Modélisation par Processus Gaussiens

➔ Notations

- Computer code $f : \mathbb{R}^D \longrightarrow \mathbb{R}$
- Inputs $\mathbf{x} = (x^1, \ldots, x^D) \in \mathbb{R}^D$
- Output *y(x)*
- Observations $(\mathbf{x}_i, y_i)_{i=1,\ldots,n}$
  - => learning sample $X_s = \left[\mathbf{x}_1^T, \ldots, \mathbf{x}_n^T\right]^T$ $Y_s = [y_1, \ldots, y_n]^T$

➔ Model: Output seen as realization of stationary Gaussian process

$$Y(\mathbf{x}) = f_0(\mathbf{x}) + W(\mathbf{x})$$

with :

- f_o the mean function or trend $f_0(\mathbf{x}) = \sum_{j=1}^{J} \beta_j f_j(\mathbf{x}) = F(x)\beta$

- ***W(x)*** a stationary centred Gaussian process (*E[W(x)] = 0*) with variance σ^2 and correlation function ***R*** :

$$\mathrm{Cov}(W(x), W(x')) = c(x, x') = \sigma^2 R(x - x')$$

➔ Joint distribution for the sample locations X_s and a new location x^*

$$[Y(X_S),Y(x^*)] \sim N\left(\begin{pmatrix} F_S \\ f_0(x^*) \end{pmatrix}, \begin{pmatrix} \Sigma_S & k(x^*) \\ k(x^*) & \sigma^2 \end{pmatrix}\right)$$

with $F_s = [f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)]^T$ the vector of the mean function at sample locations

Σ_S the covariance matrix at sample locations Xs

k(x*) the covariance vector between x and sample locations Xs

➔ Conditional distribution

$$Y(x^*)_{|Y(X_S)=Y_S} \sim N(\mu(x^*), \tilde{\sigma}^2(x^*))$$

$$avec \begin{cases} \mu(x^*) = E[Y(x^*)|Y(X_S)=Y_S] = f_0(x^*) + k(x^*)^T \Sigma_S^{-1}(Y_S - F_S) \\ \tilde{\sigma}^2(x^*) = Var[Y(x^*)|Y(X_S)=Y_S] = \sigma^2 - k(x^*)^T \Sigma_S^{-1} k(x^*) \end{cases}$$

The conditional mean $\mu(x^*)$ serves as the predictor at location x^*

The conditional variance $\tilde{\sigma}^2(x^*)$ serves as the prediction variance

➔ Maximum likelihood estimators for the hyperparameters

- Correlation parameters, called hyperparameters, ψ and R denoted as R_ψ

- Provided that ψ is known, regression parameters obtained by generalized least square estimator :

$$\hat{\beta} = (F_s R_\psi^{-1} F_s)^{-1} F_s^T R_\psi^{-1} Y_s$$

- MLE estimator of σ^2 is deduced

$$\widehat{\sigma^2} = \frac{1}{n}(Y_s - F_s\hat{\beta})^T R_\psi^{-1}(Y_s - F_s\hat{\beta})$$

- Estimation of hyperparameters consists in solving the minimization problem :

$$\psi^* = \arg\min_\psi \widehat{\sigma^2} \det(R_\psi)^{\frac{1}{n}}$$

Surrogate model validation

- **Validation of metamodel accurracy**
  - **Study of residuals computed:**
    - on a test sample
    - **Or by cross validation**
  - **Predictivity coefficient *Q²* :**
    - ***Q²*** estimated by cross validation on practical cases

$$Q^2 = 1 - \frac{\sum_{i=1}^{n} \left[\hat{Y}\left(x^{(i)}\right) - Y^{(i)} \right]^2}{\sum_{i=1}^{n} \left[Y^{(i)} - \frac{1}{n} \sum_{i=1}^{n} Y^{(i)} \right]^2}$$

$Y^{(i)} = y(x^{(i)})$ output on *observed data*

$\hat{Y}^{(i)} = E[Y(x^{(i)}) | Y(X_{S,-i}) = Y_{S,-i}]$ *GP* metamodel prediction *by cross validation*

Closer to one the Q², better the accuracy.