بسم الله الرحمن الرحیم

جلسه ۲۳ / ۸ / ۱۴۰۰

likelihood statistics

$\{D_i\}$ $i = 1, N$

$\{\theta_j\}$ $j = 1, M$ ← model free Parameters

$\{\theta\}_{Best}$

$$P(\{\theta\}|\{D\}) \simeq \mathcal{L}(\{D\}|\{\theta\})\, P(\{\theta\})$$

فعلاً $p(\{\theta\}) = cts$ یعنی فعلاً هیچ ارجحیت خاصی به مقادیر خاص پارامترها نمی‌دهیم

فرض: قضیه حد مرکزی (اغلب حفظ می‌شود) Gaussian statistics

استقلال اندازه‌گیری‌ها (اغلب نقض می‌شود)

$$\mathcal{L}(\{D\}|\{\theta\}) \simeq \prod_{i=1}^{N} e^{-\frac{(y_i - Y(x_i, \{\theta\}))^2}{2\sigma_i^2}}$$

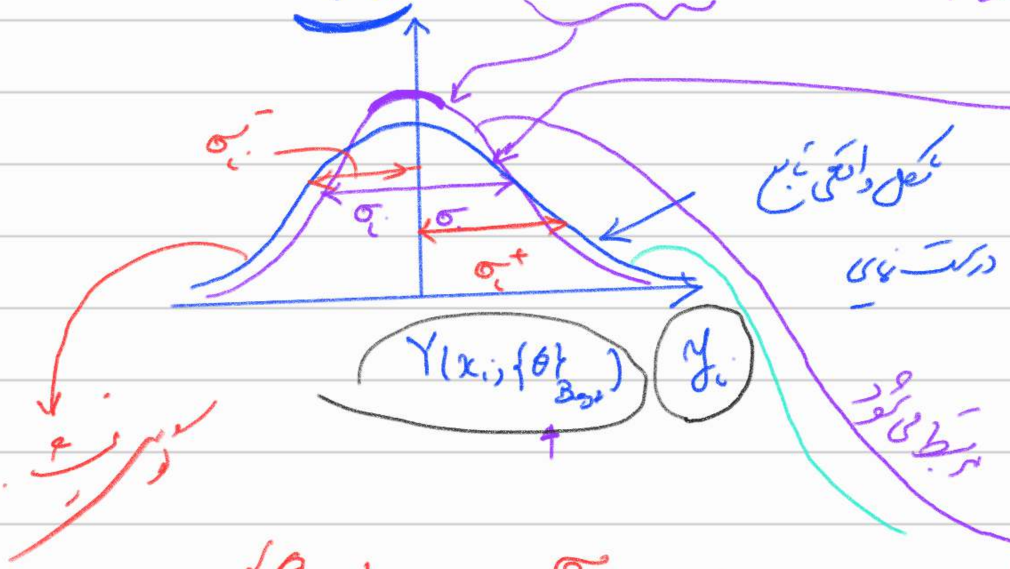
Theory

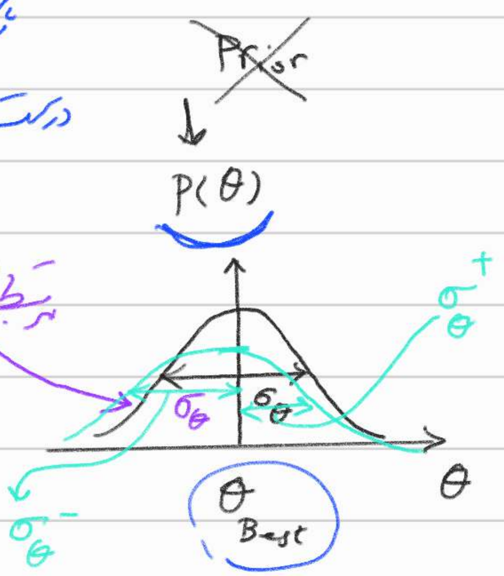
○ اگر هدف صرفاً یافتن مقدار بهینه باشد $\{\theta\}_{Best}$ ← در هم‌سنگی دیگر راه‌حل مقدار نظری سودی باشد از فرض نشود کافی است

○ اگر علاوه بر یافتن مقدار بهینه $\{\theta\}_{Best}$ بازه‌ی تراز تطابق (Confidence Interval)

