457.646 Topics in Structural Reliability

In-Class Material: Class 25

VII. Random fields (contd.)

④ Karhunen-Loève (KL) expansion (Gaussian RFs)

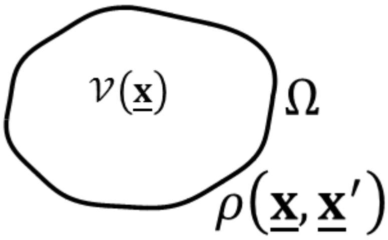
→ Describe RF in terms of finite # of shape functions

defined over ________ domain

(no geometric discretization)

→ Discretization based on

_____________ structure $\rho(\mathbf{x},\mathbf{x}')$



Goal: Want to descrive $\rho(\mathbf{x},\mathbf{x}')$ by

$$\rho(\mathbf{x},\mathbf{x}') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}')$$

Orthogonal shape (base) functions

Can find λ, φ by solving an integral eigenvalue problem, i.e.

$$\int_{\Omega} \rho(\mathbf{x},\mathbf{x}')\varphi_i(\mathbf{x}')d\mathbf{x}' = \lambda_i \varphi_i(\mathbf{x})$$ (Fredholem integral eqn – 2^{nd} kind)

Note $\rho(\mathbf{x},\mathbf{x}')$ is bounded, symmetric, (+) definite.

If so, one can find

$\varphi_i(\mathbf{x})$: orthogonal $\int \varphi_i(\mathbf{x})\varphi_j(\mathbf{x})d\mathbf{x} = \delta_{ij}$

λ_i: real & positive

Can drop λ_i's if $\lambda_r \cong 0$

Then using $\varphi_i(\mathbf{x})$, and λ_i, $i=1,\ldots,r$, one can describe Gaussian RF $v(x)$ by

KL expansion of Gaussian RF

$$\boxed{v(\mathbf{x}) \approx \hat{v}(\mathbf{x}) = \mu(\mathbf{x}) + \sigma(\mathbf{x}) \sum_{i=1}^{r} (u_i \sqrt{\lambda_i}\varphi_i(\mathbf{x})), \quad x \in \Omega} \Rightarrow v(\mathbf{x}) \Rightarrow \{u_1, \cdots, u_r\}$$

$u_i \rightarrow N(0,1)$, u_i s.i

Let's check!

Instructor: Junho Song
junhosong@snu.ac.kr

i. Gaussian? Yes, function of u_i's

ii. $E[\hat{v}(\mathbf{x})] = \mu(\mathbf{x})$? $E[\hat{v}(\mathbf{x})] =$

iii. $Var[\hat{v}(\mathbf{x})] = E[(\quad)^2]$

$$= E[\quad \sum_{i=1}^{r}\sum_{j=1}^{r} \quad]$$

$$= \sigma^2(\mathbf{x})\sum_{i=1}^{r}\sum_{j=1}^{r} \quad \sqrt{\lambda_i}\sqrt{\lambda_j}\varphi_i(\mathbf{x})\varphi_j(\mathbf{x})$$

$$= \sigma^2(\mathbf{x})\sum_{i=1}^{r}\lambda_i\varphi_i^2(\mathbf{x})$$

$$= \sigma^2(\mathbf{x})$$

(because $\rho(\mathbf{x},\mathbf{x}) =$ $\quad = \quad$)

iv. $\rho_{\hat{v}\hat{v}}(\mathbf{x},\mathbf{x}') \stackrel{?}{=} \rho(\mathbf{x},\mathbf{x}')$

$$= E[(\hat{v}(\mathbf{x}) - \mu(\mathbf{x}))(\hat{v}(\mathbf{x}') - \mu(\mathbf{x}'))] / \sigma(\mathbf{x})\sigma(\mathbf{x}')$$

$$= E[\sum_{i=1}^{r}\sum_{j=}^{r} u_i\sqrt{\lambda_i}\varphi_i(\mathbf{x})u_j\sqrt{\lambda_j}\varphi_j(\mathbf{x}')]$$

