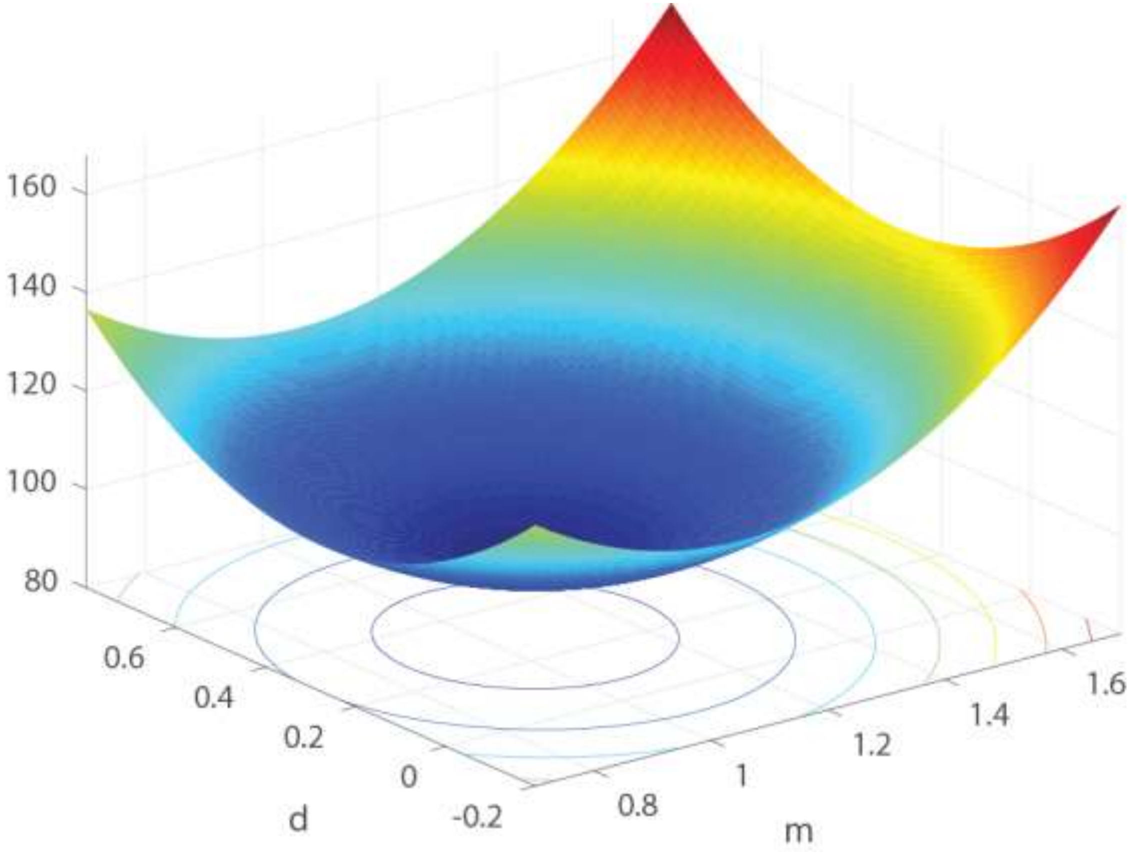
Line Fitting (Ax=b)



Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

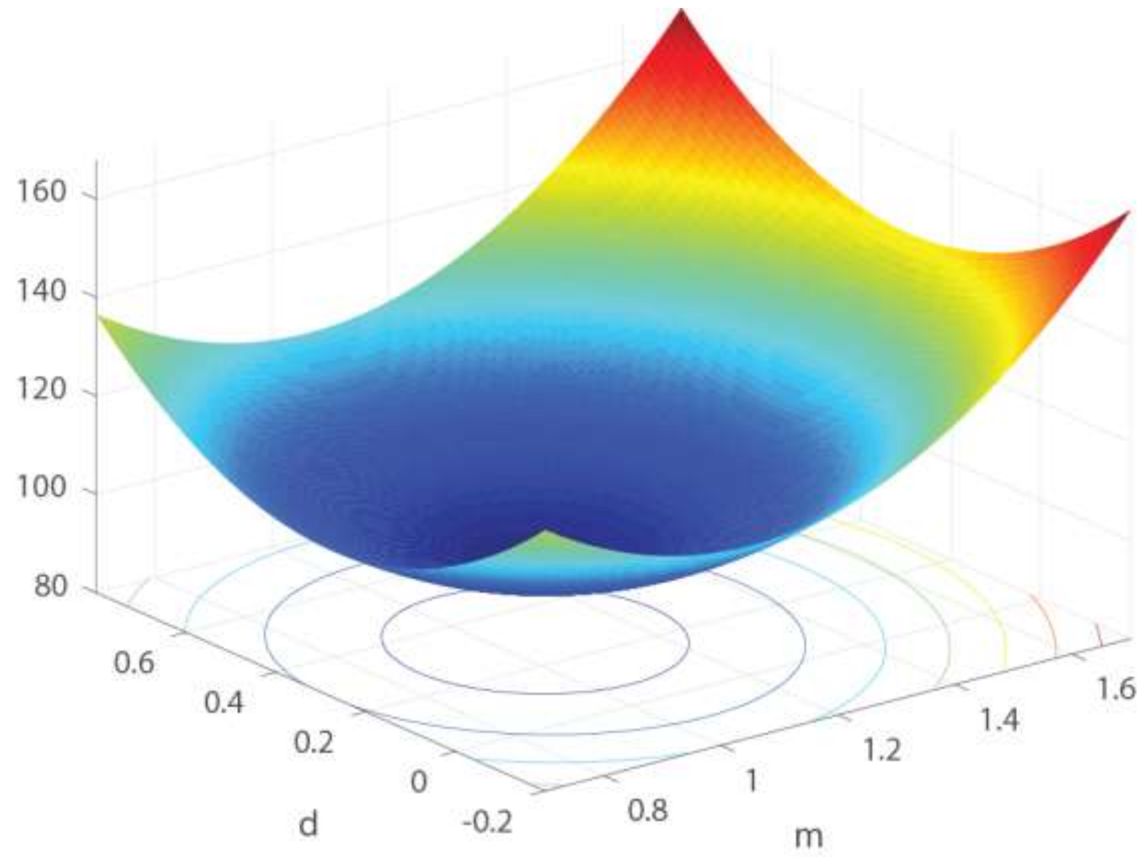
Optimal point: $\frac{\partial E}{\partial m} = 0, \quad \frac{\partial E}{\partial d} = 0$

$$v_1 \approx mu_1 + d$$

$$v_2 \approx mu_2 + d$$

$$v_n \approx mu_n + d$$

Line Fitting (Ax=b)



Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

Optimal point: $\frac{\partial E}{\partial m} = 0, \quad \frac{\partial E}{\partial d} = 0$

$$v_1 \approx mu_1 + d$$
$$v_2 \approx mu_2 + d$$
$$v_n \approx mu_n + d$$

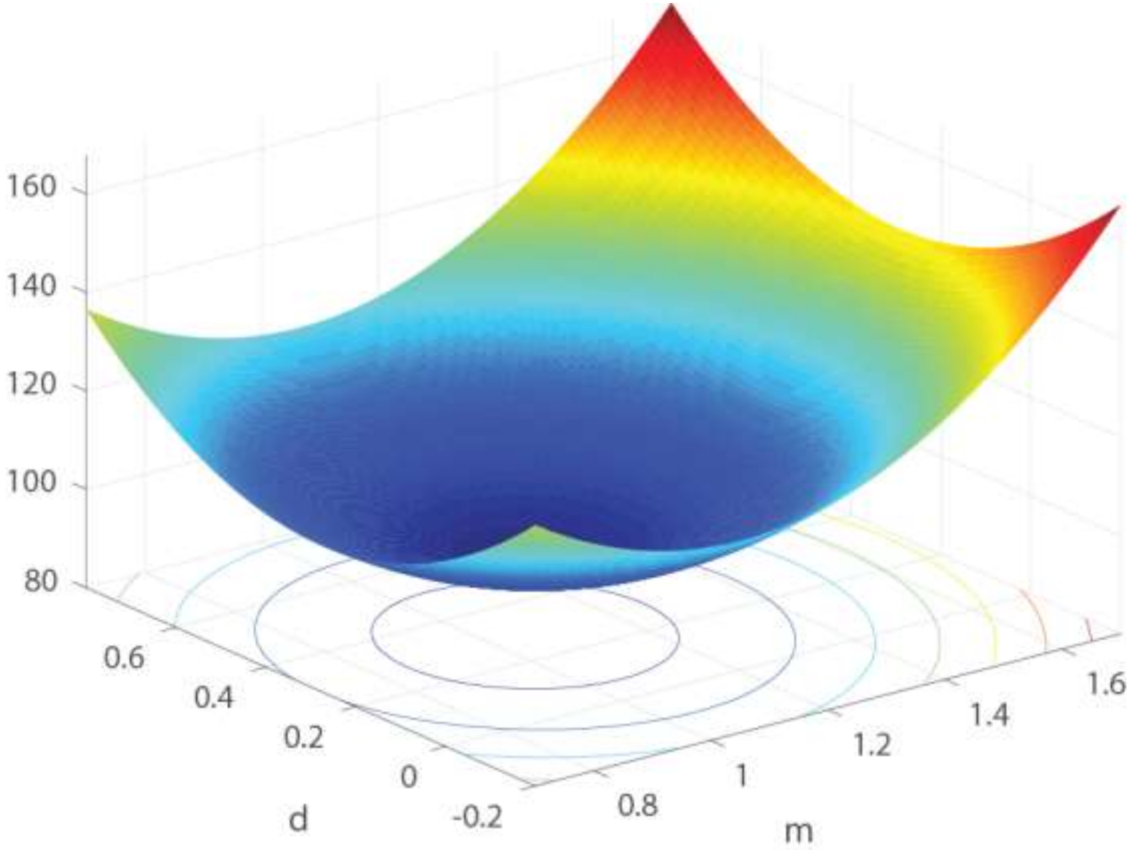
$\rightarrow$

$$\begin{bmatrix} u_1 & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_1$$
$$\begin{bmatrix} u_2 & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_2$$
$$\begin{bmatrix} u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_n$$

Line Fitting (Ax=b)

Total error: $E = \sum_{i=1}^{n} (v_i - (mu_i + d))^2$

Optimal point: $\frac{\partial E}{\partial m} = 0, \quad \frac{\partial E}{\partial d} = 0$



$$v_1 \approx mu_1 + d$$
$$v_2 \approx mu_2 + d$$
$$v_n \approx mu_n + d$$

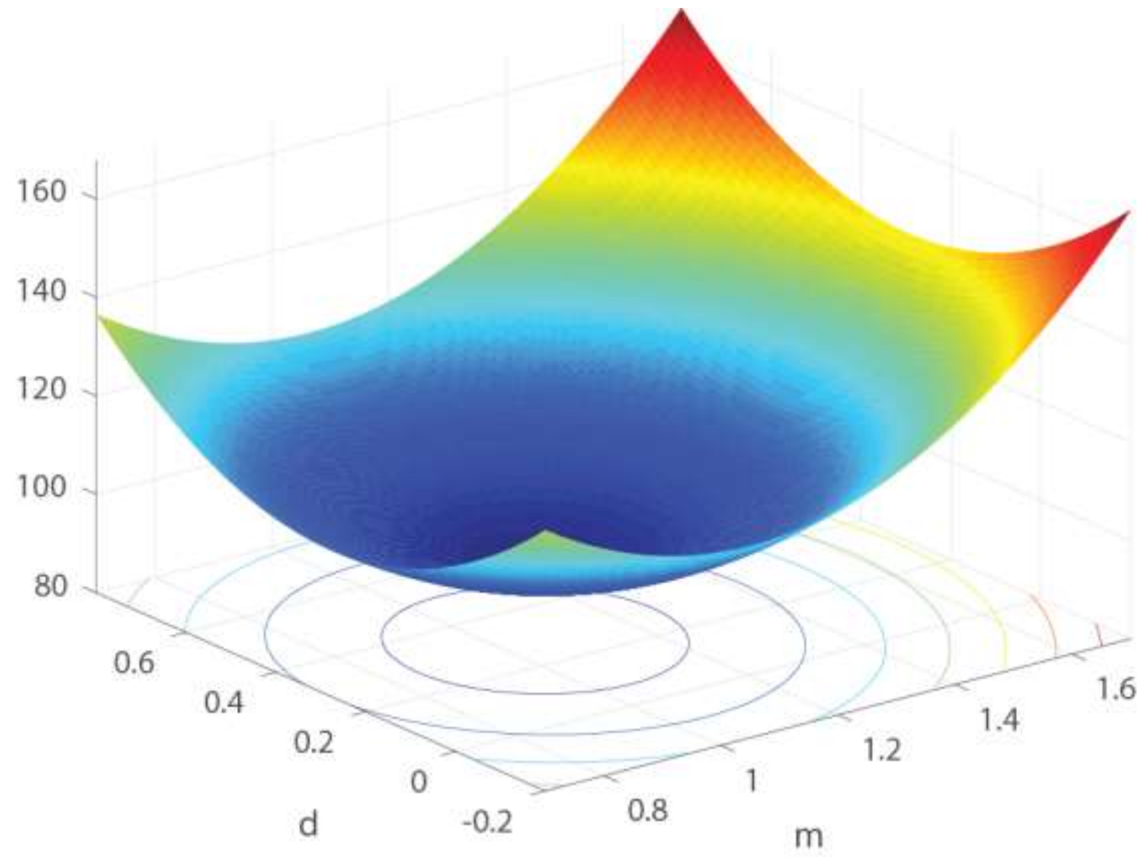
$\rightarrow$

$$\begin{bmatrix} u_1 & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_1$$
$$\begin{bmatrix} u_2 & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_2$$
$$\begin{bmatrix} u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_n$$

$\rightarrow$

$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Line Fitting (Ax=b)



Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

Optimal point: $\frac{\partial E}{\partial m} = 0, \quad \frac{\partial E}{\partial d} = 0$

$v_1 \approx mu_1 + d$

$v_2 \approx mu_2 + d$

$v_n \approx mu_n + d$

→

$\begin{bmatrix} u_1 & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_1$

$\begin{bmatrix} u_2 & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_2$

$\begin{bmatrix} u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx v_n$

→

$\underset{A}{\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix}} \underset{x}{\begin{bmatrix} m \\ d \end{bmatrix}} \approx \underset{b}{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}}$

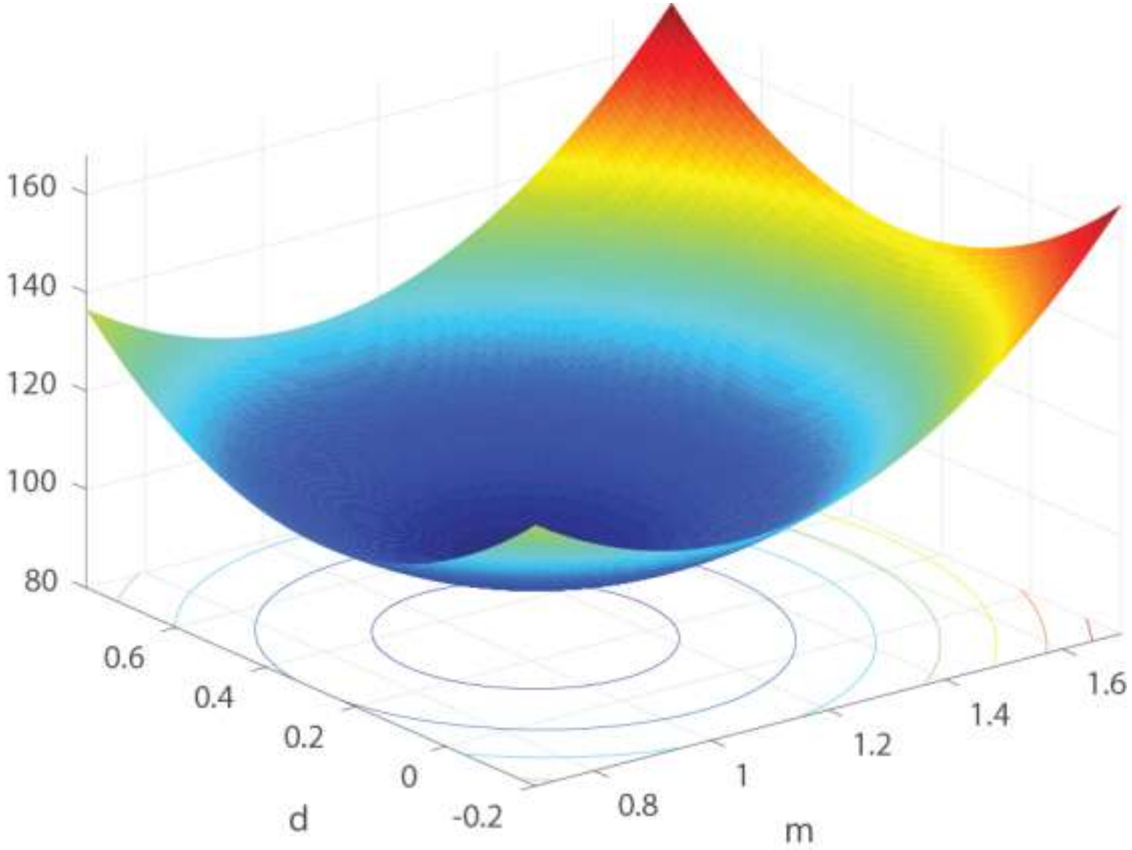
We can't invert A.

Line Fitting (Ax=b)

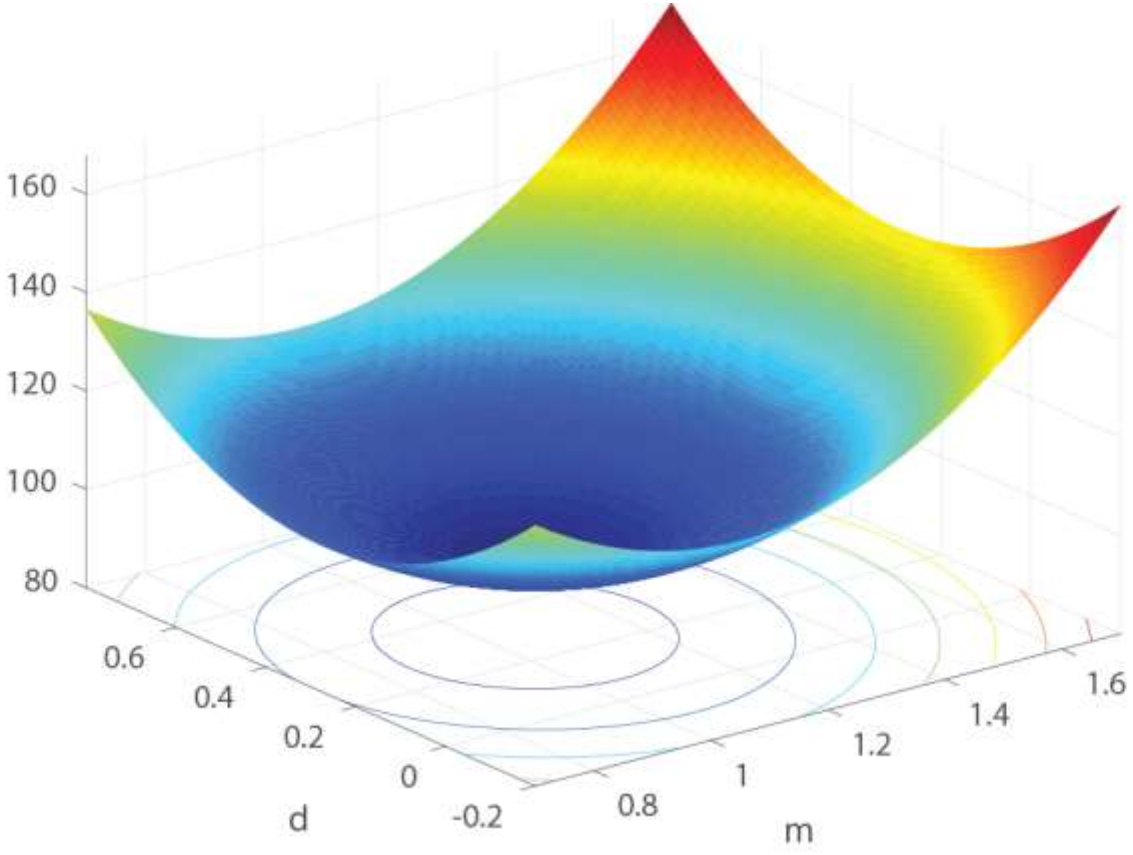
Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

$$E = \left(\mathrm{Ax} - \mathrm{b}\right)^{T} \left(\mathrm{Ax} - \mathrm{b}\right) = \left\|\mathrm{Ax} - \mathrm{b}\right\|^2$$

$$\underbrace{\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix}}_{\mathrm{A}} \underbrace{\begin{bmatrix} m \\ d \end{bmatrix}}_{\mathrm{x}} \approx \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}}_{\mathrm{b}}$$