بازه خطای پارامتر که مورد توجه باشد لذا گوسی بودن یا نبودن در نتیجه نهایی موثر است

$$P(y_i) \equiv \mathcal{L}(D_i|\{\theta\}) = \exp\left[-\frac{(y_i - Y(x_i,\{\theta\}))^2}{2\sigma_i}\right]$$





$\{\theta_{Best}\} \longrightarrow \sigma_{\theta_{Best}} = ?$

حالت گوسی دقیق $\boxed{\theta_{Best} \pm \sigma_{\theta_{Best}}}$

حالت

$\theta_{Best} \begin{matrix} +\sigma^+_{\theta_{Best}} \\ -\sigma^-_{\theta_{Best}} \end{matrix}$

در کتاب ۹

fisher forecast

به طور کلی قصد داریم ارزیابی از بازه خطای پارامتر که داشته باشیم بدون اینکه اندازه گیری انجام شود

فرض گوسی بودن را در نظر می‌گیریم

و صرفا ارزیابی σ_θ در دستور کار قرار می‌گیرد

θ_{Best} مهم نیست. چون آزمایش انجام نداده‌ایم

حتی علی الاصول نمی‌توانیم در مورد مقدار θ_{Best} صحبت کنیم

در قضیه حد مرکزی حاصل جمع (متوسط) تعداد زیادی از متغیرهای مستقل به شرط محدود بودن واریانس هر کدام، منجر به یک توزیع گوسی خوب کمیت متوسط این متغیر گرفته می‌شود.

$$\mathcal{L}(\{D\}|\{\theta\}) = e^{-\frac{\chi^2(\{\theta\})}{2}} \equiv e^{-\frac{\Delta^T \cdot C_D^{-1} \cdot \Delta}{2}}$$

$$\Delta_i \equiv y_i - Y(x_i, \{\theta\})$$

$$\underline{C_D = CoV_D} = \begin{bmatrix} \langle \delta y_1 \delta y_1 \rangle & \langle \delta y_1 \delta y_2 \rangle & \cdots & \langle \delta y_1 \delta y_i \rangle \\ \langle \delta y_2 \delta y_1 \rangle & \langle \delta y_2 \delta y_2 \rangle & & \cdots \\ \vdots & & & \\ \langle \delta y_N \delta y_1 \rangle & \cdots & & \langle \delta y_N \delta y_N \rangle \end{bmatrix}_{N\times N}$$

$$\Delta^T_{1\times N}, \quad C_{D_{N\times N}}, \quad \Delta_{N\times 1}$$

$$\Delta^T \cdot C_D^{-1} \cdot \Delta = (1\times 1)$$

$$\boxed{\sigma_{ii}^2 \equiv \langle \delta y_i \, \delta y_i \rangle \quad \text{خطای مربوط به اندازه‌گیری}}$$

$$\sigma_{\theta_i}^2 = \langle \delta\theta_i \, \delta\theta_i \rangle = \langle (\theta_i - \theta_i^{Best})(\theta_i - \theta_i^{Best}) \rangle$$

$$\sigma_{ij}^2 = \langle \delta y_i \, \delta y_j \rangle$$ — اندازه‌گیری‌ها مستقل از هم هستند

ماتریس کوواریانس اندازه‌گیری قطری است.

$$\sigma^2_{\theta_i\theta_j} = \langle \delta\theta_i \; \delta\theta_j \rangle$$

خود پارامترهای آزاد مدل هم از هم مستقل نیستند

$\cancel{\theta_j} \leftarrow \theta_i$ (و الکی داریم سی پارامتر که)

خیلی مهم است

$$Cov_\theta = C_\theta = \begin{bmatrix} \langle\delta\theta_1\,\delta\theta_1\rangle & \cdots & \langle\delta\theta_1\,\delta\theta_M\rangle \\ \langle\delta\theta_2\,\delta\theta_1\rangle & \langle\delta\theta_2\,\delta\theta_2\rangle & \cdots \\ \vdots & & \ddots \\ \langle\delta\theta_M\,\delta\theta_1\rangle & \cdots & \langle\delta\theta_M\,\delta\theta_M\rangle \end{bmatrix}_{M\times M}$$

$$\boxed{\Delta\theta_i \equiv \theta_i - \theta_i^{Best}}$$

$$\boxed{\text{نهایتاً هدف‌مان } \{\theta\}_{Best} \text{ و } Cov_\theta \text{ و } \sigma_{\theta_i\theta_j}}$$

حواسمان باشد که با اینها Cov_D ، σ_{ij} ، و فرض گرفتن بودن χ^2 شروع می‌کنیم.

