



Modelos con Coeficientes Variando VCMs

Bladimir Valerio Morales Torrez¹
Profesor: Dr. Christian Caamaño²

¹⁻² Universidad del Bío Bío

2022

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación
- 4 Inferencia
- 5 Aplicación
- 6 Conclusiones

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación
- 4 Inferencia
- 5 Aplicación
- 6 Conclusiones

Modelos Aditivos Generalizados

Los modelos aditivos generalizados (modelos no paramétricos), propuestos por Hastie and Tibshirani (1986) dan a conocer una nueva clase de modelos de regresión, extendiendo y flexibilizando el modelo de regresión clásico, reemplazando la función lineal por una función aditiva suave y no paramétrica, la cual es estimada por el algoritmo de puntuación local (local scoring algorithm).

Modelos con Coeficientes Variando

Hastie and Tibshirani (1993), proponen otra generalización al modelo de regresión lineal clásico, denominados modelos de coeficientes variando (Varying-Coefficient Models VCMs), los cuales contienen regresores lineales pero permiten que sus coeficientes cambien suavemente con el valor de otras variables, que se denominan "modificadores de efecto".

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación
- 4 Inferencia
- 5 Aplicación
- 6 Conclusiones

Modelo Lineal Generalizado con Coeficientes Variando

Sea una v.a. Y cuya distribución depende de un parámetro η (predicador lineal) y se relaciona con la media $\mu = \mathbb{E}(Y)$ mediante la función enlace $\eta = g(\mu)$, el modelo lineal generalizado con coeficientes variando (VCGLMs) se puede representar como:

$$y_i = \eta_i + \varepsilon_i \quad (i = 1, \dots, n), \quad (1)$$

$$\eta_i = \beta_0 + x_i^{(1)} \beta_1(t_{1i}) + \dots + x_i^{(s)} \beta_s(t_{si})$$

Modelo Lineal Generalizado con Coeficientes Variando

- Las variables explicativas t_{k_i} cambian el comportamiento de los coeficientes de las covariables $x_i^{(k)}$ a través de las funciones $\beta_k(\cdot)$ para $k = 1, \dots, s$.
- La dependencia de β_k en t_{k_i} implica un tipo de interacción entre t_{k_i} y $x_i^{(k)}$.
- Las variables t_{k_i} pueden no ser tan diferentes a $x_i^{(k)}$, pero también puede ser variables como el tiempo.

Casos del modelo

El CVMs generaliza los siguientes tipos de modelos:

- Si $\beta_k(t_{k_i}) = \beta_k$ es una función constante, entonces ese término es lineal en $x_i^{(k)}$. Si todos los términos son lineales se reduce a un modelo lineal generalizado clásico.
- Si $x_i^{(k)} = c$ con c constante entonces el k – *esimo* término es $\beta_k(t_{k_i})$ una función no conocida. Si todos los términos tienen esta característica se tiene un modelo aditivo generalizado.
- Una función lineal $\beta_k(t_{k_i}) = \beta_k t_{k_i}$ produce una interacción de la forma $\beta_k x_i^{(k)} t_{k_i}$.

Casos del modelo

- En el término del modelo $x_i^{(k)} \beta_k(t_{k_i})$, cuando x_i es una variable binaria (0 – 1). Supongamos que hay un término $\beta_0(t_{0_i})$ en el modelo. Esto equivale a tener una curva separada correspondiente a cada uno de los dos valores de x_i .
- A menudo las t_{k_i} serán la misma variable, (ej. edad o tiempo), que podría modificar los efectos de $x_i^{(k)}$. Si los datos consisten en $y_i, x_i^{(1)}, \dots, x_i^{(s)}$ sobre n puntos de tiempo $t \in (t_1, \dots, t_n)$. Entonces se podría modelar esto a lo largo del tiempo

$$\eta_t = \beta_0(t_i) + x_i^{(1)}(t_{1_i})\beta_1(t_{1_i}) + \dots + x_i^{(s)}(t_{s_i})\beta_s(t_{s_i})$$

Llamado modelo lineal generalizado dinámico o condicionalmente paramétrico.

