Lecture 14+: Functional ANOVA

Hao Helen Zhang

Function ANOVA Decomposition

Similar to classical ANOVA, any multivariate function f is decomposed as

$$f(\boldsymbol{x}) = \beta_0 + \sum_{j=1}^{p} f_j(x_j) + \sum_{j<k} f_{jk}(x_j, x_k) + \cdots + f_{1\cdots p}(x_1, \cdots, x_p)$$

- Side conditions guarantee uniqueness of decomposition (Wahba 1990, Gu 2002)

Projection on Linear Space

Definition: Assume $\mathcal{F}$ is a linear space. A is a *projection* if it is a linear map: $\mathcal{F} \to \mathcal{F}$ such that

$$A^2 = A.$$

Any projection is associated with a direct sum decomposition.

Fact 1: Let $\mathcal{F}$ be a linear space, A is a projection, then

$$\mathcal{F} = \text{range}(A) \oplus \text{kernel}(A),$$

- $\text{range}(A) = \{f : f \in \mathcal{F}, f = Ag \text{ for some } g \in \mathcal{F}\}$ is the range of A
- $\text{kernel}(A) = \{f : Af = 0, f \in \mathcal{F}\}$ is the kernel of A
- $\oplus$ is the direct sum, which means $\text{range}(A) \Cap \text{kernel}(A) = \{0\}$

Fact 2: $f = Af + (I - A)f$ is the unique decomposition.

Product Domain

Consider a multivariate function $f(\boldsymbol{x}) = f(x_1, \ldots, x_p) \in \mathcal{F}$

- Let $\mathcal{X}_j$ be the domain for x_j, i.e., $x_j \in \mathcal{X}_j$
- $\mathcal{X} = \prod_{j=1}^{p} \mathcal{X}_j$ is the product domain for $\mathbf{x} = (x_1, \ldots, x_p)$
- $\mathcal{F}$ is a vector (or linear) space

Examples:

- continuous: $\mathcal{X}_j = [0, 1]$, and $\mathcal{X} = [0, 1]^p$, and $\mathcal{F}$ is a space of continuous functions
- discrete: $\mathcal{X}_j = \{1, 2, \cdots, K\}$ for any $j = 1, \ldots, p$.

Averaging Operator: A_j

A_j is a linear map: $\mathcal{F} \to \mathcal{F}$ that averages out x_j from the active argument list

- A_j satisfies $A_j^2 = A_j$ (it is a projection)

Examples: $p = 2$.

- Example 1: $\mathcal{X}_1 = \mathcal{X}_2 = [0,1], \mathcal{X} = [0,1]^2$
  - $A_1 f = \int_0^1 f(x_1, x_2) dx_1$
  - $A_2 f = \int_0^1 f(x_1, x_2) dx_2$
- Example 2. $\mathcal{X}_1 = \{1, \ldots, K_1\}$ and $\mathcal{X}_2 = \{1, \ldots, K_2\}$
  - $A_1 f = \sum_{x_1=1}^{K_1} f(x_1, x_2)/K_1$
  - $A_2 f = \sum_{x_2=1}^{K_2} f(x_1, x_2)/K_2$

Additional Examples

Examples: $p = 2$.

- Example 3: $\mathcal{X}_1 = \mathcal{X}_2 = [0,1], \mathcal{X} = [0,1]^2$
  - $A_1 = f(0, x_2)$
  - $A_2 = f(x_1, 0)$.
- Example 4. $\mathcal{X}_1 = \{1, \ldots, K_1\}$ and $\mathcal{X}_2 = \{1, \ldots, K_2\}$
  - $A_1 = f(1, x_2)$
  - $A_2 = f(x_1, 1)$. for $j = 1, 2$

Multiway ANOVA Decomposition

Assume $f \in \mathcal{F}$ is a linear space, A_j's are averaging operators on $\mathcal{F}$. Then

$$
\begin{aligned}
f(\boldsymbol{x}) &= \{\prod_{j=1}^{p}(I - A_j + A_j)\}f = \sum_{\mathcal{S}}\{\prod_{j\in\mathcal{S}}(I - A_j)\prod_{j\notin\mathcal{S}} A_j\}f \\
&= \sum_{\mathcal{S}} f_{\mathcal{S}} \\
&= \beta_0 + \sum_{j=1}^{p} f_j(x_j) + \sum_{j<k} f_{jk}(x_j, x_k) + \cdots + f_{1\cdots p}(x_1, \cdots, x_p)
\end{aligned}
$$

where

- $\mathcal{S} \in \{1, \ldots, p\}$ enlists the active arguments in $f_{\mathcal{S}}$
- the summation is over all of the 2^p subsets of $\{1, \ldots, p\}$.

