Data Assimilation - Bayes' Theorem

miércoles, 12 de agosto de 2020 01:44 p. m.

$$\underbrace{P(x|y)}_{\text{POSTERIOR}} \propto \overbrace{P(x)}^{\text{PRIOR}} \cdot \overbrace{\underbrace{\mathcal{L}(x|y)}_{P(y|x)}}^{\text{LIKELIHOOD}}$$

VAR

OBS

$P(x|y)$, $P(x)$, $\mathcal{L}(x|y)$ } PDF

$$P(x|y) = \frac{P(x)\cdot\mathcal{L}(x|y)}{P(y)}$$

$$= \underbrace{\frac{1}{P(y)}}_{k} \cdot P(x)\cdot\mathcal{L}(x|y) \propto \underbrace{P(x)}_{\text{PRIOR BELIEF}} \cdot \underbrace{\mathcal{L}(x|y)}_{\text{PRIOR}}$$

PRIOR OBSERVATION

DATA

DATA

$$P(x|y) \propto \underbrace{P(x)}_{\text{PRIOR BELIEF}} \cdot \overbrace{\mathcal{L}(x|y)}^{\text{OBSERVATION}}$$

IMPERFECT NUMERICAL FORECAST

OBSERVATIONS: LAND STATION, SATELITE

$\mathbb{R}$

$x^b \in \mathbb{R}$ — MODEL SIMULATION — PRIOR

$y \in \mathbb{R}$ — TERMOMETRO — OBS

1. x^b IMPERFECT
2. y NOISY

$x \sim x^t$

I HAVE NO IDEA

$\dfrac{1}{\sqrt{2\pi}\sigma_b}$

ASSUMPTIONS

1. $x \sim N(x^b, \sigma_b^2)$ $\quad P(x) \propto e^{-\frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2}}$ $\quad \varepsilon^b$

2. $y \sim N(x, \sigma_o^2)$ $\quad \mathcal{L}(x|y) \propto e^{-\frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$ $\quad \varepsilon^o$

?

$$\varepsilon^b = x - x^b \sim N(0, \sigma_b^2) \qquad \varepsilon^o = y - x \sim N(0, \sigma_o^2)$$

3. $\mathrm{cov}(\varepsilon^b, \varepsilon^o) = 0$ NO CORRELATED ERRORS

$P(x|y) \propto P(x) \cdot \mathcal{L}(x|y)$:

$P(x) \propto e^{-\frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2}}$ — MEAN, VAR

$\mathcal{L}(x|y) \propto e^{-\frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$ — MEAN, VAR

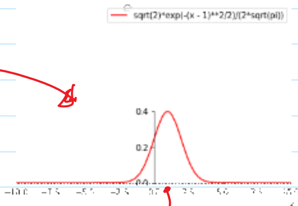
$\rightarrow N(?, ?)$

$$P(x|y) = ?$$

$$P(x|y) \propto P(x) \cdot \mathcal{L}(x|y) \propto e^{-\frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2}} \cdot e^{-\frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$$

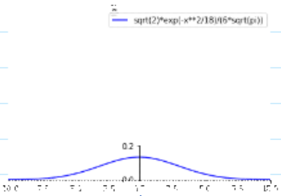
sqrt(2)*exp(-(x - 1)**2/2)/(2*sqrt(pi))

$$P(x|y) \propto P(x) \cdot \mathcal{L}(x|y) \propto e^{-\frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2}} \cdot e^{-\frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$$



y, σ_o

$$= e^{-\frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2} - \frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$$

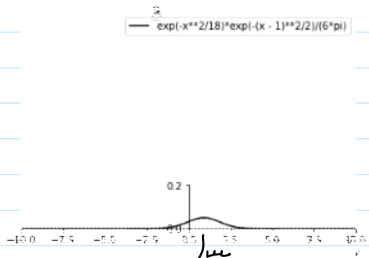


x^b, σ_b

$\mathcal{N}(x^b, \sigma_b^2)$

$$= e^{-J(x)} \quad \boxed{J(x) = \frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2} + \frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$$