Casos del modelo

- La variable modificante t_{k_i} puede ser la misma que $x_i^{(k)}$, por simplicidad suponemos que es un modelo lineal normal con un solo término.

$$y_i = x_i \beta(x_i) + \varepsilon_i$$

Este es un modelo común de suavizamiento o regresión no paramétrica de y versus x .

- Cada t_{k_i} puede tener un valor escalar o vectorial.

Casos del modelo

- En todos los casos anteriores, hay muchas formas de modelar las funciones $\beta_k(t_{k_i})$. Por ejemplo, podríamos usar representaciones paramétricas flexibles tales como polinomios, series de Fourier o polinomios por partes.
- De manera más general se pueden utilizar funciones no paramétricas, mediante el uso de métodos kernel, penalización o formulación bayesiana estocástica.

Modelo Lineal con Coeficientes Variando

Si se toma el modelo Gaussiano, donde $g(\mu) = \mu$ y la variable aleatoria y_i tiene distribución normal con media η_i , el modelo (1) es de la forma:

$$y_i = \beta_0 + x_i^{(1)} \beta_1(t_{1i}) + \dots + x_i^{(s)} \beta_s(t_{si}) + \varepsilon_i \quad (2)$$

donde y_i es la variable de respuesta y ε_i es el error aleatorio con $\mathbb{E}(\varepsilon) = 0$ y $\text{Var}(\varepsilon) = \sigma^2$

Modelo Lineal con Coeficientes Variando

El modelo (1) puede ser representado matricialmente como:

$$y = \tilde{N}_1\beta_1 + \dots + \tilde{N}_s\beta_s + \varepsilon, \quad (3)$$

donde y es un vector aleatorio de respuestas observadas ($n \times 1$), $\tilde{N}_k = X^{(k)}N_k$, $X^{(k)} = \text{diag}_{1 \leq i \leq n}(x_i^{(k)})$ y N_k es la matriz de incidencia de ($n \times r$) con el (i, l) -ésimo elemento igual al indicador $I(t_{k_i} = t_{k_l}^0)$, con $t_{k_l}^0$ que denota los valores distintos y ordenados de la variable explicativa t_{k_i} y $\beta_k = (\psi_{k_1}, \dots, \psi_{k_r})^\top$ es un ($r_k \times 1$) vector de parámetros con $\psi_{k_l} = \beta_k(t_{k_l}^0)$, para $l = 1, \dots, r_k$.

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación**
- 4 Inferencia
- 5 Aplicación
- 6 Conclusiones

Algunas consideraciones

- Los t_k pueden tomar valores diferentes para cada observación, por esta razón, se impone restricciones a las funciones de los coeficientes variando, por ejemplo, a trozos con forma paramétrica conocida o funciones suaves pero no paramétricas.
- Una aproximación puede ser a través de bases paramétricas como funciones polinómicas o trigonométricas, que normalmente no proporcionan suficiente flexibilidad y adaptabilidad local
- Es preferible utilizar un conjunto de bases splines.

Algunas consideraciones

- Para la estimación se procede igual que con un modelo lineal, sólo que con varias variables definidas por los productos de cada $x_i^{(k)}$, y $\beta_k(t_{k_i})$. Así, las herramientas inferenciales estándar sirven para evaluar conjuntos de coeficientes, puntos de influencia, etc.
- Las características de las curvas ajustadas pueden ser diferentes con pequeños cambios en las posiciones de los nudos, sobre todo si sólo se tienen pocos.

Mínimos Cuadrados Penalizados

Para la estimación de β_k se propone minimizar la suma de cuadrados de los errores penalizados

$$\sum_{i=1}^n \left\{ y_i - \sum_{j=1}^s x_i^{(j)} \beta_j(t_{j_i}) \right\}^2 - \sum_{j=1}^s \lambda_j \int \beta_j''(t_{j_i})^2 dt_j \quad (4)$$

El primer término mide la bondad del ajuste y el segundo penaliza la rugosidad de cada β_k con un parámetro fijo λ_k .