ANOVA Interpretation

- $\beta_0 = \prod_{j=1}^{p} A_j(f)$ is a constant (overall mean).
- $f_j = f_{\{j\}} = (I - A_j) \sum_{k \neq j} A_k(f)$ is the x_j main effect.
- $f_{jk} = f_{\{j,k\}} = (I - A_j)(I - A_k) \sum_{l \neq k,j} A_l(f)$ is x_j-x_k interaction

Side conditions:

$$A_j f_{\mathcal{S}} = 0, \quad \forall j \in \mathcal{S}.$$

Assume $p = 2$, then

$$
\begin{aligned}
\beta_0 &= A_1 A_2 f, \\
f_1 &= (I - A_1) A_2 f = A_2 f - A_1 A_2 f = A_2 f - \beta_0, \\
f_2 &= (I - A_2) A_1 f = A_1 f - A_1 A_2 f = A_1 f - \beta_0, \\
f_{12} &= (I - A_1)(I - A_2) f \\
&= f(x_1, x_2) - f_1(x_1) - f_2(x_2) + \beta_0.
\end{aligned}
$$

And this decomposition is unique,

$$f(x_1, x_2) = \beta_0 + f_1(x_1) + f_2(x_2) + f_{12}(x_1, x_2)$$

Continuous Domain Example: $p = 2$

Domain $\mathcal{X}_1 = \mathcal{X}_2 = [0,1]$. $A_j = \int_0^1 f(x_1, x_2) dx_j$ for $j = 1, \ldots, 2$

- $\beta_0 = A_1 A_2 f = \int_0^1 \int_0^1 f(x_1, x_2) dx_1 x_2$
- $f_1 = (I - A_1) A_2 f = \int_0^1 f(x_1, x_2) dx_2 - \beta_0$
- $f_2 = (I - A_2) A_1 f = \int_0^1 f(x_1, x_2) dx_1 - \beta_0$
- $f_{12} = (I - A_1)(I - A_2) f$
  $= f(x_1, x_2) - \int_0^1 f dx_1 - \int_0^1 f dx_2 + \int_0^1 \int_0^1 f(x_1, x_2) dx_1 x_2$

Continuous Domain Example: $p = 2$

Domain $\mathcal{X}_1 = \mathcal{X}_2 = [0, 1]$.

- Example: $A_1 = f(0, x_2)$ and $A_2 = f(x_1, 0)$.
  - $\beta_0 = A_1 A_2 f = f(0, 0)$
  - $f_1 = (I - A_1) A_2 f = f(x_1, 0) - f(0, 0)$
  - $f_2 = (I - A_2) A_1 f = f(0, x_2) - f(0, 0)$
  - $f_{12} = (I - A_1)(I - A_2) f = f(x_1, x_2) - f(x_1, 0) - f(0, x_2) + f(0, 0)$

Discrete Domain Example: $p = 2$

Domain $\mathcal{X}_1 = \{1, \ldots, K_1\}$ and $\mathcal{X}_2 = \{1, \ldots, K_2\}$.

- Example: $A_j = \sum_{x_j=1}^{K_j} f(x_1, x_2)/K_j$ for $j = 1, \ldots, 2$
  - $\beta_0 = A_1 A_2 f = f_{\cdot\cdot}$
  - $f_1 = (I - A_1) A_2 f = f_{x_1\cdot} - f_{\cdot\cdot}$
  - $f_2 = (I - A_2) A_1 f = f_{\cdot x_2} - f_{\cdot\cdot}$
  - $f_{12} = (I - A_1)(I - A_2) f = f(x_1, x_2) - f_{x_1\cdot} - f_{\cdot x_2} + f_{\cdot\cdot}$,

  $f_{\cdot\cdot}$ is the overall mean, $f_{x_1\cdot} = \sum_{x_2=1}^{K_2} f(x_1, x_2)/K_2$ is the marginal average over x_2, $f_{\cdot x_2} = \sum_{x_1=1}^{K_1} f(x_1, x_2)/K_1$ is the marginal average over x_1.
- Example: $A_1 = f(1, x_2)$ and $A_2 = f(x_1, 1)$.
  - $\beta_0 = A_1 A_2 f = f(1, 1)$
  - $f_1 = (I - A_1) A_2 f = f(x_1, 1) - f(1, 1)$
  - $f_2 = (I - A_2) A_1 f = f(1, x_2) - f(1, 1)$
  - $f_{12} = (I - A_1)(I - A_2) f = f(x_1, x_2) - f(x_1, 1) - f(1, x_2) + f(1, 1)$

Truncated Models

- Additive models

$$f(\boldsymbol{x}) = \beta_0 + \sum_{j=1}^{p} f_j(x_j),$$

Claim X_j as unimportant if the function $f_j = 0$

- Two-way interaction model

$$f(\boldsymbol{x}) = \beta_0 + \sum_{j=1}^{p} f_j(x_j) + \sum_{j<k} f_{jk}(x_j, x_k).$$

The interaction effect between X_j and X_k is unimportant if $f_{jk} = 0$.