$$

Jacobian matrix

Review: Multivariate Newton's method

$\underline{x}_0$ = initial vector

$\underline{x}_{k+1} = \underline{x}_k - \left(D_F(\underline{x}_k)\right)^{-1} F(\underline{x}_k)$ for $k = 0,1,2,\ldots$

$$\underline{s} = -\left(D_F(\underline{x}_k)\right)^{-1} F(\underline{x}_k)$$

$$D_F(\underline{x}_k)\underline{s} = -F(\underline{x}_k)$$

$A \qquad \underline{x} \qquad\qquad \underline{b}$

Solving $A\underline{x} = \underline{b}$

$\underline{s}$ is the solution of $D_F(\underline{x}_k)\underline{s} = -F(\underline{x}_k)$

Review: Multivariate Newton's method

$\underline{x}_0$ = initial vector

$\underline{x}_{k+1} = \underline{x}_k - \left(D_F(\underline{x}_k)\right)^{-1} F(\underline{x}_k)$ for $k = 0,1,2,\ldots$

$\Downarrow$

$\underline{x}_0$ = initial vector

solve $D_F(\underline{x}_k)\underline{s} = -F(\underline{x}_k)$

$\underline{x}_{k+1} = \underline{x}_k + \underline{s}$ for $k = 0,1,2,\ldots$

Solving nonlinear least squares

- Consider the system of *m* (nonlinear) equations in *n* unknowns: (*m* > *n*)

$$r_1(x_1, \cdots, x_n) = 0$$
$$\vdots$$
$$r_m(x_1, \cdots, x_n) = 0$$

- Sum of the squares of the errors

$$r_1(\underline{x})^2 + r_2(\underline{x})^2 + \cdots + r_m(\underline{x})^2$$

$$\uparrow$$

$$\boxed{\textit{Find a solution } \underline{x} \textit{ that minimizes the sum!}}$$

Solving nonlinear least squares

- Consider the system of *m* (nonlinear) equations in *n* unknowns: ($m > n$)

$$\begin{gathered} r_1(x_1, \cdots, x_n) = 0 \\ \vdots \\ r_m(x_1, \cdots, x_n) = 0 \end{gathered}$$

- D_r

$$\begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \cdots & \frac{\partial r_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial r_m}{\partial x_1} & \cdots & \frac{\partial r_m}{\partial x_n} \end{bmatrix} m \times n$$

Analogy

- $F(\underline{x})$ $n \times 1$
- $D_F(\underline{x})$ $n \times n$

- $r(\underline{x})$ $m \times 1$
- $D_r(\underline{x})$ $m \times n$

Example: the exponential model

- $y = c_1 e^{c_2 t}$
  - Unknowns? $\underline{x} = [c_1, c_2]$
  - Equations? 7

year	cars ($\times 10^6$)
1950	53.05
1955	73.04
1960	98.31
1965	139.78
1970	193.48
1975	260.20
1980	320.39
t	y

$$r_1: \quad c_1 e^{1950 c_2} - 53.05 \times 10^6 = 0$$
$$r_2: \quad c_1 e^{1955 c_2} - 73.04 \times 10^6 = 0$$
$$\vdots$$
$$r_m: \quad c_1 e^{1980 c_2} - 320.39 \times 10^6 = 0$$

Find c_1 and c_2 that minimizes $r_1^2 + \cdots + r_m^2$

$$y = c_1 e^{c_2 t}$$

- $\underline{r}(\underline{x}_k) \quad \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$$\begin{aligned}
r_1&: \quad c_1 e^{1950 c_2} - 53.05 \times 10^6 \\
r_2&: \quad c_1 e^{1955 c_2} - 73.04 \times 10^6 \\
&\quad\vdots \\
r_m&: \quad c_1 e^{1980 c_2} - 320.39 \times 10^6
\end{aligned}$$

$$\begin{bmatrix} c_1 e^{c_2 t} - y_1 \\ \vdots \\ c_1 e^{c_2 t} - y_m \end{bmatrix} \quad m \times 1$$

$$y = c_1 e^{c_2 t}$$

- $D_r(\underline{x}_k) \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

  $r_i = c_1 e^{c_2 t_i} - y_i,\ i = 1..m$

  $\frac{\partial r_i}{\partial c_1} \Rightarrow e^{c_2 t_i}$

  $\frac{\partial r_i}{\partial c_2} \Rightarrow c_1 t_i e^{c_2 t_i}$

$$\begin{bmatrix} e^{c_2 t_1} & c_1 t_1 e^{c_2 t_1} \\ \vdots & \vdots \\ e^{c_2 t_m} & c_1 t_m e^{c_2 t_m} \end{bmatrix}_{m \times n}$$

$$\boxed{D_r = \begin{bmatrix} \frac{\partial r_1}{\partial c_1} & \cdots & \frac{\partial r_1}{\partial c_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial r_m}{\partial c_1} & \cdots & \frac{\partial r_m}{\partial c_n} \end{bmatrix}_{m \times n}}$$