Mínimos Cuadrados Penalizados

De manera matricial se define la SCE_p :

$$\left(y - \sum_{j=1}^s \tilde{N}_j \beta_j \right)^{\top} \left(y - \sum_{j=1}^s \tilde{N}_j \beta_j \right) - \sum_{j=1}^s \lambda_j \beta_j^{\top} K_j \beta_j \quad (5)$$

donde K es la matriz definida positiva de $(r \times r)$ que depende de las matrices tridiagonales Q y R

Verosimilitud Penalizada

Bajo el modelo de coeficientes variando de la ecuación (2) se tiene que el log. de verosimilitud de $y_i \sim (\sum_{j=1}^s x_i^{(j)} \beta_j(t_{j_i}); \sigma^2) = (\sum_{j=1}^s \tilde{N}_k \beta_k)$ es:

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \left\{ y_i - \sum_{j=1}^s x_i^{(j)} \beta_j(t_{j_i}) \right\}^2 \quad (6)$$

donde $\boldsymbol{\theta} = (\beta_1^\top, \dots, \beta_s^\top, \sigma^2)^\top$ es el vector de parámetros a estimar.

Verosimilitud Penalizada

Matricialmente se tiene:

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \left(\mathbf{y} - \sum_{j=1}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right)^\top \left(\mathbf{y} - \sum_{j=1}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right) \quad (7)$$

Luego se tiene la función de log. verosimilitud penalizada que esta definida por:

$$\ell_p(\boldsymbol{\theta}, \boldsymbol{\lambda}) = \ell(\boldsymbol{\theta}) - \sum_{j=1}^s \frac{\lambda_j}{2} \boldsymbol{\beta}_j^\top \mathbf{K}_j \boldsymbol{\beta}_j$$

Vector Score Penalizado

Considerando λ_k fijo la función de Score penalizada esta dada por:

$$U_p(\boldsymbol{\theta}) = \begin{pmatrix} U_p^{\beta_k}(\boldsymbol{\theta}) \\ U_p^{\sigma^2}(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} \frac{\partial \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \beta_k} \\ \frac{\partial \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \sigma^2} \end{pmatrix} \quad (8)$$

Vector Score Penalizado

$$\frac{\partial \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\beta}_k} = \frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \left(y - \tilde{\mathbf{N}}_k \boldsymbol{\beta}_k - \sum_{j=1: j \neq k}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right) - \lambda_k \mathbf{K}_k \boldsymbol{\beta}_k$$

$$\frac{\partial \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \left(y - \sum_{j=1}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right)^\top \left(y - \sum_{j=1}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right)$$

Matriz Hessiana Penalizada

Esta definida como las segundas derivadas del vector de Score.

$$\frac{\partial^2 \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\beta}_k \partial \boldsymbol{\beta}_k^\top} =$$

* Si $k = k'$

$$\frac{\partial^2 \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\beta}_k \partial \boldsymbol{\beta}_k^\top} = -\frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k - \lambda_k \mathbf{K}_k \quad (9)$$

* Si $k \neq k'$

$$\frac{\partial^2 \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\beta}_k \partial \boldsymbol{\beta}_k^\top} = -\frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k \quad (10)$$

Matriz Hessiana Penalizada

$$\frac{\partial^2 \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \sigma^2 \partial \boldsymbol{\beta}_k^\top} = -\frac{1}{\sigma^4} \tilde{\mathbf{N}}_k^\top \left(y - \tilde{\mathbf{N}}_k \boldsymbol{\beta}_k - \sum_{j=1; j \neq k}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right)$$

$$\frac{\partial^2 \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \sigma^2 \partial \sigma^2} = \frac{n}{2\sigma^4} - \frac{1}{2\sigma^6} \left(y - \sum_{j=1}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right)^\top \left(y - \sum_{j=1}^s \tilde{\mathbf{N}}_j \boldsymbol{\beta}_j \right)$$