Line Fitting (Ax=b)



Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\mathrm{A} \quad \mathrm{x} \approx \mathrm{b}$$

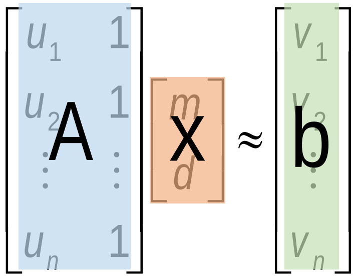
$$E = (\mathrm{Ax} - \mathrm{b})^T (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^T \mathrm{A}^T \mathrm{Ax} - 2\mathrm{x}^T \mathrm{A}^T \mathrm{b} + \mathrm{b}^T \mathrm{b}$$

$$\frac{\partial E}{\partial \mathrm{x}} = \qquad ?$$

Line Fitting (Ax=b)

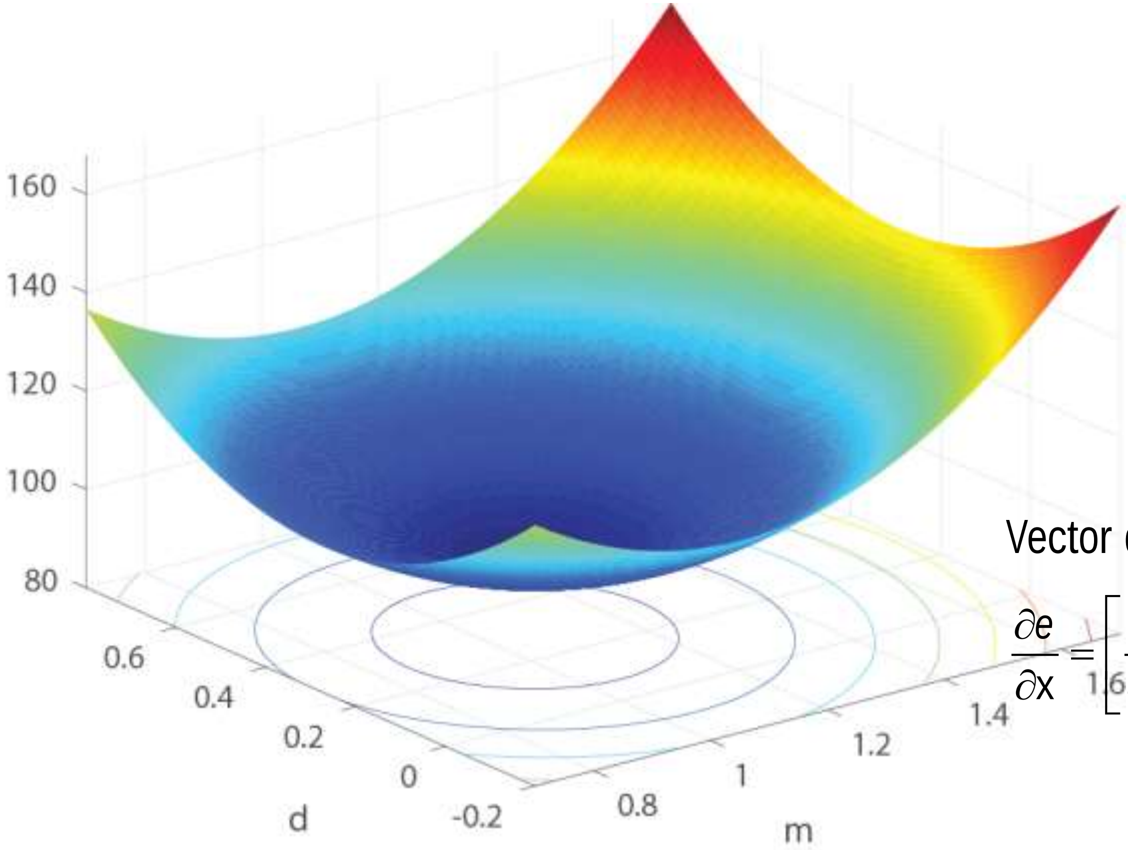
Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$



$$E = (\mathrm{Ax} - \mathrm{b})^T (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^T \mathrm{A}^T \mathrm{Ax} - 2\mathrm{x}^T \mathrm{A}^T \mathrm{b} + \mathrm{b}^T \mathrm{b}$$

$$\frac{\partial E}{\partial \mathrm{x}} = \qquad ?$$



Vector derivative:

$$\frac{\partial e}{\partial \mathrm{x}} = \left[\frac{\partial e}{\partial x_1} \quad \cdots \quad \frac{\partial e}{\partial x_n} \right]$$

Line Fitting (Ax=b)

Total error: $E = \sum_{i=1}^{n} (v_i - (mu_i + d))^2$

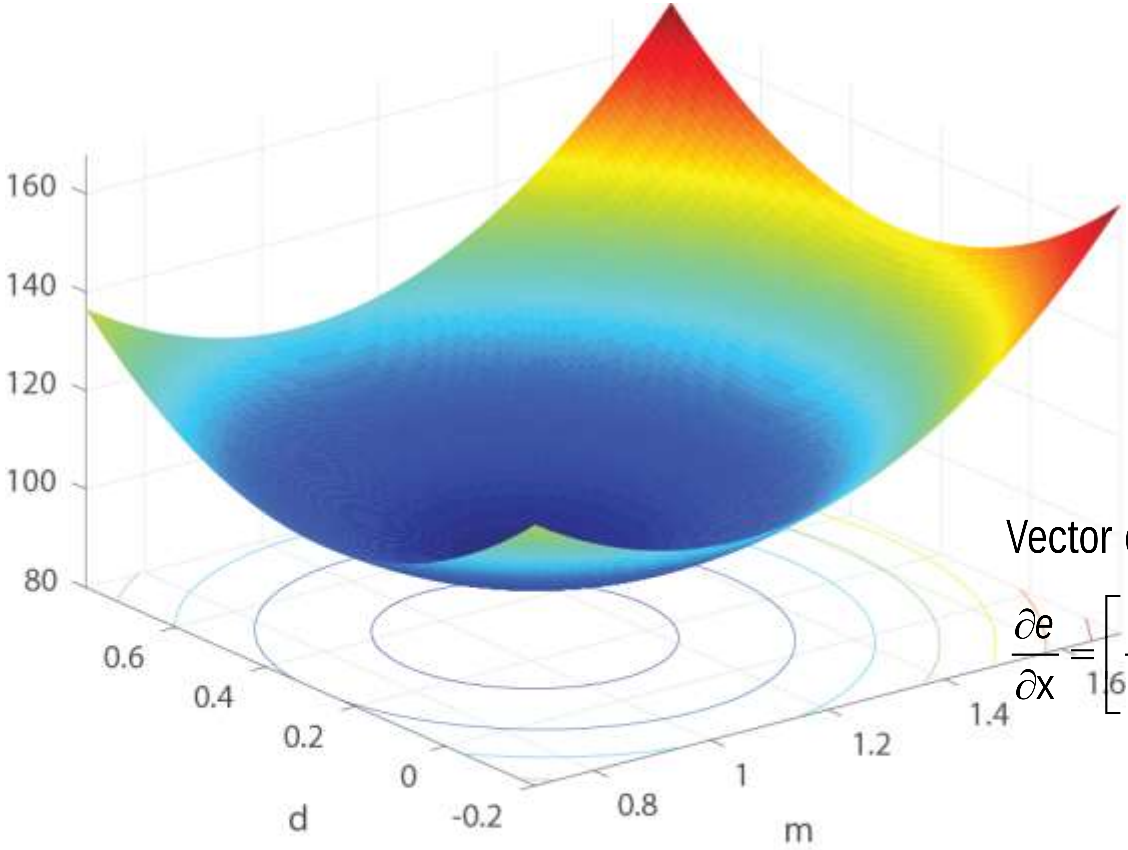
$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\mathrm{A}\,\mathrm{x} \approx \mathrm{b}$$