$\mathcal{N}(x, \sigma_o^2)$



x^a, σ_a

$\mathcal{N}(x^a, \sigma_a^2)$

MAP — MAXIMUM A POSTERIORI ESTIMATE

$$x^a = \underset{x}{\arg\max}\ P(x|y)$$

(+)

$$P(x|y) \propto e^{-J(x)}$$

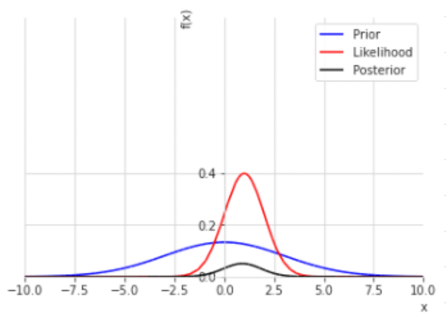
$x^a = ?$ } ANALYSIS

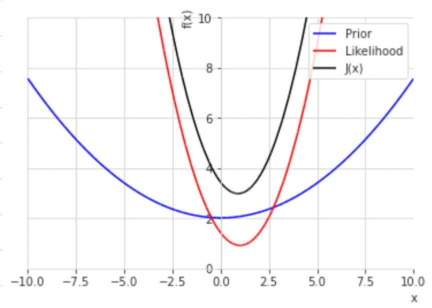
$\sigma_a = ?$

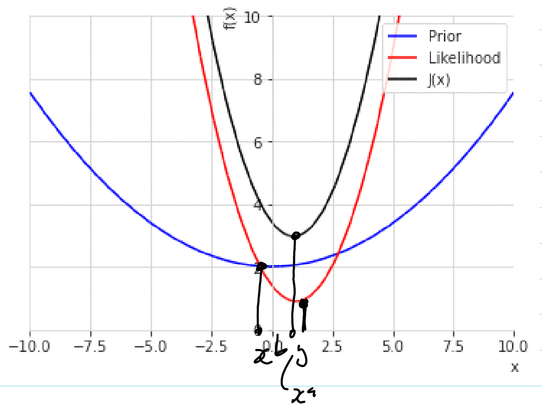
$$P(x|y) \propto e^{-\frac{1}{2}J(x)} \qquad J(x) = \frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2} + \frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}$$

=

$$x^a = \underset{x}{\arg\max}\ P(x|y) \qquad x^a = \underset{x}{\arg\min}\ J(x)$$







① $J(x) = \frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2} + \frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}$ $\qquad J'(x) = \frac{x-x^b}{\sigma_b^2} - \frac{y-x}{\sigma_o^2}$

$$J'(x) = \frac{x}{\sigma_b^2} - \frac{x^b}{\sigma_b^2} - \frac{y}{\sigma_o^2} + \frac{x}{\sigma_o^2} = \left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right)x - \left[\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right]$$

$$J'(x^a) = 0 \Rightarrow \quad x^a = \underbrace{\left(\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right)^{-1}}\left(\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right)$$

$$\left[\frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}\right]^{-1} = \frac{\sigma_o^2 \cdot \sigma_b^2}{\sigma_b^2 + \sigma_o^2} \qquad x^a = \left[\frac{\sigma_o^2 \cdot \sigma_b^2}{\sigma_b^2 + \sigma_o^2}\right]\left[\frac{x^b}{\sigma_b^2} + \frac{y}{\sigma_o^2}\right]$$

$$\Rightarrow \quad x^a = \underbrace{\frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2}}_{\alpha_1} \cdot x^b + \underbrace{\frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}}_{\alpha_2} \cdot y$$