$$= \sum_{i=1}^{r}\sum_{j=1}^{r} E[\quad]\sqrt{\lambda_i}\sqrt{\lambda_j}\varphi_i(\mathbf{x})\varphi_j(\mathbf{x}')$$

$$= \sum_{i=1}^{r}\lambda_i\varphi_i(\mathbf{x})\varphi_i(\mathbf{x}')$$

$$= \varphi(\mathbf{x},\mathbf{x}')$$

- # of RV's:
- Represented by function
- No necessary
- Most efficient (in terms of # of)
- Requires solution of an integral eigenvalue problem.

Application examples:

Yi, S., and J. Song (2018) Particle filter based monitoring and prediction of spatiotemporal corrosion using successive measurements of structural responses. *Sensors*, Vol. 18(11), 3909.

Lee, S.-H., and J. Song (2017). System identification of spatial distribution of structural parameters using modified Transitional Markov Chain Monte Carlo (m-TMCMC) method. *ASCE Journal of Engineering Mechanics*. Vol. 143(9), 04017099-1~18.

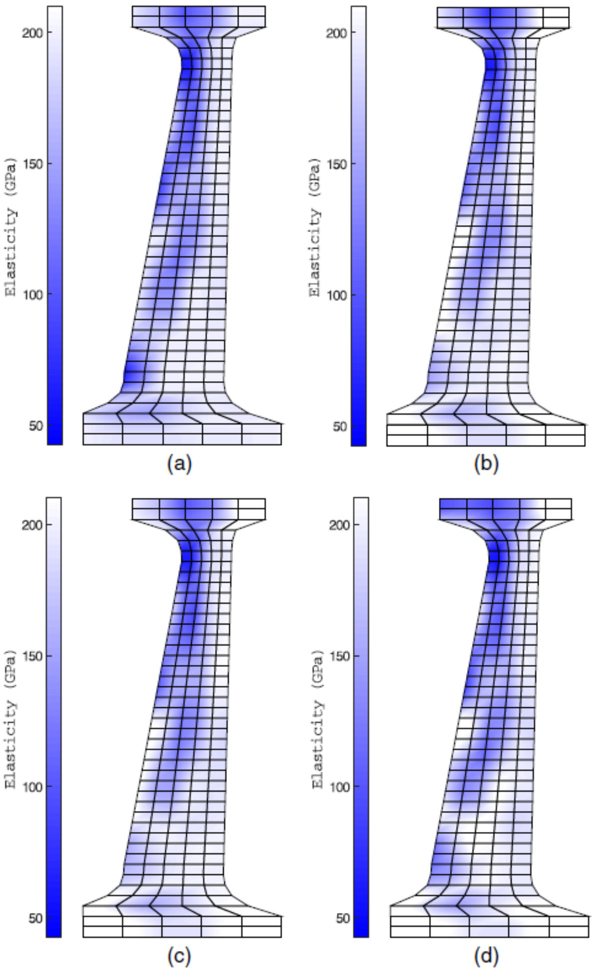


Fig. 25. Exact and estimated spatial distribution by m-TMCMC-SI: Scenario 8: (a) exact; (b) 1% error; (c) 3% error; (d) 5% error

⑤ Orthogonal expansion (eigen-expansion, but correlated rv's)

⑥ Optimal linear estimation (OLE)~ linear regression

⑦ Expansion OLE

⋮ See Sudret & ADK (2000)

◎ **Nataf RF**

$v(\mathbf{x}) \Rightarrow F(v, \mathbf{x})$, $\rho_{ZZ}(\mathbf{x}, \mathbf{x}')$

$v(\mathbf{x}) = F_v^{-1}\{\Phi(\hat{Z}(\mathbf{x}))\}$, $Z(\mathbf{x}) \sim N(\mathbf{0}, \rho_{ZZ}(\mathbf{x}, \mathbf{x}'))$ ($Z(\mathbf{x}) \rightarrow$ Gaussian RF)