$$E = (\mathrm{Ax} - \mathrm{b})^T (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^T \mathrm{A}^T \mathrm{Ax} - 2\mathrm{x}^T \mathrm{A}^T \mathrm{b} + \mathrm{b}^T \mathrm{b}$$

$$\frac{\partial E}{\partial \mathrm{x}} = \quad ?$$



Vector derivative:

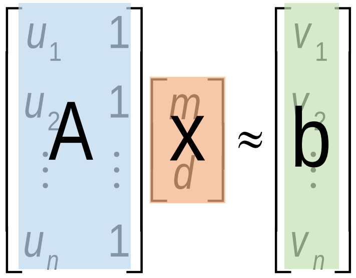
$$\frac{\partial e}{\partial \mathrm{x}} = \begin{bmatrix} \frac{\partial e}{\partial x_1} & \cdots & \frac{\partial e}{\partial x_n} \end{bmatrix}$$

Ex) $e = \mathrm{c}^T \mathrm{x} = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

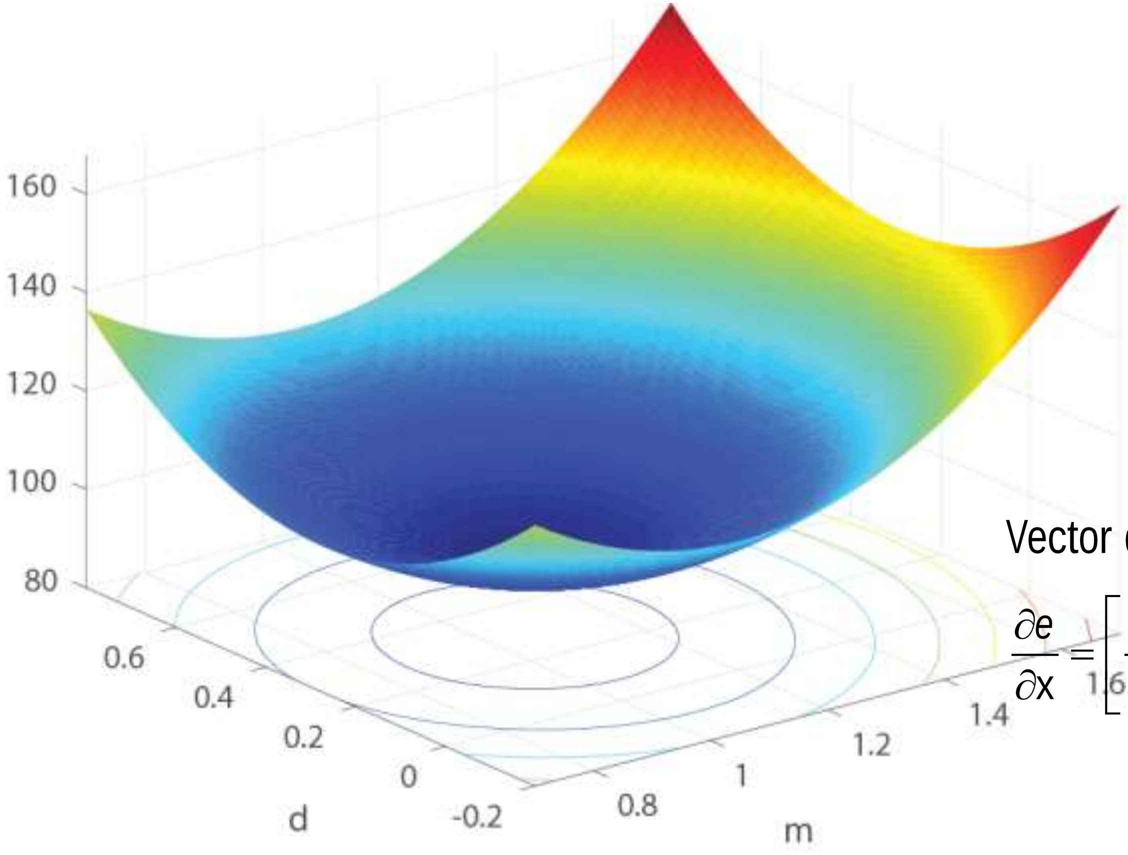
$$\frac{\partial \mathrm{c}^T \mathrm{x}}{\partial \mathrm{x}} =$$

Line Fitting (Ax=b)

Total error: $E = \sum_{i=1}^{n} (v_i - (mu_i + d))^2$



$$E = (\mathrm{Ax} - \mathrm{b})^T (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$
$$= \mathrm{x}^T \mathrm{A}^T \mathrm{Ax} - 2\mathrm{x}^T \mathrm{A}^T \mathrm{b} + \mathrm{b}^T \mathrm{b}$$



$$\frac{\partial E}{\partial \mathrm{x}} = \quad ?$$

Vector derivative:

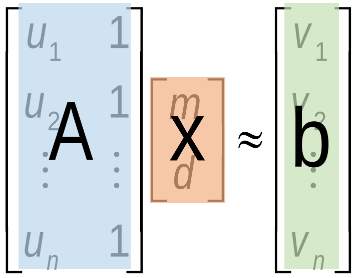
$$\frac{\partial e}{\partial \mathrm{x}} = \begin{bmatrix} \frac{\partial e}{\partial x_1} & \cdots & \frac{\partial e}{\partial x_n} \end{bmatrix}$$

Ex) $e = \mathrm{c}^T \mathrm{x} = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$$\frac{\partial \mathrm{c}^T \mathrm{x}}{\partial \mathrm{x}} = \frac{\partial}{\partial \mathrm{x}} (c_1 x_1 + \cdots + c_n x_n)$$