Matriz de Información de Fisher Penalizada

Se define como

$$\mathcal{I}_p(\boldsymbol{\theta}) = -\mathbb{E} \left\{ \frac{\partial^2 \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right\}$$

donde:

$$\mathcal{I}_p(\boldsymbol{\theta}) = \begin{pmatrix} \mathcal{I}_p^{\boldsymbol{\beta}_1 \boldsymbol{\beta}_1^\top}(\boldsymbol{\theta}) & \dots & \mathcal{I}_p^{\boldsymbol{\beta}_1 \boldsymbol{\beta}_s^\top}(\boldsymbol{\theta}) & \mathcal{I}_p^{\boldsymbol{\beta}_1 \sigma^2}(\boldsymbol{\theta}) \\ \vdots & \ddots & \vdots & \vdots \\ \mathcal{I}_p^{\boldsymbol{\beta}_s \boldsymbol{\beta}_1^\top}(\boldsymbol{\theta}) & \dots & \mathcal{I}_p^{\boldsymbol{\beta}_s \boldsymbol{\beta}_s^\top}(\boldsymbol{\theta}) & \mathcal{I}_p^{\boldsymbol{\beta}_s \sigma^2}(\boldsymbol{\theta}) \\ \mathcal{I}_p^{\sigma^2 \boldsymbol{\beta}_1}(\boldsymbol{\theta}) & \dots & \mathcal{I}_p^{\sigma^2 \boldsymbol{\beta}_s}(\boldsymbol{\theta}) & \mathcal{I}_p^{\sigma^2 \sigma^2}(\boldsymbol{\theta}) \end{pmatrix}$$

Matriz de Información de Fisher Penalizada

* Para $k = k'$

$$\begin{aligned}\mathcal{I}_p^{\beta_k \beta_k^\top}(\theta) &= -\mathbb{E} \left\{ -\frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k - \lambda_k \mathbf{K}_k \right\} \\ &= \frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k + \lambda_k \mathbf{K}_k\end{aligned}\quad (11)$$

* Para $k \neq k'$

$$\begin{aligned}\mathcal{I}_p^{\beta_k \beta_{k'}^\top}(\theta) &= -\mathbb{E} \left\{ -\frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_{k'} \right\} \\ &= \frac{1}{\sigma^2} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_{k'}\end{aligned}\quad (12)$$

Matriz de Información de Fisher Penalizada

$$\mathcal{I}_p^{\beta_k \sigma^2}(\theta) = -\mathbb{E} \left\{ -\frac{1}{\sigma^4} \tilde{\mathbf{N}}_k^\top \left(y - \tilde{\mathbf{N}}_k \beta_k - \sum_{j=1; j \neq k}^s \tilde{\mathbf{N}}_j \beta_j \right) \right\}$$

$$= 0_{(r \times 1)}$$

$$\mathcal{I}_p^{\sigma^2 \sigma^2}(\theta) = -\mathbb{E} \left\{ \frac{n}{2\sigma^4} - \frac{1}{2\sigma^6} \left(y - \sum_{j=1}^s \tilde{\mathbf{N}}_j \beta_j \right)^\top \left(y - \sum_{j=1}^s \tilde{\mathbf{N}}_j \beta_j \right) \right\}$$

$$= \frac{n}{2\sigma^4}$$

Estimadores de Máxima Verosimilitud Penalizada

Si consideramos λ y σ^2 fijos, para hallar el estimador de máxima verosimilitud penalizada se debe igualar el vector score penalizado a cero.

$$U_p^{\beta_k}(\theta) = \frac{\partial \ell_p(\theta, \lambda)}{\partial \beta_k} = 0$$

$$\frac{1}{\sigma^2} \tilde{N}_k^\top \left(y - \tilde{N}_k \beta_k - \sum_{j=1: j \neq k}^s \tilde{N}_j \beta_j \right) - \lambda_k K_k \beta_k = 0$$

$$\hat{\beta}_k = \left(\tilde{N}_k^\top \tilde{N}_k + \lambda_k K_k \right)^{-1} \tilde{N}_k^\top \left(y - \sum_{j=1: j \neq k}^s \tilde{N}_j \beta_j \right) \quad (13)$$