$\boxed{\alpha_1 + \alpha_2 = 1}$

$\alpha_1 \in [0,1]$

$\alpha_2 \in [0,1]$

CONVEX COMBINATION

$$x^a = \alpha\, x^b + (1-\alpha)\, y$$

WHAT MODEL SAYS

WHAT OBSERVATION SAYS

$$\boxed{\alpha = \frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2}}$$

② $P(x|y) \propto e^{-\frac{1}{2}\frac{(x-x^a)^2}{\sigma_a^2}}$ $\qquad \boxed{\sigma_a = ?}$

$$P(x|y) \propto P(x) \cdot \mathcal{L}(x|y) \propto e^{-\frac{1}{2}\frac{(x-x^b)^2}{\sigma_b^2} - \frac{1}{2}\frac{(y-x)^2}{\sigma_o^2}}$$

$$\propto e^{-\frac{1}{2}\left[\frac{(x-x^b)^2}{\sigma_b^2}+\frac{(y-x)^2}{\sigma_o^2}\right]}$$

$$\propto e^{-\frac{1}{2}\left[\frac{x^2}{\sigma_b^2}-\frac{2xx^b}{\sigma_b^2}+\frac{x^{b2}}{\sigma_b^2}+\frac{y^2}{\sigma_o^2}-\frac{2yx}{\sigma_o^2}+\frac{x^2}{\sigma_o^2}\right]}$$

1 2 3 4

$$\propto e^{-\frac{1}{2}\left[\frac{x^2}{\sigma_b^2}-\frac{2xx^b}{\sigma_b^2}-\frac{2yx}{\sigma_o^2}+\frac{x^2}{\sigma_o^2}+\frac{x^{b2}}{\sigma_b^2}+\frac{y^2}{\sigma_o^2}\right]}$$

$$\propto e^{-\frac{1}{2}\left[\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)x^2-\left(\frac{x^b}{\sigma_b^2}+\frac{y}{\sigma_o^2}\right)2x+\frac{x^{b2}}{\sigma_b^2}+\frac{y^2}{\sigma_o^2}\right]}$$

$$\boxed{\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)x^a=\frac{x^b}{\sigma_b^2}+\frac{y}{\sigma_o^2}}$$

$$\propto e^{-\frac{1}{2}\left[\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)x^2-\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)2xx^a+\frac{x^{b2}}{\sigma_b^2}+\frac{y^2}{\sigma_o^2}\right]}$$

$$\propto e^{-\frac{1}{2}\left[\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)x^2-\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)2xx^a\right]}\cdot \underbrace{e^{-\frac{1}{2}\left[\frac{x^{b2}}{\sigma_b^2}+\frac{y^2}{\sigma_o^2}\right]}}_{\text{CONSTANT}}$$

$$\propto K e^{-\frac{1}{2}\left[\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)x^2-\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)2xx^a\right]}$$

K

σ_a ?

PROP TO

$$\left[\propto e^{-\frac{1}{2}\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)\left[x^2-2xx^a\right]}\cdot \underbrace{e^{-\frac{1}{2}\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)x^{a2}}}_{Z}\right.$$

$$\propto e^{-\frac{1}{2}\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)\left(x^2-2xx^a+x^{a2}\right)}$$

$$\propto e^{-\frac{1}{2}\underbrace{\left(\frac{1}{\sigma_b^2}+\frac{1}{\sigma_o^2}\right)}_{(-1)}(x-x^a)^2}\propto e^{-\frac{1}{2}\frac{(x-x^a)^2}{\left(\frac{\sigma_o^2\cdot\sigma_b^2}{\sigma_o^2+\sigma_b^2}\right)}}$$

1ST MOMENT:

$$x^a=\left(\frac{\sigma_o^2\cdot\sigma_b^2}{\sigma_o^2+\sigma_b^2}\right)\left(\frac{x^b}{\sigma_b^2}+\frac{y}{\sigma_o^2}\right)$$

$$\underline{2^{nd}\ \text{MOMENT}} \qquad \sigma_a^2 = \frac{\sigma_0^2 \cdot \sigma_b^2}{\sigma_0^2 + \sigma_b^2}$$

$\ast$ $\underline{\text{EXAMPLE}}$ $\underline{\text{LINEAR MODEL}}$

$$y = \sum_{i=0}^{N} \alpha_i \cdot \phi_i(x) + \varepsilon \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