یک نکته مهم:

$$P(\{\theta\}) = cts$$

$$P(\{\theta\}|\{D\}) = \mathcal{L}(\{D\}|\{\theta\}) \simeq \underbrace{e^{-\frac{\Delta^T \cdot C_D^{-1} \cdot \Delta}{2}}}_{\text{برحسب داده‌ها نوشته شده است}} \simeq \underbrace{e^{-\frac{\Delta\theta^T \cdot C_\theta^{-1} \cdot \Delta\theta}{2}}}_{\text{برحسب پارامتر نوشته شده}}$$

※ برگردیم به عبارت $\chi^2(\{\theta\})$ - می‌خواهیم کمینه کنیم:

فرض Cov_D قطری باشد

$$\chi^2(\{\theta\}) = \sum_{i=1}^{N} \frac{\left[y_i - Y(x_i,\{\theta\})\right]^2}{\sigma_i^2}$$

عناصر قطری Cov_D

Ex 1

دلخواه

فرض کردیم $\underbrace{Y(x_i,\{\theta\})}$ $= \sum_{j=1}^{M} \theta_j f_j(x_i)$ فرض دلخواه است

در حالت خطی تو ترانه هرگز ۲ باشد

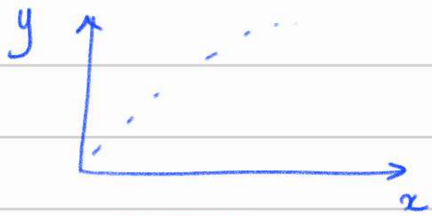
I	V
I_1	V_1
I_2	V_2
$\vdots$	
I_N	V_N

پارا ورودی

$$Y(x,\{\theta\}) = RI = R\ I$$

$M=1$

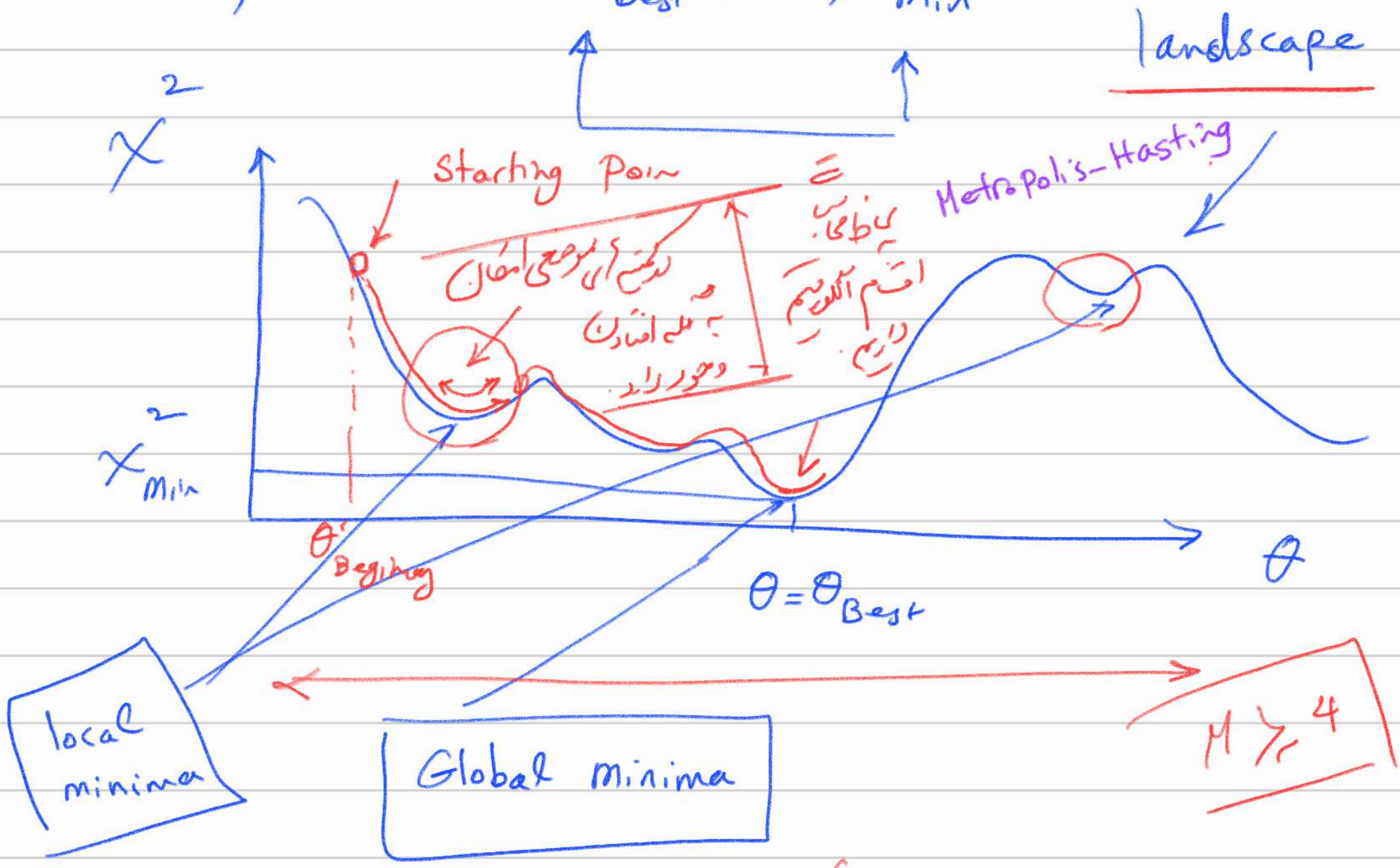
$(x,y) \to (I,V)$



$$\chi^2(\{\theta\}) = \sum_{i=1}^{N} \frac{\left[y_i - \sum_{k=1}^{M}\theta_k f_k(x_i)\right]^2}{\sigma_i^2}$$

$$= \sum_{i=1}^{N}\left[b_i - \sum_{k=1}^{M} A_{ik}\theta_k\right]^2 = \left[b_i - A_{ik}\theta_k\right]^2$$

$$\chi^2\left(\{\theta\} = \{\theta\}_{Best}\right) = \chi^2_{Min}$$



$O(6)$

$\boxed{O(10^{24})}$

مرتبه محاسبه زمان بر است