⇒ Construct $Z(\mathbf{x})$ and discrete to $\hat{Z}(\mathbf{x})$

⇒ $v(\mathbf{x}) = F^{-1}\{\Phi(\hat{Z}(\mathbf{x}))\}$

Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

VIII. Response Surface Method (CRC Ch.19 & Mike Tipping's chapter)

◎ Reliability Analysis, Uncertainty Quantification & Response Surface

Reliability Analysis

$P_f = \int\limits_{g(\mathbf{x})\le 0} f_{\mathbf{x}}(\mathbf{x})d\mathbf{x}$ → e.g. FORM/SORM $g(\mathbf{x}_i)$, $\nabla g(\mathbf{x}_i)$

→ e.g. Sampling $q_i = I(\mathbf{x}_i)$ or $\dfrac{I(\mathbf{x}_i)\cdot f(\mathbf{x}_i)}{h(\mathbf{x}_i)}$

where $I(\mathbf{x}_i) = \begin{cases} 1 & g(\mathbf{x}_i) \le 0 \\ 0 & g(\mathbf{x}_i) > 0 \end{cases}$

Uncertainty Quantification

"Process of determining the effect of input uncertainties"

on response metrics of interest (Eldred et al. 2008)

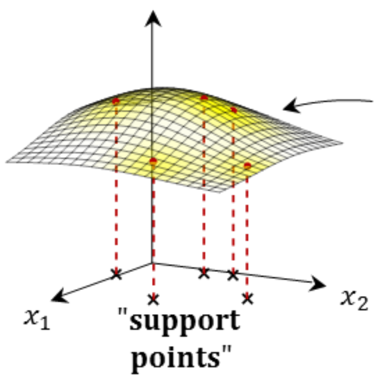
e.g. $E[g(\mathbf{x})^m] = \int\limits_{\mathbf{x}} g(\mathbf{x})^m f_{\mathbf{x}}(\mathbf{x})d\mathbf{x}$

① $g(\mathbf{x})$ Sometimes

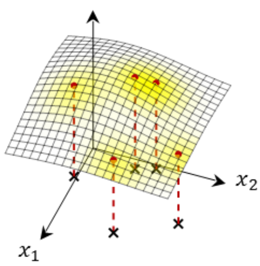
- Computationally costly for MCS
- No analytical gradients but many RVs

  ⇒ FORM/SORM difficult
- Experiments expensive (statistical analysis of experiment data infeasible)

② Idea: $g(\mathbf{x}) \approx \eta(\mathbf{x})$ ($\eta(\mathbf{x})$ ← "response surface" or "surrogate" model)



$\eta(\underline{\mathbf{x}})$ **fitted to** $\{\underline{\mathbf{x}}^{(i)}, g(\underline{\mathbf{x}}^{(i)})\}$



⇒ $\eta(\mathbf{x})$ usually constructed in terms of basic functions that can be computed more easily (polynomials, exp, ⋯)

⇒ Should fit $g(\mathbf{x}^{(i)})$ sufficiently well especially in the region that contributes most to P_f or $E[g(\mathbf{x})^m]$

Instructor: Junho Song
junhosong@snu.ac.kr

③ History

- Box and Wilson (1954): influential
- Applied mostly in chemical, industrial eng. etc.

  (Mostly for "experimental design")
- Rackwitz (1982) $\Rightarrow$ Use RS for Structural Reliability Analysis
- Has been applied to random field, nonlinear structural dynamics, etc.