$\phi_i(x) = x_i$

α_i

$$\begin{aligned}
y_1 &= \alpha_1 x_1^{(1)} + \alpha_2 x_2^{(1)} \cdots + \alpha_N x_N^{(1)} + \varepsilon_1 \\
y_2 &= \alpha_2 x_1^{(2)} + \alpha_2 x_2^{(2)} \cdots + \alpha_N x_N^{(2)} + \varepsilon_2 \\
&\vdots \\
y_m &= \alpha x_1^{(m)} + \alpha_2 x_2^{(m)} + \cdots + \alpha_N x_N^{(m)} + \varepsilon_m
\end{aligned}$$

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_N^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_N^{(2)} \\ & & \ddots & \\ x_1^{(m)} & x_2^{(m)} & \cdots & x_N^{(m)} \end{bmatrix}}_{X} \cdot \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}}_{\alpha} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{bmatrix}}_{\varepsilon \in \mathbb{R}^{m \times 1}}$$

$\operatorname{cov}(\varepsilon_1, \varepsilon_2) = 0$

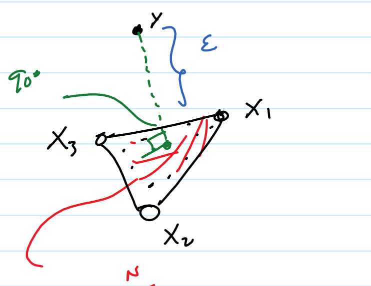
$\operatorname{cov}(\varepsilon_i, \varepsilon_j) = 0 \quad i \neq j$

$\operatorname{cov}(\varepsilon_i, \varepsilon_i) = \sigma^2 \quad i = j$

$$\Longrightarrow \quad y = X\alpha + \varepsilon \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2 \cdot I)$$

X: $x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_N$

$$\boxed{y = \sum_{i=1}^{N} \alpha_i x_i + \varepsilon} \qquad x_i \in \mathbb{R}^{m \times 1}$$



$$x = \sum_{i=1}^{N} \alpha_i x_i$$

$$\boxed{\varepsilon(\alpha) = y - X\alpha}$$

$$y - X\alpha \sim \mathcal{N}(0, \sigma^2 \cdot I)$$

$$y - X\alpha \sim \mathcal{N}(0, \sigma^2 I) \equiv \mathcal{L}(\alpha \mid X, y) \propto e^{-\frac{1}{2}\| y - X\alpha \|^2_{(\sigma^2 I)^{-1}}}$$

$$P(x) \propto e^{-\frac{1}{2}(x-\mu)^T Q^{-1}(x-\mu)} = e^{-\frac{1}{2}\|x-\mu\|^2_{Q^{-1}}}$$

↳ KERNEL OF NORMAL $\mathbb{R}^{n\times 1}$

$$y - X\alpha \sim \mathcal{N}(0, \sigma^2 I) \equiv \mathcal{L}(\alpha \mid X, y) \propto e^{-\frac{1}{2}\|y - X\alpha\|^2_{(\sigma^2 I)^{-1}}}$$

$$\underbrace{}_{}$$

$$P(\alpha \mid X, y) \propto \overset{\alpha\,1}{P(\alpha)} \cdot \mathcal{L}(x \mid y) \qquad \mathbb{E}\left((y - X\alpha)(y - X\alpha)^T\right) = \sigma^2 I$$

$$P(\alpha \mid X, y) \propto \mathcal{L}(x \mid y)$$

$$\propto e^{-\frac{1}{2}\|y - X\alpha\|^2_{(\sigma^2 I)}} = e^{-J(\alpha)}$$

$$J(\alpha) = \frac{1}{2}\|y - X\alpha\|^2_{(\sigma^2 I)} = \frac{1}{2}(y - X\alpha)^T (\sigma^2 I)^{-1}(y - X\alpha)$$

$$= \frac{1}{2\sigma^2}(y - X\alpha)^T (y - X\alpha)$$

$$\alpha^* = \arg\max_\alpha P(\alpha \mid X, y) \equiv \alpha^* = \arg\min_\alpha J(\alpha)$$

$$\nabla_\alpha \left(J(\alpha)\right) = \nabla_\alpha \left(\frac{1}{2\sigma^2}(y - X\alpha)^T (y - X\alpha)\right)$$