$$\left.\frac{\partial \chi^2(\{\theta\})}{\partial \theta_j}\right|_{\{\theta\}=\{\theta\}_{Best}} = 0 \quad , \quad b_i \equiv \frac{y_i}{\sigma_i}$$

$$A_{ik} \equiv \frac{f_k(x_i)}{\sigma_i}$$

$$\chi^2_{Min} = \checkmark \longrightarrow \boxed{Min \|b - A\theta\| \hookrightarrow \{\theta\}_{Best}}$$

Normal Equation

Singular Value Decomposit

Numerical methods

Determinist

Monte_Carlo

چون شکل تابع $Y = \sum_{k=1}^{M} f_k(x_i)\,\theta_k$ گرفتیم .. داریم.

$$\left.\frac{\partial \chi^2}{\partial \theta_j}\right|_{\{\theta\} = \{\theta\}_{Bst}} = 0 \qquad = 2\sum_{i=1}^{N}\left[y_i - \sum_{k=1}^{M}\theta_k f_k(x_i)\right]\frac{f_j(x_i)}{\sigma_i^2} = 0$$

$$\sum_{i=1}^{N}\frac{y_i f_j(x_i)}{\sigma_i^2} = \sum_{i=1}^{N}\sum_{k=1}^{M}\frac{\theta_k f_k(x_i) f_j(x_i)}{\sigma_i^2}$$

$$\beta_j \equiv \sum_{i=1}^{N}\frac{y_i f_j(x_i)}{\sigma_i^2}, \qquad \alpha_{jk} \equiv \sum_{i=1}^{N}\frac{f_j(x_i) f_k(x_i)}{\sigma_i^2}$$

$$\beta_j = \sum_{k=1}^{M}\alpha_{jk}\,\theta_k$$

$$\beta_j = \alpha_{jk}\,\theta_k$$

حست کاملاً تحلیلی

$$\boxed{[\theta] = [\alpha]^{-1}[\beta]}$$

Normal Eq

$$\boxed{[\theta]_{Best} = [\alpha]^{-1}[\beta]}$$

ادامهٔ بحث خطای مربوط به $[\theta]_{Best}$ ها:

$$\beta_j \equiv \sum_{i=1}^{N} \frac{y_i f_j(x_i)}{\sigma_i^2} \quad , \quad \alpha_{jk} = \sum_{i=1}^{N} \frac{f_j(x_i) f_k(x_i)}{\sigma_i^2}$$

$$\sigma^2(\theta_i) = \sum_{j=1}^{N} \sigma_j^2 \left(\frac{\partial\theta_i}{\partial y_j}\right)^2 + \sum_{j=1}^{N} \sum_{l \neq j}^{N} \left(\frac{\partial\theta_i}{\partial y_j}\frac{\partial\theta_i}{\partial y_l}\right) Cov(y_j, y_l)$$

خطای مربوط به اندازه‌گیری‌ها

عناصر غیرقطری

خیلی وقت‌ها کنار می‌گذاریم در مقیاس‌های معمول.

$$\frac{\partial\theta_i}{\partial y_j} = \sum_{k=1}^{M} [\alpha_{ik}]^{-1} \frac{\partial\beta_k}{\partial y_j}$$

$$\boxed{\left(\frac{\partial\theta_i}{\partial y_j}\right)^2 = \sum_{l=1}^{M}\sum_{k=1}^{M} [\alpha_{ik}]^{-1}[\alpha_{il}]^{-1} \frac{\partial\beta_k}{\partial y_j}\frac{\partial\beta_l}{\partial y_j}}$$

$$\sigma^2(\theta_i) = \sum_{l=1}^{M} \sum_{k=1}^{M} [\alpha_{ik}]^{-1} [\alpha_{il}]^{-1} \underbrace{\sum_{j=1}^{N} \frac{\partial \beta_k}{\partial y_j} \frac{\partial \beta_l}{\partial y_j}}$$

$$= \sum_{l=1}^{M} \sum_{k=1}^{M} [\alpha_{ik}]^{-1} [\alpha_{il}]^{-1} [\alpha_{kl}]$$

δ_{il}

$$\boxed{\sigma^2(\theta_i) = [\alpha_{ii}]^{-1}} \neq \alpha_{ii}^{-1}$$

این فرمول به صورت تحلیلی و صرفاً در حالتی که

$$\boxed{Y_{Th}(x_i; \{\theta\}) = \sum_{j=1}^{M} \theta_j f_j(x_i)}$$

صادق است

در حالت کلی Y_{the} می‌تواند خودش حل معادلات دیفرانسیل و انتگرالی به صورت عددی به دست بیاید

پس دیگر به شکل بالا نوشته نمی‌شود. ← کاملاً عددی مسئله جلو می‌رود.

Ex 2. $\quad Y(x, \{\theta\}) = \theta_1 + \theta_2 x^1 + \theta_3 x^2 + \dots \theta_M x^{M-1}$

چندجمله‌ای مرتبه $M-1$ از x که M تا پارامتر آزاد دارد.

$$\mathcal{L}(\{D\} | \{\theta\}) = \prod_{i=1}^{N} \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\frac{1}{2} \frac{\left[y_i - \sum_{k=1}^{M} \theta_k x_i^{k-1} \right]^2}{\sigma_i^2} \right]$$

$M = 3$

$$\chi^2(\theta_1, \theta_2, \theta_3) = \sum_{i=1}^{N} \frac{\left[y_i - \sum_{k=1}^{3} \theta_k x_i^{k-1}\right]^2}{\sigma_i^2}$$

$\theta_1 = ?$ $\quad \theta_2 = ?$ $\quad \theta_3 = ?$ ، Y_{theory} و تحلیل دادیم