Instructor: Junho Song
junhosong@snu.ac.kr

457.646 Topics in Structural Reliability

In-Class Material: Class 26

VIII. Response Surface Method (Contd.)

◎ Basic formulation of RS models

Two approaches: Regression $\Rightarrow$ use assumed mathematical model & fit it to data

e.g. $\eta(\mathbf{x}) = \sum_{i=1}^{p} \theta_i x_i^m$

Interpolation $\Rightarrow$ Interpolate using nearby data points

e.g. K-nearest points

Regression

True response of $g(\mathbf{x})$: $Z(\mathbf{x})$

$$Z(\mathbf{x}) = \eta(\underbrace{\theta_1,\cdots,\theta_p}_{\text{Model parameters}};\ \underbrace{\mathbf{x}}_{\text{Input}}) + \underbrace{\varepsilon}_{\text{Zero mean (random) error term}}$$

Zero mean (random) error term

$\Rightarrow E[z-\eta] = E[\varepsilon] = 0$

"unbiased" model

How to find $\boldsymbol{\theta}$? What do data tell us?

Ref: Tipping, M.E. (2004)

"Bayesian inference: an introduction to principles and practice in machine learning" Advanced lectures on machine learning, pp.41-62

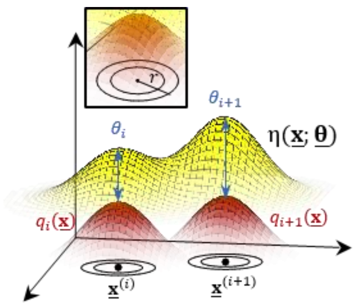
(Free codes and papers at www.miketipping.com)

$$\eta = \theta_1 \exp(x) + \theta_2 \ln x + \theta_3 \cdots$$

◎ Additive models (Linear in)

Find $Z = \eta(\mathbf{x};\boldsymbol{\theta}) + \varepsilon$

$$= \sum_{i=1}^{p} \underbrace{\theta_i}_{\text{Model Parameter}} \underbrace{q_i(\mathbf{x})}_{\text{Basis Function (Shape function)}} + \varepsilon$$



e.g. $q_i(\mathbf{x}) \propto \text{PDF of } N(\mathbf{x}^{(i)}, r^2\mathbf{I})$

from $\{\mathbf{x}^{(i)}, Z^{(i)}\},\ i = 1,\cdots,m$

Instructor: Junho Song
junhosong@snu.ac.kr

$\mathbf{Z} = \mathbf{Q}\boldsymbol{\theta} + \boldsymbol{\varepsilon}$

!!!

$$\begin{Bmatrix} Z^{(1)} \\ Z^{(2)} \\ \vdots \\ Z^{(m)} \end{Bmatrix} = \begin{bmatrix} q_1(\mathbf{x}^{(1)}) & \cdots & \cdots & q_p(\mathbf{x}^{(1)}) \\ \vdots & & & \\ q_1(\mathbf{x}^{(m)}) & \cdots & \cdots & q_p(\mathbf{x}^{(m)}) \end{bmatrix} \begin{Bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{Bmatrix} + \begin{Bmatrix} \varepsilon^{(1)} \\ \vdots \\ \varepsilon^{(m)} \end{Bmatrix}$$

$m \times 1$ $\quad$ $m \times p$ $\quad$ $p \times 1$ $\quad$ $m \times 1$

Five approaches (Tipping 2004)

① "Least-Square" Approximation (classic)

⇒ Minimize sum of squared errors

$$E_D = \frac{1}{2}\sum_{i=1}^{m}(Z^{(i)} - \eta(\mathbf{x}^{(i)}, \theta))^2$$
$$= \frac{1}{2}(\mathbf{Z} - \mathbf{Q}\boldsymbol{\theta})^T(\mathbf{Z} - \mathbf{Q}\boldsymbol{\theta})$$
$$= \frac{1}{2}\mathbf{Z}\mathbf{Z}^T + \frac{1}{2}(\mathbf{Q}\boldsymbol{\theta})^T(\mathbf{Q}\boldsymbol{\theta}) - \mathbf{Z}^T\mathbf{Q}\boldsymbol{\theta}$$

$$\frac{\partial E_D(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -\mathbf{Z}^T\mathbf{Q} + (\mathbf{Q}\boldsymbol{\theta})^T\mathbf{Q} = 0$$