Multivariate Newton

$$\underline{x}_0 = \text{initial vector}$$

$$\underline{x}_{k+1} = \underline{x}_k - \left(D_F(\underline{x}_k)\right)^{-1} F(\underline{x}_k) \text{ for } k = 0,1,2,\ldots$$

$$\left.\begin{array}{l} F(\underline{x}_k) \to r(\underline{x}_k) \\ D_F(\underline{x}_k) \to D_r(\underline{x}_k) \end{array}\right. \to \textit{normal equations}$$

$$A^T A \underline{\tilde{x}} = A^T b$$

Gauss-Newton

$$\underline{x}_0 = \text{initial vector}$$

$$\underline{x}_{k+1} = \underline{x}_k - \left(D_r(\underline{x}_k)^T D_r(\underline{x}_k)\right)^{-1} D_r(\underline{x}_k)^T r(\underline{x}_k) \text{ for } k = 0,1,2,\ldots$$

$$\underline{v}_k = -\left(D_r(\underline{x}_k)^T D_r(\underline{x}_k)\right)^{-1} D_r(\underline{x}_k)^T r(\underline{x}_k)$$

$$\underbrace{\left(\underset{n\times m}{D_r(\underline{x}_k)^T}\ \underset{m\times n}{D_r(\underline{x}_k)}\right)}\ \underbrace{\underset{n\times 1}{\underline{v}_k}} = \underbrace{-\underset{n\times m}{D_r(\underline{x}_k)^T}\ \underset{m\times 1}{r(\underline{x}_k)}}$$

Gauss-Newton

$\underline{x}_0$ = initial vector

$$\underline{x}_{k+1} = \underline{x}_k \underbrace{- \left(D_r(\underline{x}_k)^T D_r(\underline{x}_k)\right)^{-1} D_r(\underline{x}_k)^T r(\underline{x}_k)}_{\underline{v}_k} \text{ for } k = 0,1,2,\ldots$$

$$\underline{x}_{k+1} = \underline{x}_k + \underline{v}_k$$

Multivariate Newton

$\underline{x}_0$ = initial vector

solve $D_F(\underline{x}_k)\underline{s} = -F(\underline{x}_k)$

$\underline{x}_{k+1} = \underline{x}_k + \underline{s}$ for $k = 0,1,2,\ldots$

Gauss-Newton

$\underline{x}_0$ = initial vector

solve $\left(D_r(\underline{x}_k)^T D_r(\underline{x}_k)\right)\underline{v} = -D_r(\underline{x}_k)^T r(\underline{x}_k)$

$\underline{x}_{k+1} = \underline{x}_k + \underline{v}$ for $k = 0,1,2,\ldots$

Gauss-Newton method

To minimize

$$r_1(\underline{x})^2 + r_2(\underline{x})^2 + \cdots + r_m(\underline{x})^2$$

Set $\underline{x}^0$ = initial vector,

for k = 0, 1, 2, …

$$A = D_r(\underline{x}^k)$$

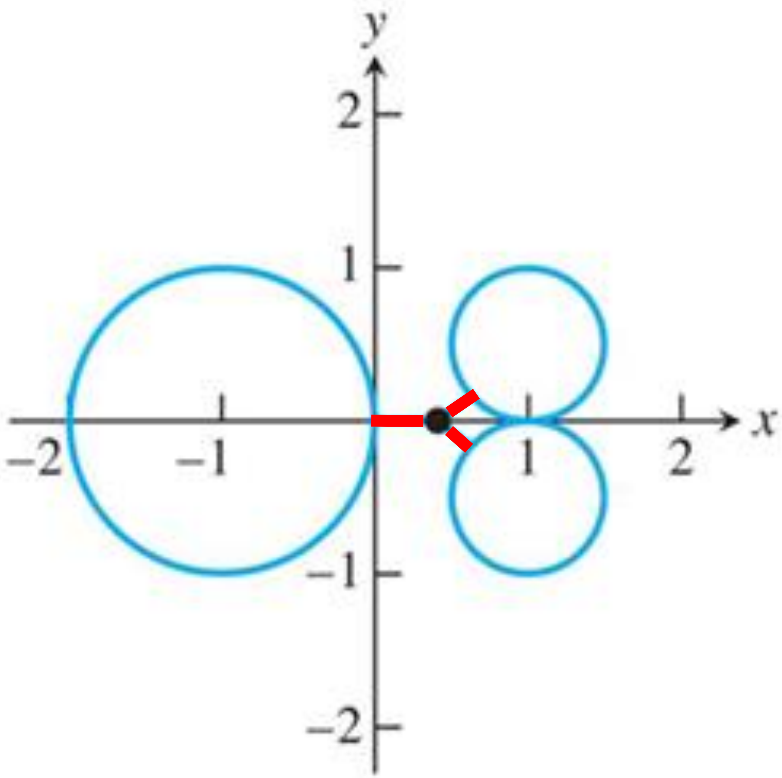
$$A^T A \underline{v}^k = -A^T r(\underline{x}^k)$$

$$\underline{x}^{k+1} = \underline{x}^k + \underline{v}^k$$

end

Example

- Given centers (−1,0), (1, 0.5), (1, −0.5) and radii 1, 0.5, 0.5, find the point for which the sum of the squared distances to the three circles is minimized.