$K=1$

$$0 = \frac{\partial \chi^2}{\partial \theta_1} = -2 \sum_{i=1}^{N} \frac{\left[y_i - \sum_{k=1}^{3} \theta_k x_i^{k-1}\right] x_i^0}{\sigma_i^2}$$

$$0 = \frac{\partial \chi^2}{\partial \theta_2} = -2 \sum_{i=1}^{N} x_i \frac{\left[y_i - \sum_{k=1}^{3} \theta_k x_i^{k-1}\right]}{\sigma_i^2}$$

$$0 = \frac{\partial \chi^2}{\partial \theta_3} = -2 \sum_{i=1}^{N} x_i^2 \frac{\left[y_i - \sum_{k=1}^{3} \theta_k x_i^{k-1}\right]}{\sigma_i^2}$$

$$\theta_1 = \frac{1}{\Delta} \begin{vmatrix} \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \\ \sum_{i=1}^{N} y_i \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^3}{\sigma_i^2} \\ \sum_{i=1}^{N} y_i \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^3}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^4}{\sigma_i^2} \end{vmatrix}$$

$$\theta_2 = \frac{1}{\Delta} \begin{vmatrix} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} & \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \\ \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N} y_i \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^3}{\sigma_i^2} \\ \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N} y_i \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^4}{\sigma_i^2} \end{vmatrix}$$

$$\theta_3 = \frac{1}{\Delta}\begin{vmatrix} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2} \\ \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N} y_i \frac{x_i}{\sigma_i^2} \\ \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^{N} \frac{x_i^3}{\sigma_i^2} & \sum_{i=1}^{N} y_i \frac{x_i^2}{\sigma_i^2} \end{vmatrix}$$

$$\Delta = \begin{vmatrix} \sum_{i=1}^{N} \frac{1}{\sigma_i^2} & \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} \\ \sum \frac{x_i}{\sigma_i^2} & \sum \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} \\ \sum \frac{x_i^2}{\sigma_i^2} & \sum \frac{x_i^3}{\sigma_i^2} & \sum x_i^4/\sigma_i^2 \end{vmatrix}$$

$$\boxed{\sigma_{\theta_i}^2 = \sum_{j=1}^{N} \sigma_j^2 \left(\frac{\partial \theta_i}{\partial y_j}\right)^2}$$

تقارن (ضمنی فرض کردیم بودن دادگان است)

$$\left(\sigma_{\theta_i}^{+}\right)^2 = \sum_{j=1}^{N} \left(\sigma_j^{+}\right)^2 \left(\frac{\partial \theta_i}{\partial y_j}\bigg|_{y_j^{+}}\right)^2$$

اگر هم متقارن نباشد

$$(\sigma^-_{\theta_i})^2 = \sum (\sigma^-_j)^2 \left(\frac{\partial \theta_i}{\partial y_j} \Big|_{y^-_j} \right)^2$$

$$\boxed{\begin{aligned} \frac{\partial \theta_i}{\partial y_j}\Big|_{y^+_j} &\equiv \frac{\theta_i(y_j + \Delta y_j) - \theta_i(y_j)}{\Delta y_j} + \cdots \\ \frac{\partial \theta_i}{\partial y_j}\Big|_{y^-_j} &\equiv \frac{\theta(y_j) - \theta(y_j - \Delta y_j)}{\Delta y_j} \to \cancel{O(\Delta y)} \end{aligned}}$$

یک تخمین خوب از Confidence Interval of free Parameter رو میده.

Ex 3 linear fitting

$\{D\} = \{(x_i \pm \sigma_{x_i}, y_i \pm \sigma_{y_i})\}, \quad i = 1, N$

$\{\theta\} = \{m, c\}. \qquad Y_{theory}(\{\theta\}, x) = m x_i + c$

شیب (m)

عرض از مبدا (c)

$m_{Best} \pm \sigma_m = ?$

$c_{Best} \pm \sigma_c = ?$

$p(m) = cts$ No - Prior information

$p(c) = cts$ No - Prior information

Centeral limit theorem

$Cov_D \equiv$ Diagonal.

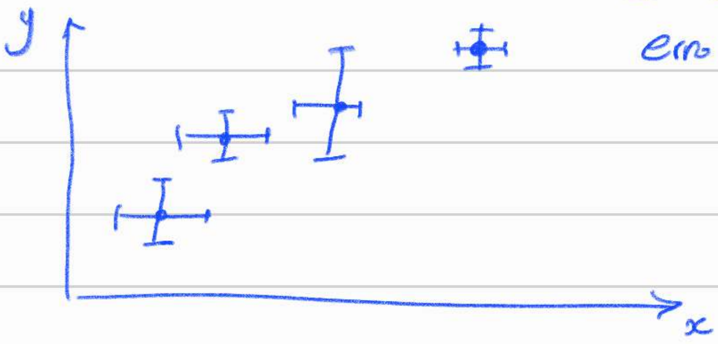
Posterior $\longrightarrow$ Likelihood $\longrightarrow$ Gaussian Distribution

$$\mathcal{L}(\{D\}|\{\theta\}) = e^{-\frac{\chi^2(m,c)}{2}}$$

$$\chi^2(m,c) = \sum_{i=1}^{N} \frac{[y_i - (mx_i + c)]^2}{\underbrace{\sigma_{y_i}^2 + \sigma_{x_i}^2}}$$