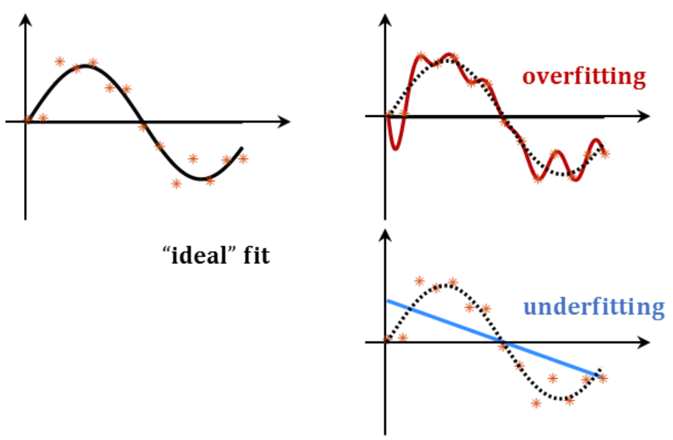
Solve for $\boldsymbol{\theta}$,

$$\boxed{\boldsymbol{\theta}_{LS} = (\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{Z}}$$

※ over-fitting?

e.g. $Z = \sin x + \varepsilon$

$\sin x \rightarrow$ true model, $\varepsilon \rightarrow$ noise $\qquad$ Figure 1 in Tipping (2004)



② Regularization (by giving penalty on large $\boldsymbol{\theta}$)

$$\hat{E}(\boldsymbol{\theta}) = E_D(\boldsymbol{\theta}) + \lambda \ E_W(\boldsymbol{\theta})$$

Standard choice

$$E_W(\boldsymbol{\theta}) = \frac{1}{2}\sum_{i=1}^{p}\theta_i^2$$

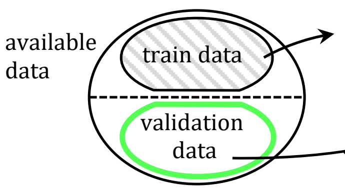
regularization parameter $\lambda \uparrow$ Discourage large value of $\boldsymbol{\theta}$

$\Rightarrow$ Smooth function

$$\frac{\partial E_D(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0 \ \Rightarrow \boxed{\boldsymbol{\theta}_{PLS} = (\mathbf{Q}^T\mathbf{Q} + \lambda\mathbf{I})^{-1}\mathbf{Q}^T\mathbf{Z}}$$

※ Appropriate value of λ?

A common approach: Use "validation" data



Find $\boldsymbol{\theta}_{PLS}$ for a given λ, $\boldsymbol{\theta}_{PLS}(\lambda)$

Construct surrogate $\eta(\mathbf{x};\lambda) = \sum_{i=1}^{p}\theta_i^{PLS}(\lambda)q_i(\mathbf{x})$

Compute error against validation data
$\varepsilon = Z - \eta$ at each $\mathbf{x}^{(i)} \in$ validation data set

and choose λ that makes $\frac{1}{2}\Sigma(Z-\eta)^2$ minimum

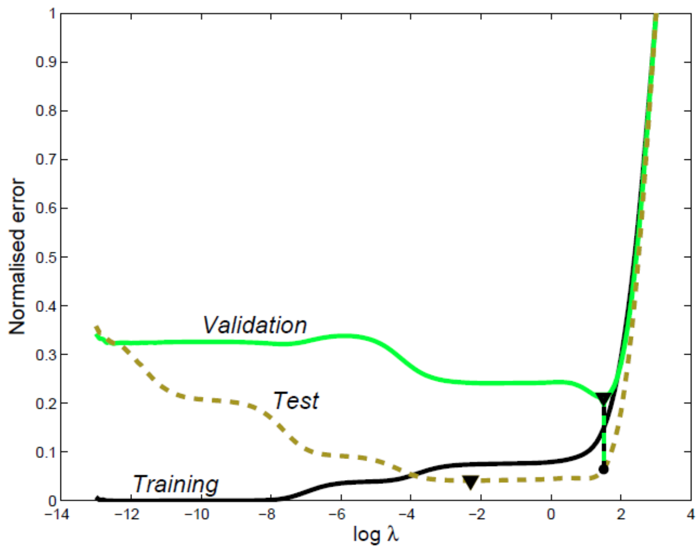


Fig. 3. Plots of error computed on the separate 15-example training and validation sets, along with 'test' error measured on a third noise-free set. The minimum test and validation errors are marked with a triangle, and the intersection of the best λ computed via validation is shown.