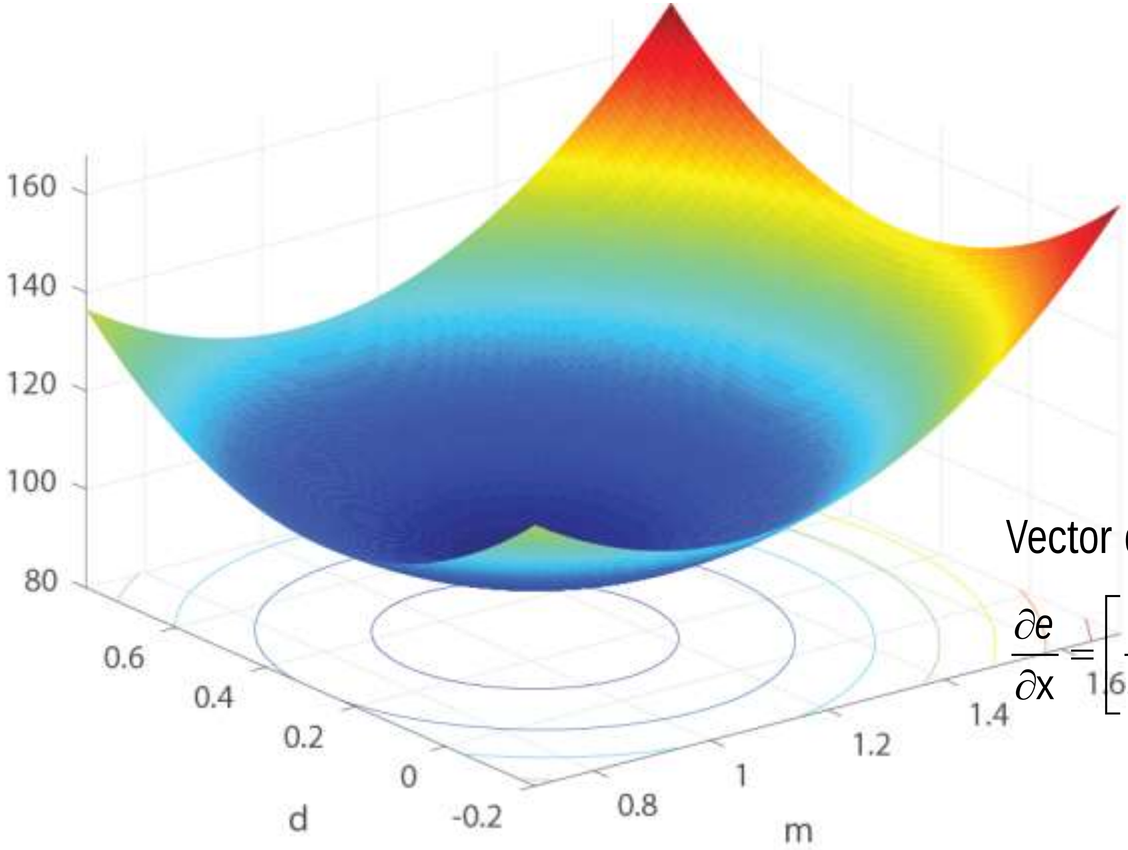
Line Fitting (Ax=b)

Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$



$$E = (\mathrm{Ax} - \mathrm{b})^{\mathrm{T}} (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$
$$= \mathrm{x}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{Ax} - 2\mathrm{x}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{b} + \mathrm{b}^{\mathrm{T}}\mathrm{b}$$

$$\frac{\partial E}{\partial \mathrm{x}} = \quad ?$$



Vector derivative:

$$\frac{\partial e}{\partial \mathrm{x}} = \begin{bmatrix} \frac{\partial e}{\partial x_1} & \cdots & \frac{\partial e}{\partial x_n} \end{bmatrix}$$

Ex) $e = \mathrm{c}^{\mathrm{T}}\mathrm{x} = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$$\frac{\partial \mathrm{c}^{\mathrm{T}}\mathrm{x}}{\partial \mathrm{x}} = \frac{\partial}{\partial \mathrm{x}}(c_1x_1 + \cdots + c_nx_n)$$
$$= \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix}$$

Line Fitting (Ax=b)

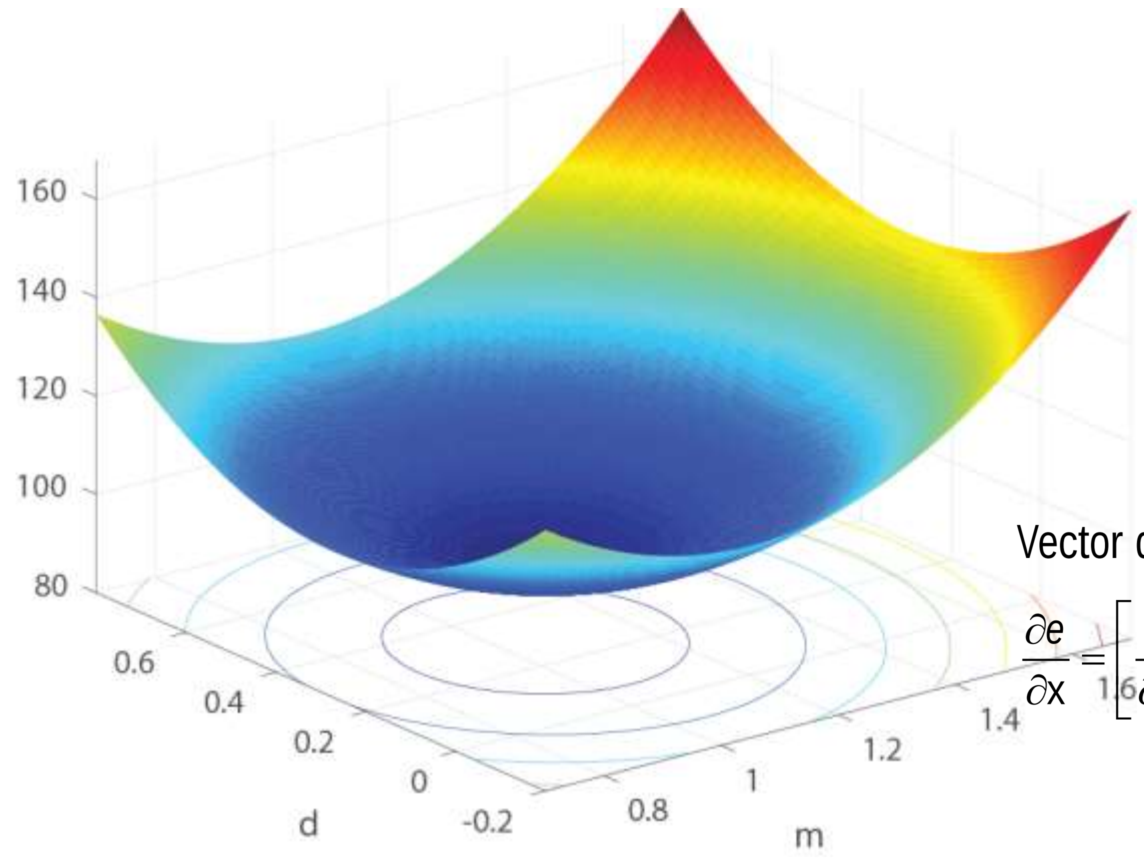
Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \mathrm{A}\,\mathrm{x} \approx \mathrm{b}$$

$$E = (\mathrm{Ax} - \mathrm{b})^{\mathrm{T}}(\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{Ax} - 2\mathrm{x}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{b} + \mathrm{b}^{\mathrm{T}}\mathrm{b}$$

$$\frac{\partial E}{\partial \mathrm{x}} = 2\mathrm{A}^{\mathrm{T}}\mathrm{Ax} - 2\mathrm{A}^{\mathrm{T}}\mathrm{b} = 0$$