اغلب در نظر نمی‌گیریم

error Propagation



Suppose $\begin{cases} \sigma_{y_i} = \sigma = cts \\ \sigma_{x_i} = 0 \end{cases}$

خطای مربوط به y همه هم‌اندازه

$\sigma_{x_i} = 0 \longleftarrow$ Independent Parameter

$$\frac{\partial \chi^2}{\partial m} = 0 \implies 0 = -2\sum y_i x_i + 2m\sum x_i^2 + 2c\sum x_i$$

$$\frac{\partial \chi^2}{\partial c} = 0 \implies 0 = -2\sum y_i + 2m\sum x_i + 2cN$$

$$\boxed{m_{Best} = \frac{\sum (x_i - \bar{x})\, y_i}{\sum (x_i - \bar{x})^2}} \qquad \bar{x} = \frac{1}{N}\sum x_i$$

$$\boxed{c_{Best} = \bar{y} - m_{Best}\,\bar{x}}, \qquad \bar{y} = \frac{1}{N}\sum y_i$$

$$\sigma_m^2 = ? = \left(\sigma_m^{(sys)}\right)^2 + \left(\sigma_m^{(stat)}\right)^2$$

$$\sigma_c^2 = ? = \left(\sigma_c^{(sys)}\right)^2 + \left(\sigma_c^{(stat)}\right)^2$$

$$\left(\sigma_m^{(stat)}\right)^2 = \underbrace{\left(\sigma_m^{(Propagat)}\right)^2}_{} + \underbrace{\left(\sigma_m^{(Intrinsic)}\right)^2}_{}$$

$x_i, y_i \to m$ قبلاً دیدیم

خطای آماری انتشاری ناشی از

متغیر اولیه ورودی متغیرهای ثانویه.

مشتق قبل دیدیم

شهود پیدا کنند

$x_1, x_2 \ldots x_N$

خطای آماری ذاتی

$$\left\{\begin{aligned} &\bar{x} = \frac{1}{N}\sum x_i \\ &\sigma_{\bar{x}}^2 = \frac{1}{N(N-1)}\sum (x_i - \bar{x})^2 \end{aligned}\right.$$

توزیع ذهن بود

$$\sigma_m^2(\mathrm{int}) = \underbrace{\left(\frac{\partial m}{\partial y_i}\,\sigma_{y_i}\right)^2 + \left(\frac{\partial m}{\partial x_i}\,\sigma_{x_i}\right)^2}$$ → صرف‌نظر کردن

شکل آن مثل انتشار خطای متغیرهای اولیه (بدون نمونه) ولی $\underline{\sigma_{y_i}}$ و $\underline{\sigma_{x_i}}$ به آن

خطاهای اولیه نیستند بر اساس صرفاً آماری به دست می‌آیند.

فرض

$\sigma_{x_i} \simeq 0$, $\sigma_{y_i} = \sigma$ → خطای یکسان همه فرض شده است

$$\boxed{\sigma_m^2(\mathrm{int}) = \frac{\sigma^2}{D}} \quad ☆$$

$$D \equiv \sum (x_i - \bar{x})^2$$

$$\cancel{y_i \Rightarrow \sigma^2 \neq \frac{1}{N(N-1)} \sum_1 (y_i - \bar{y})^2 \qquad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i}$$

$$\sigma^2 \equiv \frac{1}{N-2} \sum \Big[y_i - \overbrace{(m_{Best}\, x_i + C)}^{Y_{Theo}(x_i;\{\theta\})}\Big]^2$$

$$☆ \quad \boxed{\sigma_c^2(\mathrm{int}) = \left(\frac{1}{N} + \frac{\bar{x}^2}{D}\right)\frac{1}{N-2} \sum \Big[y_i - (m_{Best}\, x_i - C)\Big]^2} \quad ☆$$