※ Probabilistic Regression

$$Z = \eta + \varepsilon \; !$$

e.g. $\varepsilon \sim N(0, \sigma^2)$ $\quad\quad\therefore Z \sim N(\eta, \sigma^2)$

Using this information one can construct likelihood function

$$L(\mathbf{Z} | \mathbf{x}, \boldsymbol{\theta}, \sigma^2) = \prod_{i=1}^{n} f(Z^{(i)} | \mathbf{x}^{(i)}, \boldsymbol{\theta}, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\{Z^{(i)} - \eta(\mathbf{x}^{(i)}; \boldsymbol{\theta})\}^2}{2\sigma^2} \right]$$

③ Maximum Likelihood Estimation

Find $\boldsymbol{\theta}$ that maximizes L( ) ⇔ Find $\boldsymbol{\theta}$ that minimizes –lnL( )

$$-\ln L(\quad) = \frac{n}{2}\ln(2\pi\sigma^2) + \frac{1}{2\sigma^2} \boxed{\sum_{i=1}^{n} \{Z^{(i)} - \eta(\mathbf{x}^{(i)}, \boldsymbol{\theta})\}^2}$$

$\mathrm{E}_D(\boldsymbol{\theta})$
⇒ error measure
for $\underline{\boldsymbol{\theta}}_{LS}$

Therefore, MLE based on s.i. __________ error assumption (i.e. $\varepsilon \sim N(\quad)$)

Gives $\quad \boxed{\boldsymbol{\theta}_{MLE} = \boldsymbol{\theta}_{LS}}$

(cf. Assuming errors are dependent? $\varepsilon \sim N(\mathbf{0}, \boldsymbol{\Sigma})$)

$$\rho_{ij} = \exp\left(-\frac{\left\| \mathbf{x}^{(i)} - \mathbf{x}^{(j)} \right\|}{L} \right) \Rightarrow$$ "Kriging" Method (Satner et al. 2003)

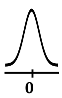
※ Bayesian Methods $\; f = c \cdot L \cdot p$

Introduce a prior distribution

$$p(\boldsymbol{\theta} | \alpha) = \prod_{i=1}^{p} \left(\frac{\alpha}{2\pi} \right)^{1/2} \exp\left\{ -\frac{\alpha}{2} \theta_i^2 \right\}$$

$$= \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi}\, 1/\sqrt{\alpha}} \exp\left\{ -\frac{\theta_i^2}{2(1/\alpha)} \right\}$$

(degree of belief about smooth model)

$\alpha \uparrow$ Variability reduces  ⇒ certain that θ is around 0

⇒Become smooth

$\therefore \; \alpha \propto \lambda$

Instructor: Junho Song
junhosong@snu.ac.kr

④ Maximum a posteriori (MAP) estimation (a Bayesian "shortcut")

$$f = c \cdot L \cdot p$$

$$\underbrace{P(\boldsymbol{\theta}|\mathbf{Z},\alpha,\sigma^2)}_{\text{Posterior}} = c \cdot \underbrace{L(\mathbf{Z}|\boldsymbol{\theta},\sigma^2)}_{\text{Likelihood function}} \cdot \underbrace{p(\boldsymbol{\theta}|\alpha)}_{\text{prior}}$$