Estimadores de Máxima Verosimilitud Penalizada

Para σ^2 , se tiene:

$$\begin{aligned} U_p^{\sigma^2}(\boldsymbol{\theta}) &= \frac{\partial \ell_p(\boldsymbol{\theta}, \lambda)}{\partial \sigma^2} = 0 \\ -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \left(y - \sum_{j=1}^s \tilde{N}_j \beta_j \right)^\top \left(y - \sum_{j=1}^s \tilde{N}_j \beta_j \right) &= 0 \\ \hat{\sigma}^2 &= \frac{1}{n} \left(y - \sum_{j=1}^s \tilde{N}_j \hat{\beta}_j \right)^\top \left(y - \sum_{j=1}^s \tilde{N}_j \hat{\beta}_j \right) \end{aligned} \quad (14)$$

Valores ajustados

Dado el vector $\hat{y} = (\hat{y}_1, \dots, \hat{y}_n)^\top$ se define los valores ajustados

$$\begin{aligned}\hat{y} &= \tilde{N}_k \hat{\beta}_n \\ &= \tilde{N}_k \left(\tilde{N}_k^\top \tilde{N}_k + \lambda_k K_k \right)^{-1} \tilde{N}_k^\top \left(y - \sum_{j=1:j \neq k}^s \tilde{N}_j \beta_j \right) \\ &= H(\lambda) \left(y - \sum_{j=1:j \neq k}^s \tilde{N}_j \beta_j \right)\end{aligned}$$

donde $H(\lambda) = \tilde{N}_k \left(\tilde{N}_k^\top \tilde{N}_k + \lambda_k K_k \right)^{-1} \tilde{N}_k^\top$ es la matriz de proyección de dimensión $(n \times n)$, que no es idempotente.

Grados de libertad efectivos

Indican el número efectivo de parámetros considerados en el modelo, definido por:

$$GLE(\lambda) = \text{traza}\{H(\lambda)\}$$

- * Si $\lambda \rightarrow 0$ se está priorizando el ajuste del modelo con la interpolación de las n observaciones, $GLE(\lambda) \rightarrow n$.
- * Si $\lambda \rightarrow \infty$ el modelo equivale al modelo lineal clásico, $GLE(\lambda) \rightarrow s$.
- * El dominio de GLE se definen en $[s, n]$

Grados de libertad efectivos

Para el VCM se tiene:

$$\begin{aligned} GLE(\lambda_k) &= \text{traza}\{H(\lambda_k)\} \\ &= \text{traza}\left\{\tilde{N}_k \left(\tilde{N}_k^\top \tilde{N}_k + \lambda_k K_k\right)^{-1} \tilde{N}_k^\top\right\} \end{aligned}$$

Sea $B_k = \tilde{N}_k^\top N_k$ y $L_k = B_k^{-1/2} K_k B_k^{-1/2}$

$$\begin{aligned} GLE(\lambda_k) &= \text{traza}\left\{B_k (B_k + \lambda_k K_k)^{-1}\right\} \\ &= \text{traza}\left\{(I_r + \lambda_k L_k)^{-1}\right\} = \sum_{j=1}^r \frac{1}{1 + \lambda_k L_{kj}} \end{aligned}$$

Criterio de Información de Akaike AIC

Es un criterio de bondad de ajuste para caclular un modelo agrupado, que consiste en minimizar la función:

$$AIC(\lambda) = -2\ell_p(\boldsymbol{\theta}, \lambda) + 2\{s + 1 + GLE(\lambda)\}$$

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación
- 4 Inferencia**
- 5 Aplicación
- 6 Conclusiones