Unknowns? (*x, y*)

Equations? 3 distances between the point to the circles

$$r_1(x, y) = \sqrt{(x - x_1)^2 + (y - y_1)^2} - R_1$$

$$r_2(x, y) = \sqrt{(x - x_2)^2 + (y - y_2)^2} - R_2$$

$$r_3(x, y) = \sqrt{(x - x_3)^2 + (y - y_3)^2} - R_3.$$

$\underline{x}_0$ = initial vector $[x_0, y_0]^T$

$\underline{x}_k = [x, y]^T$

solve $\left(D_r(\underline{x}_k)^T D_r(\underline{x}_k)\right)\underline{v} = -D_r(\underline{x}_k)^T r(\underline{x}_k)$

$\underline{x}_{k+1} = \underline{x}_k + \underline{v}$ for $k = 0, 1, 2, \ldots$

$r_1(x,y)$

$r_2(x,y)$

$r_3(x,y)$

$r(x,y) = [r_1\ r_2\ r_3]^T$

$$D_r(x,y) = \begin{bmatrix} \partial r_1/_x & \partial r_1/_y \\ \partial r_2/_x & \partial r_2/_y \\ \partial r_3/_x & \partial r_3/_y \end{bmatrix} = \begin{bmatrix} \frac{x-x_1}{S_1} & \frac{y-y_1}{S_1} \\ \frac{x-x_2}{S_2} & \frac{y-y_2}{S_2} \\ \frac{x-x_3}{S_3} & \frac{y-y_3}{S_3} \end{bmatrix}$$

$$r_1(x,y) = \sqrt{(x-x_1)^2 + (y-y_1)^2} - R_1$$

$$\frac{\partial r_1}{x} = \frac{1}{2}\left((x-x_1)^2 + (y-y_1)^2\right)^{-\frac{1}{2}}(2x - 2x_1)$$

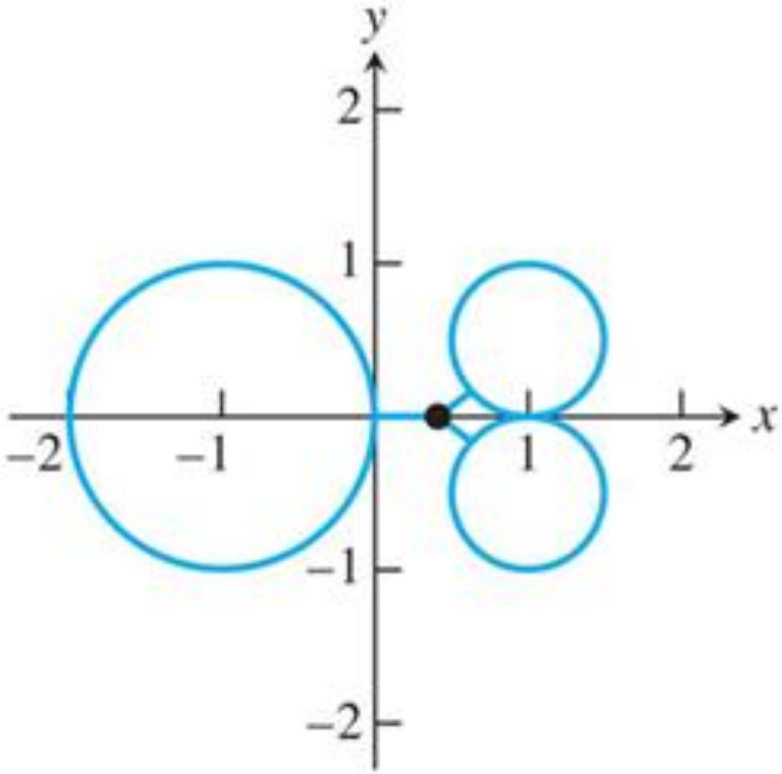
$$= \frac{x-x_1}{S_1} \qquad S_1 = \sqrt{(x-x_1)^2 + (y-y_1)^2}$$

$\underline{x}_0$ = initial vector

solve $\left(D_r(\underline{x}_k)^T D_r(\underline{x}_k)\right)\underline{v} = -D_r(\underline{x}_k)^T r(\underline{x}_k)$

$\underline{x}_{k+1} = \underline{x}_k + \underline{v}$ for $k = 0, 1, 2, \ldots$

- Using Gauss-Newton iteration with initial vector (0, 0), get $(x, y) = (0.4129, 0)$.



程式練習

And, please upload your program on moodle.

- Find the point (x, y) for which the sum of the squared distances from the point to the three circles is minimized.
- Circle 1: (0, 1) R = 1
- Circle 2: (1, 1) R = 1
- Circle 3: (0, -1) R = 1