بسم الله الرحمن الرحیم

9- Fisher forecast

(Fisher information Matrix)

0- Main Refrences

arxiv:0906.0993

optimization course: 1400830, 1400902, 1400907

1400909A, 1400928, 1400930

1: Recall that for a typical Data modeling problem:

We are interested in $\{\theta\}:\{\theta_1 \dots \theta_M\}$

$\{D\}:\{D_1 \dots D_N\}$

☆ $\{\theta\}_{Best}$ s? and $\{\sigma_\theta\}$ s?

More General Case

☆ $\{\sigma_\theta\} \longrightarrow \text{Cov}: \langle \delta\theta \otimes \delta\theta \rangle$

☆ $$\text{Cov}_{\text{Parameters}} : \begin{bmatrix} \sigma_{\theta_1}^2 & \sigma_{\theta_1\theta_2} & \sigma_{\theta_1\theta_3} & \dots & \sigma_{\theta_1\theta_M} \\ \sigma_{\theta_2\theta_1} & \sigma_{\theta_2}^2 & \dots & & \\ & & \sigma_{\theta_3}^2 & \dots & \\ \dots & & & \dots & \sigma_{\theta_M}^2 \end{bmatrix}_{M\times M}$$

$$\star \quad \underbrace{P(\{\theta\}|D)}_{\checkmark} = \frac{\mathcal{L}(D|\{\theta\})\, P(\{\theta\})}{\underbrace{\int d\{\theta\}\, \mathcal{L}(D|\{\theta\})\, P(\{\theta\})}_{\text{Evidence measure}}}$$

Ensemble average:

$$\widetilde{\langle\{\theta\}\rangle} \overset{?}{=} \int d\{\theta\}\, \{\theta\}\, P(\{\theta\}|D)$$

Most probable value $\rightarrow$

$$\boxed{\{\theta_{Best}\} \Rightarrow \left.\frac{dP(\{\theta\}|D)}{d\{\theta\}}\right|_{\{\theta\}=\{\theta_{Best}\}} = 0}$$

2: Recall that: from observation (Experiment) Point of View

$\{D\} \cdot \{(x_i, y_i)\} \quad i = 1 \ldots N$

$$\mathrm{Cov}\big|_{Data} = \langle \delta y \otimes \delta y \rangle$$

$$= \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdot & - & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & & & \sigma_{2N} \\ \sigma_{31} & \cdot\,\cdot & \sigma_3^2 & \cdot & \\ & & & \ddots & \\ \sigma_{N1} & & & & \sigma_N^2 \end{bmatrix}_{N\times N}$$

$$\boxed{\sigma_{ij}^2 = (\sigma_{ij}^{stat.})^2 + (\sigma_{ij}^{sys})^2}$$

3: By Fisher Information Matrix, we are interested in determining

$C_\theta \equiv \mathrm{Cov}\big|_{Parameters}$ even without doing

Measurement !

☆ We don't want to achieve (find) the best fit values for model's free parameters : $\{\theta_{Best}\}$: { It is given by a-pariori. that is it! }

☆ Significance of Confidence Interval on free Parameters : For a given Model (Theory), we look for the capability of Model (observation) for constraining the model parameters

☆ We totally assume that the posterior of model's free parameters is compatible by { Multi-Variate Gaussian function }

④ Mathematical Notion of "Fisher Forecast"

☆ $\{D\} : \{(x_i, y_i), \underbrace{\text{Cov}|_{\text{Data}}}_{\text{Cov}_D}\}$ $i = 1 \ldots N$

☆ $\{\theta\} : \{\theta_j\}$, $j = 1, \ldots, M$

☆ $P(\{\theta\})$: Prior on Model's free Parameters $\longrightarrow$ Constant

☆ MultiVariate Gaussian form for
- Posterior
- likelihood

☆ $\mathcal{L}(\{D\}|\{\theta\})$ = Multivariate Gaussian form

$$= \frac{1}{\sqrt{(2\pi)^N \text{Det}(\text{Cov}_D)}} e^{-\frac{\Delta^T \cdot \text{Cov}_D^{-1} \cdot \Delta}{2}}$$

$$\begin{cases} \Delta \equiv y_{obs} - y_{\text{Theory}} & (N \times 1) \\ \Delta^T \longrightarrow & (1 \times N) \\ \text{Cov}_D \longrightarrow & (N \times N) \end{cases}$$

☆ $\mathcal{L}(\{D\}|\{\theta\}) \propto e^{-\frac{\chi^2(\{\theta\})}{2}}$

$$\boxed{\chi^2(\{\theta\}) \equiv \Delta^T \cdot \text{Cov}_D^{-1} \cdot \Delta}$$

☆ $\mathcal{L}_{\text{Rel}}(\{D\}|\{\theta\}) \equiv \frac{\mathcal{L}(\{D\}|\{\theta\})}{\mathcal{L}_{\text{Max}}(\{D\}|\{\theta\})}$