Vector derivative:

$$\frac{\partial e}{\partial \mathrm{x}} = \begin{bmatrix} \frac{\partial e}{\partial x_1} & \cdots & \frac{\partial e}{\partial x_n} \end{bmatrix}$$

Ex) $e = \mathrm{c}^{\mathrm{T}}\mathrm{x} = \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$$\frac{\partial \mathrm{c}^{\mathrm{T}}\mathrm{x}}{\partial \mathrm{x}} = \frac{\partial}{\partial \mathrm{x}}(c_1x_1 + \cdots + c_nx_n)$$

$$= \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix}$$

Line Fitting (Ax=b)

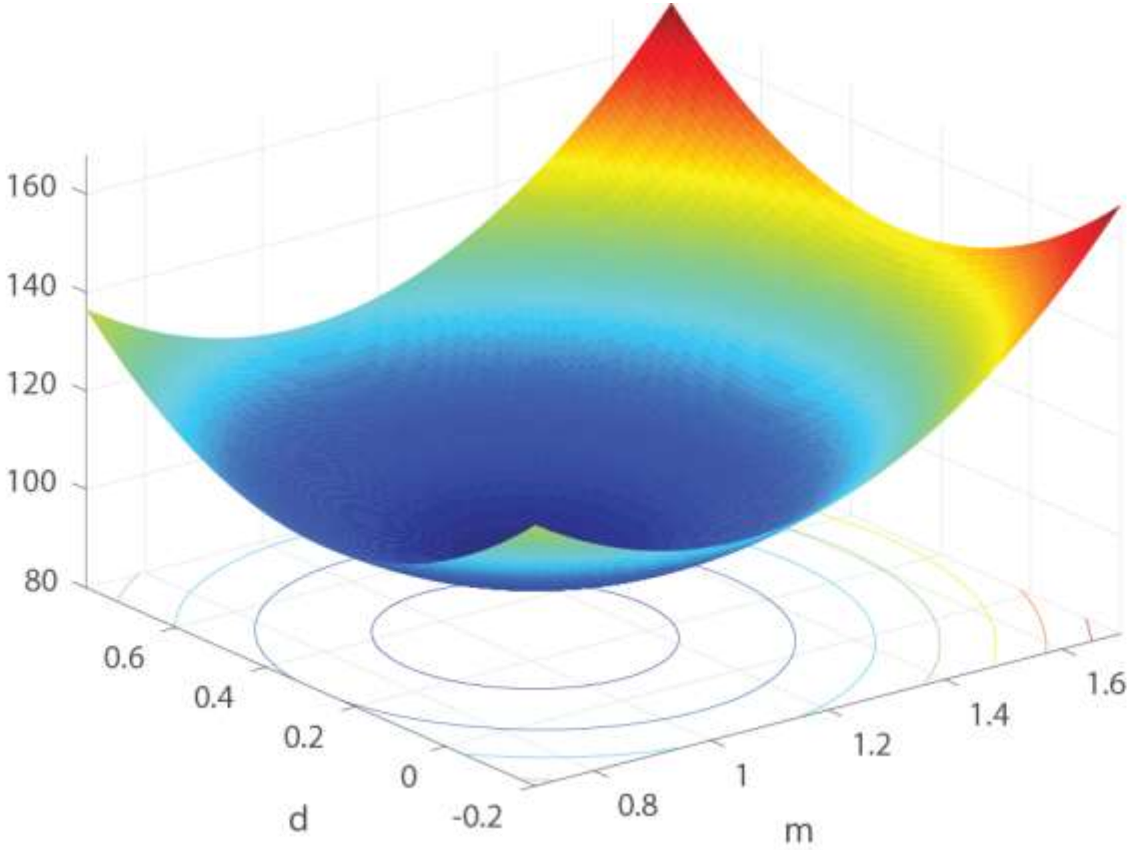
Total error: $E = \sum_{i=1}^{n} (v_i - (mu_i + d))^2$

$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\mathrm{A} \; \mathrm{x} \approx \mathrm{b}$$

$$E = (\mathrm{Ax} - \mathrm{b})^T (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^T \mathrm{A}^T \mathrm{Ax} - 2\mathrm{x}^T \mathrm{A}^T \mathrm{b} + \mathrm{b}^T \mathrm{b}$$



$$\frac{\partial E}{\partial \mathrm{x}} = 2\mathrm{A}^T \mathrm{Ax} - 2\mathrm{A}^T \mathrm{b} = 0$$

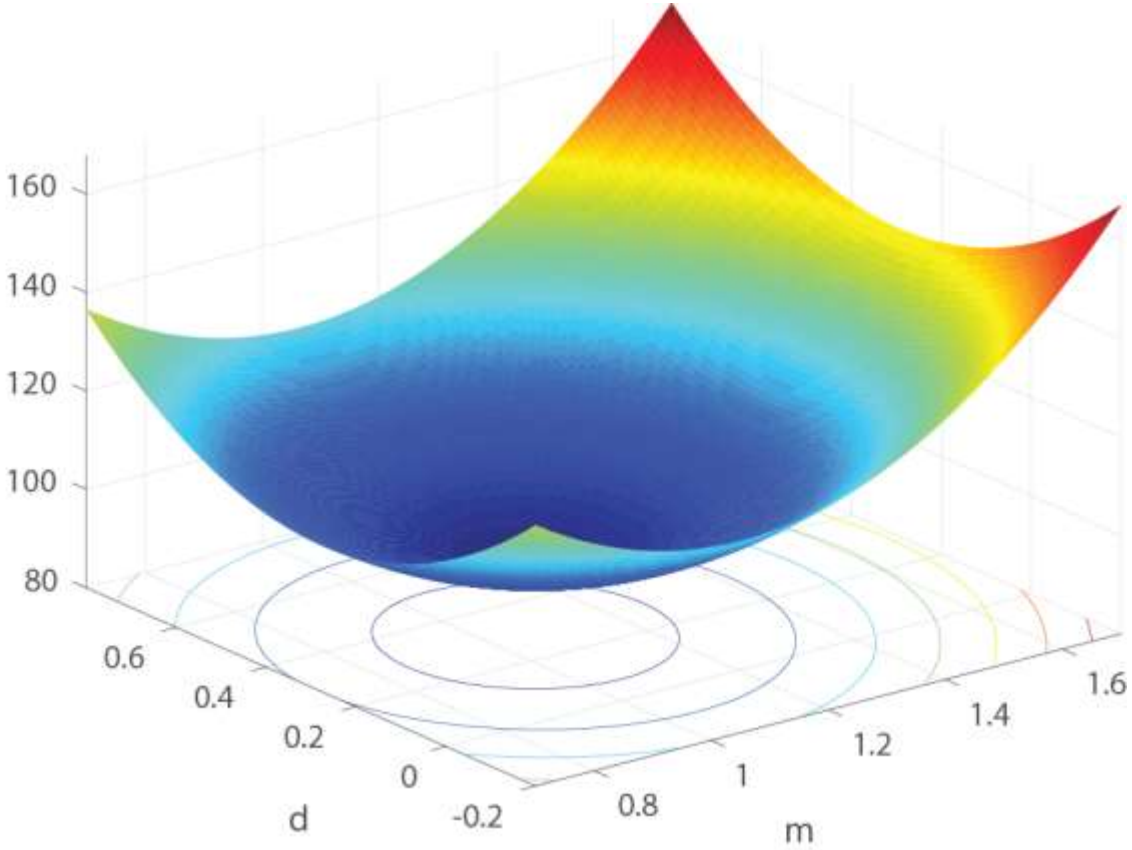
$$\longrightarrow \mathrm{A}^T \mathrm{Ax} = \mathrm{A}^T \mathrm{b}$$