$\alpha \in \mathbb{R}^{N\times 1}$ $\quad y \in \mathbb{R}^{m\times 1}$ $\quad X \in \mathbb{R}^{m\times N}$

$$\boxed{z^T y = y^T z}$$

$$\nabla_\alpha \left(J(\alpha)\right) = \nabla_\alpha \left(\frac{1}{2\sigma^2}\left(y^T y - \underbrace{(X\alpha)}_{z \in \mathbb{R}^{m\times 1}}{}^T y - y^T \underbrace{(X\alpha)}_{z \in \mathbb{R}^{m\times 1}} + (X\alpha)^T (X\alpha)\right)\right)$$

$$= \nabla_\alpha \left(\frac{1}{2\sigma^2}\left(y^T y - 2(X\alpha)^T y + \alpha^T X^T X\alpha\right)\right)$$

$$= \nabla_\alpha \left(\frac{1}{2\sigma^2}\left(\alpha^T X^T X\alpha - 2\alpha^T X^T y + y^T y\right)\right)$$

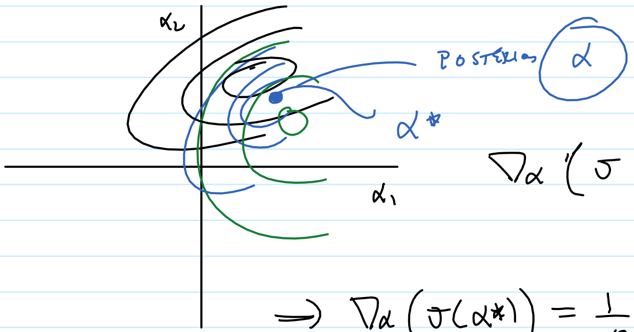
$$= \frac{1}{2\sigma^2}\left[\nabla_\alpha\left(\alpha^T X^T X\alpha\right) - 2\nabla_\alpha\left(\alpha^T X^T y\right) + \nabla_\alpha\left(\cancel{y^T y}\right)\right] \quad 0 \in \mathbb{R}^{N\times 1}$$

$$= \frac{1}{2\sigma^2}\left[2X^T X\alpha - 2X^T y\right] = \left[X^T X\alpha - X^T y\right]\cdot\frac{1}{\sigma^2}$$

$$= \frac{1}{2\sigma^2}\left[2X'X\alpha - 2X'y\right] = \left[X\ X\alpha - X\ y\right]\cdot\frac{1}{\sigma^2}$$

$$\boxed{\nabla_\alpha\left(J(\alpha)\right) = \frac{1}{\sigma^2}\cdot\left[X^TX\alpha - X^Ty\right]}$$

FIRST DERIVATIVE OF $\nabla_\alpha\left(J(\alpha)\right)$



$$\nabla_\alpha\left(J(\alpha^*)\right) = 0 \in \mathbb{R}^{N\times 1}$$

$$\Rightarrow \nabla_\alpha\left(J(\alpha^*)\right) = \frac{1}{\sigma^2}\left[\underbrace{X^TX\alpha^* - X^Ty}_{=0}\right] = 0 \in \mathbb{R}^{N\times 1}$$

$$\boxed{\alpha^* = (X^TX)^{-1}X^Ty}$$

+ $P(\alpha \mid X, y) \propto e^{-\frac{1}{2}\|\alpha-\alpha^*\|^2_{(Z)^{-1}}}$ ¿?