Find $\boldsymbol{\theta}$ where $P(\boldsymbol{\theta}|\mathbf{Z},\alpha,\sigma^2)$ is maximum

e.g. Normal s.i errors ε, $Z \sim N(\eta,\sigma^2)$

$$-\ln(f) = \frac{1}{2\sigma^2}\sum_{i=1}^{n}\{Z^{(i)} - \eta(\mathbf{x}^{(i)};\boldsymbol{\theta})\}^2 + \frac{\alpha}{2}\sum_{i=1}^{p}\theta_i^2$$

$$-\sigma^2\ln(f) = \underbrace{\frac{1}{2}\sum_{i=1}^{n}\{Z^{(i)} - \eta(\mathbf{x}^{(i)};\boldsymbol{\theta})\}^2}_{E_D(\underline{\boldsymbol{\theta}})} + \underbrace{\frac{\alpha\sigma^2}{2}}_{=\frac{1}{2}\lambda}\underbrace{\sum_{i=1}^{p}\theta_i^2}_{E_W(\underline{\boldsymbol{\theta}})}$$

$\lambda = \alpha\ \sigma^2$

the same as

※ α, σ^2 ? no need to bother w/ Bayesian?

⑤ **Full** Bayesian ("Marginalization") integrate $P(\mathbf{Z}|\boldsymbol{\theta},\alpha,\sigma^2)$ over all $\boldsymbol{\theta}$

$$P(\mathbf{Z}) = \int P(\mathbf{Z}|\boldsymbol{\theta}) \cdot P(\boldsymbol{\theta})d\boldsymbol{\theta}$$

Focus on

Total probability theorem

$$P(\mathbf{Z}|\boxed{\alpha,\sigma^2}) = \int P(\mathbf{Z}|\boldsymbol{\theta},\boxed{\alpha,\sigma^2}) \cdot P(\boldsymbol{\theta}|\boxed{\alpha,\sigma^2})d\boldsymbol{\theta}$$

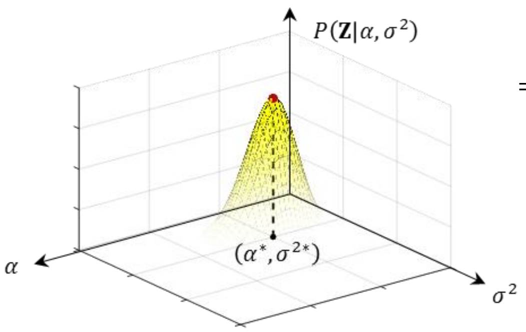
Simplified to

$$= \int P(\mathbf{Z}|\boldsymbol{\theta},\sigma^2) \cdot P(\boldsymbol{\theta}|\alpha)d\boldsymbol{\theta}$$

→ Closed-form available:

$f_N(\mathbf{Z},\alpha,\sigma^2)$ (Eq. 23 in Tipping, 2004)

※ $P(\mathbf{Z}|\alpha,\sigma^2)$: Probability that you will observe $\mathbf{Z}$ for given α,σ^2



$\Rightarrow$ Find $\alpha\,\&\,\sigma^2$ that maximizes $P(\mathbf{Z}|\alpha,\sigma^2)$

(i.e. Let data $\mathbf{Z}$ tell us the optimal point α^*,σ^{2*})

