Spectral Surrogates

- Polynomials as Surrogates provide some analytical advantages over purely data-driven approaches.
- Here, we employ polynomials that exhibit orthogonality w.r.t. a specific PDF, $\rho(\vec{\theta})$.
  - Spectral expansions are often called "polynomial chaos expansion" (PCEs)
- Throughout, assume $\vec{\theta} \sim N(0, \underset{\sim}{I}_p)$ or $\vec{\theta} \sim U([-1,1]^p)$, $\vec{\theta} \in \mathbb{R}^p$
  - we assume each θ_i is independent.
- The general spectral expansion is

$$f_{LF}(\vec{\theta}) = \sum_{\ell=0}^{K} a_\ell \Psi_\ell(\vec{\theta})$$

Gaussian θ

- Assume $\theta \in \mathbb{R}$ is Gaussian, $\theta \sim N(0,1)$.
  - The pdf, $\rho(\theta)$, is

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\theta^2\right)$$

since $E[\theta] = 0$.

- Consider the Hermitian polynomials, $H_n(x)$, given

$$H_n(x) = (-1)^n \exp(x^2/2) \cdot \frac{d^n}{dx^n} \exp(-x^2/2)$$

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x$$

- Note, for any $m, n \in \mathbb{N}_0$

$$\int_{\mathbb{R}} H_n(\theta) H_m(\theta) \rho(\theta) d\theta$$

$$= \int_{\mathbb{R}} (-1)^{n+m} e^{\theta^2} \frac{d^n}{d\theta^n}\left(e^{-\theta^2/2}\right) \frac{d^m}{d\theta^m}\left(e^{-\theta^2/2}\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\theta^2} d\theta$$

$$= \begin{cases} 0, & \text{if } n \neq m \\ n!, & \text{if } n = m \end{cases}$$

Uniform θ

- Let $\theta \sim U([-1,1])$.
  - The Legendre polynomials are orthogonal to any pdf that is a constant.
  - The first five polynomials

$$P_0(\theta) = 1, \quad P_1(\theta) = \theta, \quad P_2(\theta) = \frac{1}{2}(3\theta^2 - 1)$$

$$P_3(\theta) = \frac{1}{2}(5\theta^3 - 3\theta), \quad P_4(\theta) = \frac{1}{8}(35\theta^4 - 30\theta^2 + 3)$$

- If $\theta \sim U([-1,1]) \Rightarrow \rho(\theta) = \frac{1}{1-(-1)} = \frac{1}{2}$

$$\Rightarrow \int_{-1}^{1} P_n(\theta) P_m(\theta) \rho(\theta) d\theta = \begin{cases} 0, & \text{if } m \neq n \\ \frac{1}{2n+1}, & \text{if } m = n \end{cases}$$

Multivariate Spectral Polynomials

- Consider $\vec{\theta} \in \mathbb{R}^P$ + assume each θ_i is independent.

- Consider $\vec{\Theta} \in \mathbb{R}^p$ + assume each Θ_i is independent.
  - Let $\Psi_q(\Theta_i)$ denote the q-th order univariate polynomial for Θ_i.

- The p-variate basis functions are given by

$$\Psi_{\vec{q}'} = \{\Psi_{q_1}(\Theta_1), \Psi_{q_2}(\Theta_2), \dots, \Psi_{q_p}(\Theta_p)\},$$

where $\vec{q}' = \{q_1, q_2, \dots, q_p\}$

- The multivariate basis functions must satisfy

$$E[\Psi_{\vec{i}'}(\vec{\Theta}), \Psi_{\vec{q}'}(\vec{\theta})] = \int_{\Gamma^*} \Psi_{\vec{i}'}(\vec{\theta}) \Psi_{\vec{q}'}(\vec{\theta}) \cdot \prod_{i=1}^{p} \rho(\theta_i) \, d\vec{\theta}$$

$$= \langle \Psi_{\vec{i}'}(\vec{\theta}), \Psi_{\vec{q}'}(\vec{\theta}) \rangle_{\rho^*}, \quad \rho^* = \prod_{i=1}^{p} \rho(\theta_i)$$

$$= \vec{\delta}_{\vec{i}', \vec{q}'} \cdot \vec{\gamma}_{\vec{q}', \vec{q}'}, \quad \vec{\gamma}_{\vec{q}', \vec{q}'} = E[\Psi^2_{\vec{q}'}(\vec{\theta})]$$