Line Fitting (Ax=b)

Total error: $E = \sum_{i=1}^{n} \left(v_i - (mu_i + d)\right)^2$

$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

A x ≈ b



$$E = (\mathrm{Ax} - \mathrm{b})^{\mathrm{T}} (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{Ax} - 2\mathrm{x}^{\mathrm{T}}\mathrm{A}^{\mathrm{T}}\mathrm{b} + \mathrm{b}^{\mathrm{T}}\mathrm{b}$$

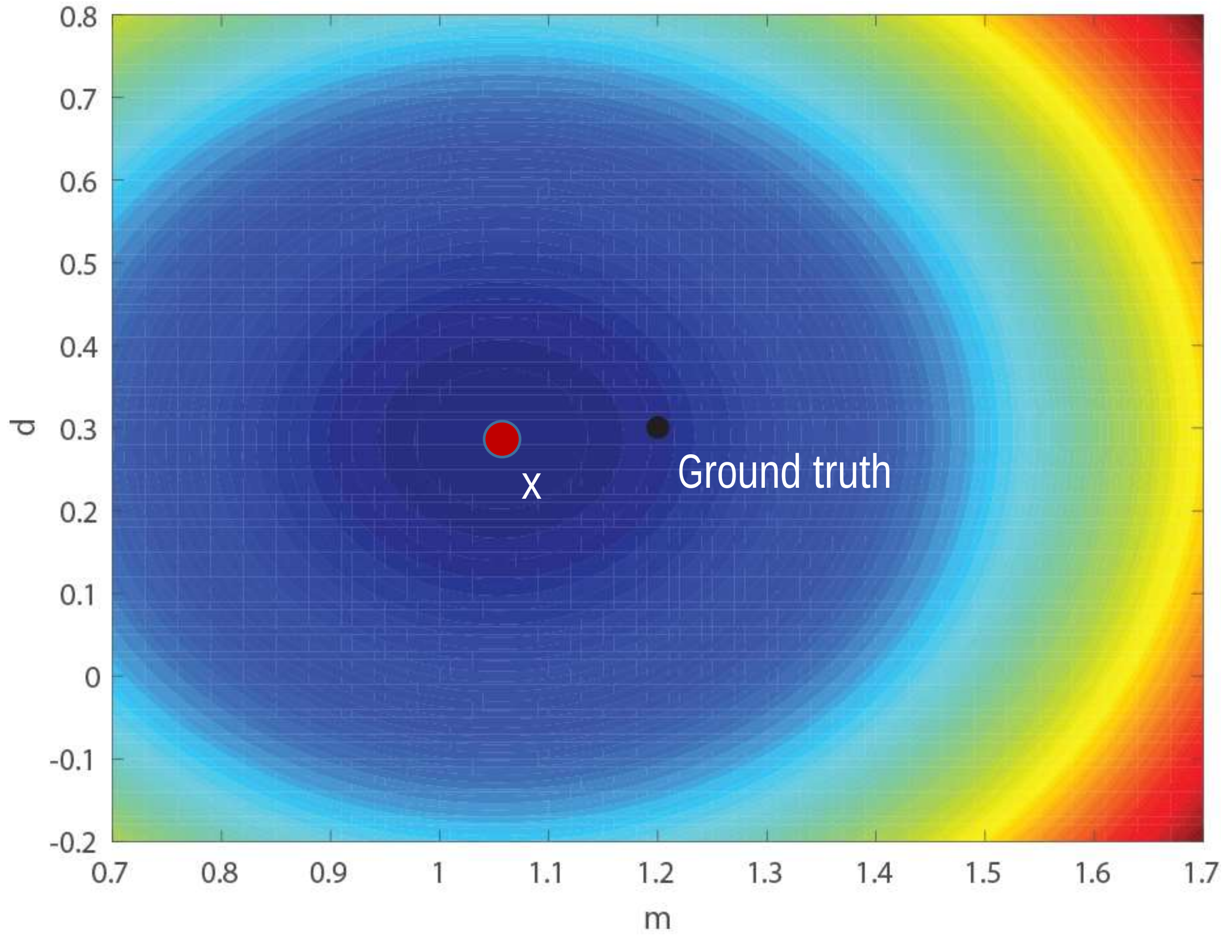
$$\frac{\partial E}{\partial \mathrm{x}} = 2\mathrm{A}^{\mathrm{T}}\mathrm{Ax} - 2\mathrm{A}^{\mathrm{T}}\mathrm{b} = 0$$

$$\longrightarrow \mathrm{A}^{\mathrm{T}}\mathrm{Ax} = \mathrm{A}^{\mathrm{T}}\mathrm{b}$$

$$\mathrm{A}^{\mathrm{T}} \; \mathrm{A} \; \mathrm{x} = \mathrm{A}^{\mathrm{T}} \; \mathrm{b}$$

Normal equation

Line Fitting (Ax=b)



Total error: $E = \sum_{i=1}^{n} (v_i - (mu_i + d))^2$

$$\begin{bmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_n & 1 \end{bmatrix} \begin{bmatrix} m \\ d \end{bmatrix} \approx \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

A x ≈ b

$$E = (\mathrm{Ax} - \mathrm{b})^T (\mathrm{Ax} - \mathrm{b}) = \|\mathrm{Ax} - \mathrm{b}\|^2$$

$$= \mathrm{x}^T \mathrm{A}^T \mathrm{Ax} - 2\mathrm{x}^T \mathrm{A}^T \mathrm{b} + \mathrm{b}^T \mathrm{b}$$

$$\frac{\partial E}{\partial \mathrm{x}} = 2\mathrm{A}^T \mathrm{Ax} - 2\mathrm{A}^T \mathrm{b} = 0$$

$$\longrightarrow \mathrm{A}^T \mathrm{Ax} = \mathrm{A}^T \mathrm{b}$$

$$A^T \; A \; x = A^T \; b$$

$$x = \left[A^T \; A \right]^{-1} A^T \; b$$