Media de los estimadores

$$\begin{aligned}
 \mathbb{E}[\hat{\beta}_k] &= \mathbb{E} \left[\left(\tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k + \lambda_k \mathbf{K}_k \right)^{-1} \tilde{\mathbf{N}}_k^\top \left(y - \sum_{j=1; j \neq k}^s \tilde{\mathbf{N}}_j \beta_j \right) \right] \\
 &= \left(\tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k + \lambda_k \mathbf{K}_k \right)^{-1} \tilde{\mathbf{N}}_k^\top \left(\mathbb{E}[y] - \sum_{j=1; j \neq k}^s \tilde{\mathbf{N}}_j \beta_j \right) \\
 &= \left(\tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k + \lambda_k \mathbf{K}_k \right)^{-1} \tilde{\mathbf{N}}_k^\top \left(\sum_{j=1}^s \tilde{\mathbf{N}}_j \beta_j - \sum_{j=1; j \neq k}^s \tilde{\mathbf{N}}_j \beta_j \right) \\
 &= \left(\tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k + \lambda_k \mathbf{K}_k \right)^{-1} \tilde{\mathbf{N}}_k^\top \tilde{\mathbf{N}}_k \beta_k \tag{15}
 \end{aligned}$$

Media de los estimadores

- Sin embargo es posible escoger el λ_k para disminuir el sesgo de los estimadores.
- Si $\lambda_k \rightarrow 0$ se tendrá un estimador insesgado $\mathbb{E}(\hat{\beta}_k) \rightarrow \beta$

Matriz de Varianzas-Covarianzas

Se tiene que $Var[y] = \sigma^2 I$, así,

$$Var[\hat{\beta}_k] = Var \left[\left(\tilde{N}_k^\top \tilde{N}_k + \lambda_k K_k \right)^{-1} \tilde{N}_k^\top \left(y - \sum_{j=1: j \neq k}^s \tilde{N}_j \beta_j \right) \right]$$

Matriz de Varianzas-Covarianzas Asintótica

Esta dado por la inversa de la matriz de información de Fisher penalizada,

$$\hat{Cov}(\hat{\theta}) \approx \mathcal{I}_p^{-1}(\hat{\theta})$$

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación
- 4 Inferencia
- 5 Aplicación**
- 6 Conclusiones

Vamos a R

- Se utilizará los datos presentados en el libro de Harezlak et al. (2018) de los rendimientos diarios de las acciones de la General Electric Company y el índice de Standard & Poor's 500 (S&P 500). Los datos están en el 'data.frame' 'capm' en el paquete 'HRW'. La variable 'Date' se tiene desde el 1 de noviembre de 1993 al 31 de marzo de 2003, por lo que hay más de nueve años de datos.
- Mencionar también que existe ejemplos detallados de este modelo en Ruppert et al. (2003).

Contenido

- 1 Introducción
- 2 Modelo con Coeficientes Variando VCM
- 3 Estimación
- 4 Inferencia
- 5 Aplicación
- 6 Conclusiones**

Conclusiones

- En general los modelos de regresión con coeficientes variando es una alternativa de modelación de datos, cuando una covariable en interacción con otra variable puede ayudar a explicar de mejor forma el modelo, teniendo siempre en cuenta el principio de parsimonia.
- Es bastante flexible en cuanto al estudio de datos que se requiera realizar, se podría utilizar en datos de corte transversal y así también en datos longitudinales (panel).

Conclusiones

- En este trabajo se presento de manera inicial el modelo lineal generalizado de coeficientes variando, lo cual nos proporciona diferentes modelos con distribuciones pertenecientes a la familia exponencial.
- También se puede contar con modelos de coeficientes variando de distribuciones asimétricos (skew) y otros.

Referencias I

- Harezlak, J., Ruppert, D., and Want, M. (2018), *Semiparametric Regression with R* Springer Science+Business Media, LLC, part of Springer Nature.
- Hastie, T., and Tibshirani, R. (1986), "Generalized Additive Models," *Statistical Science*, 1(3), 297 – 310.
URL: <https://doi.org/10.1214/ss/1177013604>
- Hastie, T., and Tibshirani, R. (1993), "Varying-Coefficient Models," *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(4), 757 –96.
- Ruppert, D., Want, M., and Carroll, R. (2003), *Semiparametric Regression* Cambridge University Press.