- The low fidelity surrogate is given by

$$f^K_{LF}(\vec{\theta}) = \sum_{|\vec{q}'|=0}^{K} a_{\vec{q}} \cdot \Psi_{\vec{q}'}(\vec{\theta})$$

  - Note: if we want up to q-order polynomials, then the total number of coefficients + basis functions is

- Note: if we want
number of coefficients + basis functions is

$$\frac{(q+p)!}{q!\,p!} = \text{nchoosek}(p, q)$$

Mean + Variance from Spectral Polynomials

- Let $f_{LF}(\vec{\theta})$ be described by a spectral polynomial

$$E\left[f_{LF}(\vec{\theta})\right] = E\left[\sum_{|\vec{q}'|=0}^{K} a_{\vec{q}'}\,\psi_{\vec{q}'}(\vec{\theta})\right]$$

$$= \sum_{|\vec{q}'|=0}^{K} a_{\vec{q}'}\, E\left[\psi_{\vec{q}'}(\vec{\theta})\right]$$

$$= a_0 + \sum_{|\vec{q}'|=1}^{K} a_{\vec{q}'}\, E\left[\psi_{\vec{q}'}(\vec{\theta})\right]$$

$$= a_0 + 0$$

$$Var\left[f_{LF}(\theta)\right] = E\left[\left(\sum_{|\vec{q}'|=0}^{K} a_{\vec{q}'}\,\psi_{\vec{q}'}(\vec{\theta}) - a_0\right)^2\right]$$

$$= E\left[\left(\sum_{|\vec{q}'|=1}^{K} a_{\vec{q}'}\,\psi_{\vec{q}'}(\vec{\theta})\right)^2\right]$$

$$= \sum_{|\vec{q}'|=1}^{K} a_{q'}^2\, E\left[\psi_{\vec{q}'}^2(\vec{\theta})\right]$$

$$= \sum_{|\vec{q}'|=1}^{K} a_{\vec{q}'}^2 \gamma_{\vec{q}'}, \qquad \text{Since } E\left[\psi_{\vec{q}'}^2(\vec{\theta})\right] = \gamma_{\vec{q}}$$

Computation of the Coefficients

i) Projection

- Note that for $f_{LF} = \sum_{|\vec{q}'|=0}^{K} a_{\vec{q}'} \psi_{\vec{q}'}(\vec{\theta})$

we can exploit orthogonality via

$$a_{\vec{q}'} = \frac{1}{\gamma_{\vec{q}'}} \int_{\Pi} f_{HF}(\vec{\theta}) \psi_{\vec{q}'}(\vec{\theta}) \rho(\vec{\theta}) \, d\vec{\theta}$$

- Using quadrature weights w^r + points $\vec{\theta}^r$

$$\Rightarrow \quad a_{\vec{q}'} \approx \frac{1}{\gamma_{\vec{q}'}} \sum_{r=1}^{R} f_{HF}(\vec{\theta}^r) \psi_{\vec{q}'}(\vec{\theta}^r) \rho(\vec{\theta}^r) w^r$$

ii) Regression

- To find $\vec{a}$ by regression, first generate $y_i = f_{HF}(\vec{\theta}^i)$, $i=1, \ldots, N_y$
- Then construct vectors + matrices

$$\vec{y} = \begin{bmatrix} f_{HF}(\vec{\theta}^{1}) \\ \vdots \\ f_{HF}(\vec{\theta}^{N_y}) \end{bmatrix}, \quad \underset{\sim}{X}_{ij} = \psi_{\vec{j}}(\vec{\theta}^{i})$$

$$\Rightarrow \quad \vec{y} = \underset{\sim}{X}\,\vec{a} \quad \Rightarrow \quad \vec{a} = (\underset{\sim}{X}^{T}\underset{\sim}{X})^{-1}\underset{\sim}{X}^{T}\,\vec{y}$$