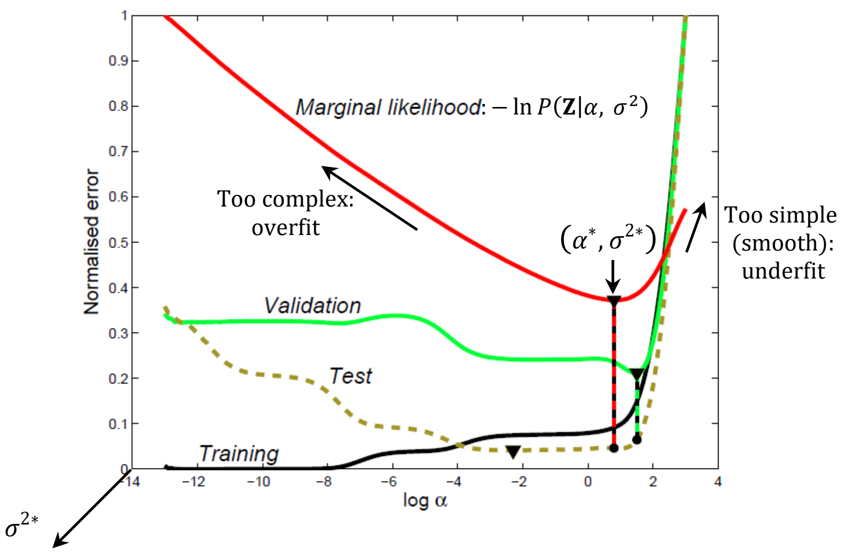


Fig. 5. Plots of the training, validation and test errors of the model as shown in Figure 3 (with the horizontal scale adjusted appropriately to convert from λ to α) along with the negative log marginal likelihood evaluated *on the training data alone* for that same model. The values of α and test error achieved by the model with highest marginal likelihood (smallest negative log) are indicated.

☆ **Okham's Razar** (or the law of parsimony)**:**

"model should be no more complex than is sufficient to explain the data"

CRC CH.19 RS
→DOE
→$q_i(\mathbf{x})$

◎ Other RS or UQ methods

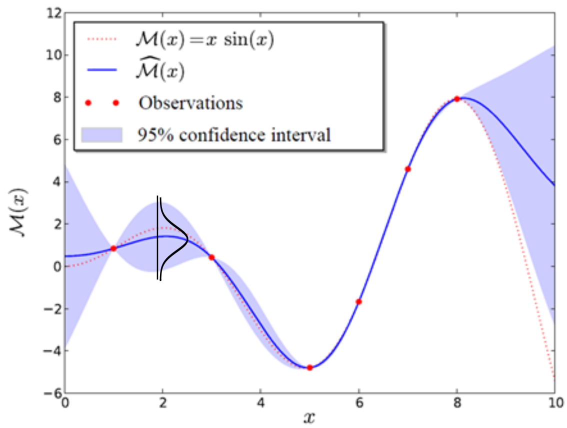
① Kriging (Santner et al. 2003)

(Dubourg et al. 2010 IFIP)

$\varepsilon \sim N(\mathbf{0}, \mathbf{\Sigma})$

e.g. $\rho_{ij} = \exp\left(-\frac{\left\|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}\right\|}{L}\right)$

- coincides at each point
- Interpolate b/w each point
- Can quantify confidence
- Regularization



(Dubourg et al. 2011)

Instructor: Junho Song
junhosong@snu.ac.kr

② Dimension Reduction (Rahman & Xu, 2004; Xu & Rahman 2004)

$$g(\mathbf{x}) \rightarrow g(\hat{\mathbf{x}}) = \sum_{i=1}^{n} g(\mu_1, \cdots, \mu_{i-1}, x_i, \mu_{i+1}, \cdots, \mu_n) - (n-1) g(\mu_1, \cdots, \mu_n)$$

$\Downarrow$

$$E[(g(x))^m] \cong E[(\hat{g}(x))^m]$$

$$= \int (\hat{g}(x))^m \underbrace{f_{\mathbf{x}}(\mathbf{x})}_{\Pi \varphi(x_i)} d\mathbf{x}$$

Transform to s.i. space; Multivariate Integral $\Rightarrow$ Multiple univariate Integral

③ Polynomials chaos (a good review by Eldred et al. 2008)

$$R = a_0 B_0 + \sum_{i_1=1}^{\infty} a_{i1} B_1(\zeta_{i1})$$

$$+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} a_{i1,i2} B_2(\zeta_{i1}\zeta_{i2}) + \cdots$$

$$= \sum_{j=0}^{p} \alpha_j \psi_j(\zeta)$$ → Orthogonal bases for given types of r.v's distribution

$$\alpha_j = \frac{< R, \psi_j >}{< \psi_j^2 >} = \frac{\int R \psi_j f(\zeta) d\zeta}{< \psi_j^2 >}$$

→ Important sampling, etc.

→ closed form available