CONSTANT $\neq \alpha$

$$P(\alpha \mid X, y) \propto e^{-\frac{1}{2\sigma^2}\left(y^Ty - 2\alpha^TX^Ty + (X\alpha)^T(X\alpha)\right)}$$

α^*

$$\propto e^{-\frac{1}{2\sigma^2}\left(\alpha^TX^TX\alpha - 2\alpha^TX^Ty\right)}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left(\alpha^TX^TX\alpha - 2\alpha^T(X^TX)\alpha^*\right)}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left(\alpha^TX^TX\alpha - 2\alpha^TX^TX\alpha^*\right)}\cdot e^{-\frac{1}{2\sigma^2}\alpha^{*T}X^TX\alpha^*}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left(\alpha^TX^TX\alpha - 2\alpha^TX^TX\alpha^* + \alpha^{*T}X^TX\alpha^*\right)}$$

$$e^{-\frac{1}{2\sigma^2}\left(\overbrace{\alpha^TX^TX\alpha}^{\alpha^T} - \alpha^TX^TX\alpha^* - \alpha^TX^TX\alpha^* + \alpha^{*T}X^TX\alpha^*\right)}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left(\alpha^T X^T X\alpha - \alpha^T X^T X\alpha^* - \underbrace{\alpha^T X^T}_{z}\underbrace{X\alpha^*}_{y} + \underbrace{\alpha^{*T} X^T}_{y^T}\underbrace{X\alpha^*}_{z}\right)}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left(\alpha^T X^T X[\alpha-\alpha^*] + \underbrace{\alpha^{*T} X^T X}_{\alpha^{*T}X^TX}\alpha^* - \alpha^{*T} X^T X\alpha\right)}$$

$$\Longrightarrow$$

$$\propto e^{-\frac{1}{2\sigma^2}\left[\alpha^T X^T X[\alpha-\alpha^*] - \alpha^{*T} X^T X[\alpha-\alpha^*]\right]}$$

$$\propto e^{-\frac{1}{2\sigma^2}\left[(\alpha-\alpha^*)^T X^T X(\alpha-\alpha^*)\right]}$$

$$\propto e^{-\frac{1}{2}\left[(\alpha-\alpha^*)^T\left[\frac{1}{\sigma^2}X^T X\right](\alpha-\alpha^*)\right]}$$

$$\propto e^{-\frac{1}{2}\left[(\alpha-\alpha^*)^T\left[\sigma^2\cdot(X^T X)^{-1}\right](\alpha-\alpha^*)\right]}$$

$$\propto e^{-\frac{1}{2}\|\alpha-\alpha^*\|^2_{\left(\frac{1}{\sigma^2}X^T X\right)^{-1}}}$$

MODO POSTERIOR

$$\underbrace{P(\alpha\,|\,X,y)}_{\text{POSTERIOR DISTRIBUTION}} \propto e^{-\frac{1}{2}\|\alpha-\alpha^*\|^2_{\left(\frac{1}{\sigma^2}X^T X\right)^{-1}}}$$

POSTERIOR COVARIANCE

BAYESIAN INFERENCE

$\boxed{\text{MAP = MEAN = MEDIAN}}$

$$P(x\,|\,y) \propto P(x)\cdot\mathcal{L}(x\,|\,y)$$

$\Longrightarrow$ COMPUTE POSTERIOR PARAMETERS

↳ KERNEL $\longrightarrow$ KERNEL POSTERIOR

1. COMPUTE MAP
2. COMPUTE COVARIANCE

$$P(x|y) \propto P(x) \cdot \mathcal{L}(x|y) \mapsto P(\alpha|X,y) \propto \mathcal{L}(\alpha|X,y)$$

$$\boxed{P(\alpha) \propto 1}$$

UNCORRELATED ERRORS

TODAY'S CLASS

α

$P(\alpha) \propto 1$

$\mathbb{E}(yy^T) = \sigma^2 I$ $\quad y - X\alpha \sim \mathcal{N}(0, \sigma^2 I)$

SVM LR LOGIT

CORRELATED

$\mathbb{E}(yy^T) = Q$ $\quad y - X\alpha \sim \mathcal{N}(0, Q)$

$\mathbb{E}(yy^T) = \sigma^2 I$ $\quad y - X\alpha \sim \mathcal{N}(0, \sigma^2 I)$

$P(\alpha) \propto$ PDF

DATA ASSIMILATION

$\mathbb{E}(yy^T) = Q$ $\quad y - X\alpha \sim \mathcal{N}(0, Q)$

$\alpha \sim \mathcal{N}(\alpha^b, B)$