Linear Regression

Write a program which illustrates simple linear regression (or a more general variant of linear regression) and implements accumulation of canonical information.

a) For some fixed parameters a and b (or, in a more general case, $a_1, \ldots, a_m$) generate a sequence of "observations" (x_i, y_i):

$$y_i = f(x_i) + \varepsilon_i,$$

where

$$f(x) = a + bx \quad \text{or} \quad f(x) = a_1 + a_2 x + a_3 x^2 + \cdots + a_m x^{m-1}$$

ε_i are i.i.d. with zero mean and $\mathsf{E}\varepsilon_i^2 = \sigma^2$. Values x_i can be generated randomly with some mean and variance.

b) Accumulate canonical information, i.e., at each step, when a new observation (x_i, y_i) is produced, update canonical information.

c) Illustrate the real function $f(x)$ and its estimate $\widehat{f(x)}$.

d) Illustrate $\mathrm{Var}(\widehat{f(x)})$, assuming that σ^2 is known.

e) Illustrate $\widehat{\mathrm{Var}(\widehat{f(x)})}$, assuming that σ^2 is NOT known.

In your report present the source code and a few (around 3) nice graphs showing estimations for "small", "intermediate", and "large" number of observations.

Example & formulas

$$y_i = f_a(x_i) + \varepsilon_i = a_1 f_1(x_i) + \cdots + a_m f_m(x_i) + \varepsilon_i$$

or

$$y_i = F_{x_i} a + \varepsilon_i,$$

where

$$F_x = \left[\begin{array}{cccc} f_1(x) & f_2(x) & \cdots & f_m(x) \end{array}\right].$$

Function used in the demo - polynomial:

$$y_i = 1 + 1 \cdot x_i - 1 \cdot x_i^2 + 0.2 \cdot x_i^3 + \varepsilon_i$$

,

$$F_x = \left[\begin{array}{cccc} 1 & x & x^2 & x^3 \end{array}\right], \qquad m = 4, \qquad a = \left[\begin{array}{c} 1 \\ 1 \\ -1 \\ 0.2 \end{array}\right]$$

Data:	(x_i, y_i),	$i = 1, \ldots, n$
Canonical information:	(T, v, V, n)	

Elementary information: (T_i, v_i, V_i, n_i)

$$n_i = 1, \qquad V_i = y_i^2, \qquad v_i = F_{x_i}^T \cdot y_i = \begin{bmatrix} f_1(x_i)\, y_i \\ \vdots \\ f_4(x_i)\, y_i \end{bmatrix},$$

$$T_i = F_{x_i}^T \cdot F_{x_i} = \begin{bmatrix} f_1(x_i)^2 & f_1(x_i)\, f_2(x_i) & \cdots & f_1(x_i)\, f_4(x_i) \\ \vdots & \vdots & \ddots & \vdots \\ f_4(x_i)\, f_1(x_i) & f_4(x_i)\, f_2(x_i) & \cdots & f_4(x_i)^2 \end{bmatrix}$$

Update:

$$(T, v, V, n) + (T_i, v_i, V_i, n_i) = (T + T_i, v + v_i, V + V_i, n + n_i)$$

Estimate $f(x)$:

$$(T, v, V, n) * x \mapsto$$
$$\widehat{f(x)} = F_x T^{-1} v,$$
$$\mathrm{Var}(\widehat{f(x)}) = \sigma^2 F_x T^{-1} F_x^T,$$
$$\widehat{\mathrm{Var}(\widehat{f(x)})} = \frac{V - v^T T^{-1} v}{n - m} \cdot F_x T^{-1} F_x^T.$$

Part of the code in MatLab

```
in = in + Info([x,y]);    % Update: Elem. Info & Combine
est = in * xv;            % Apply Info
```

Example of an illustration