☆ $\mathcal{L}_{\text{Rel}} = e^{-\frac{\Delta\chi^2}{2}}$ $\Delta\chi^2 = \chi^2(\{\theta\}) - \chi^2_{min}(\{\theta\})$

$$\chi^2_{min}(\{\theta\}) = \chi^2(\{\theta\} = \{\theta_{Best}\})$$

$$\left\{ \begin{array}{l} \left.\dfrac{\partial \chi^2}{\partial \{\theta\}}\right|_{\{\theta\}=\{\theta\}_{Best}} = 0 \\[2ex] \left.\dfrac{\partial^2 \chi^2}{\partial \{\theta\}^2}\right|_{\{\theta\}=\{\theta\}_{Best}} \succ 0 \end{array} \right.$$

☆ Connected point to Fisher Information Matrix

(A) $$\mathcal{L}_{Rel} = e^{-\frac{\Delta\chi^2}{2}} \propto \underbrace{e^{-\frac{\Delta\theta^T \cdot F \cdot \Delta\theta}{2}}}_{\text{Target Posterior Probability in Multivariate Gaussian form}} \propto e^{-\frac{\Delta\theta^T \cdot Cov_\theta^{-1} \cdot \Delta\theta}{2}}$$

$\Delta\theta = \theta - \theta_{Best}$ $(M \times 1)$

$\Delta\theta^T \longrightarrow (1 \times M)$

$\boxed{F \longrightarrow (M \times M) \quad : \text{Fisher information Matrix}}$

$\boxed{Cov_\theta = F^{-1}} \longrightarrow$ $\sigma^2_{\theta_1} = (F^{-1})_{11} \neq (F_{11})^{-1}$

$\sigma_{\theta_i\theta_j} = (F^{-1})_{ij}$

$$\langle \delta\theta_i \delta\theta_j \rangle = ? = \underbrace{(F^{-1})_{ij}}$$

$$\langle (\theta_i - \theta_i^{Best})(\theta_j - \theta_j^{Best}) \rangle$$

⑤ Second Mathematical foundation:

$$\ln \mathcal{L}_{Rel}(\{\theta\} = \{\theta\} + \{\theta\}_{Best}) = \underbrace{\ln \mathcal{L}_{Rel}(\{\theta\} = \{\theta\}_{Best})}_{1}$$

$$+ \{\delta\theta\} \cancel{\left. \frac{\partial \ln \mathcal{L}_{Rel}}{\partial \{\theta\}} \right|_{\{\theta\} = \{\theta\}_{Best}}} \to 0 \; +$$

$$+ \frac{\{\delta\theta\}^T \{\delta\theta\}}{2!} \left. \frac{\partial^2 \ln \mathcal{L}_{Rel}}{\partial\{\theta\} \partial\{\theta\}} \right|_{\{\theta\} = \{\theta\}_{Best}} + O(\delta\theta^3)$$

$$\boxed{\ln \mathcal{L}_{Rel} \simeq \frac{1}{2} \sum_{ij} \left. \frac{\partial^2 \ln \mathcal{L}_{Rel}}{\partial\theta_i \partial\theta_j} \right|_{\{\theta\} = \{\theta\}_{Best}} \delta\theta_i \delta\theta_j}$$ Ⓑ

A $\ln \mathcal{L}_{Rel} = \ln\left(e^{-\frac{\Delta\chi^2}{2}}\right) = \ln\left(e^{-\frac{\Delta\theta^T \cdot F \cdot \Delta T}{2}}\right)$

$$\boxed{\ln \mathcal{L}_{Rel} = -\frac{\Delta\theta^T \cdot F \cdot \Delta\theta}{2}}$$ C

B and C $$F_{ij} = -\left\langle \frac{\partial^2 \ln \mathcal{L}_{Rel}}{\partial\theta_i \, \partial\theta_j} \Big|_{\{\theta\}=\{\theta\}_{Best}} \right\rangle_{ens}$$

Ex 1: $M=1$

$$\mathcal{L}_{Ret} \propto e^{-\frac{(\theta-\theta_{Best})^2}{2} F_{11}}$$

$$\ln \mathcal{L}_{Rel} = -\frac{(\theta-\theta_{Best})^2 F_{11}}{2}$$

$$\frac{d \ln \mathcal{L}_{Rel}}{d\theta}\Big|_{\theta=\theta_{Best}} = -(\theta-\theta_{Best}) F_{11}\Big|_{\theta=\theta_{Best}} = 0$$

$$\frac{d^2 \ln \mathcal{L}_{Rel}}{d\theta^2}\Big|_{\theta=\theta_{Best}} = -F_{11}$$

$$F_{ij} = -\left\langle \frac{\partial^2 \ln \mathcal{L}_{Rel}}{\partial\theta_i\,\partial\theta_j} \right\rangle_{\text{ens}}\Bigg|_{\theta=\theta_{Best}}$$

$$\begin{cases} \mathcal{L}_{Rel} \sim e^{-\frac{\Delta\chi^2}{2}} \\ \mathcal{L}_{Rel} \sim e^{-\frac{\Delta\theta^T \cdot F \cdot \Delta\theta}{2}} \end{cases}$$

✓

$$F = ✓ \longrightarrow F^{-1} \longrightarrow \underbrace{\mathrm{Cov}\big|_{\theta}}_{} = ✓$$

Uncertainty in Model's free parameters

⑥ Third Mathematical Foundation

$$\mathcal{L}_{Rel} \sim e^{-\frac{\Delta\chi^2}{2}} \longrightarrow ?$$