

Detalles Metodológicos del Modelo de Corrección de Errores con MARS

José Mauricio Gómez Julián

Table of Contents

1	Marco Econométrico para el Análisis de Cointegración y Corrección de Errores	2
1.1	Determinación de Orden de Integración $I(1)$	2
1.1.1	Fundamento Teórico	2
1.1.2	Protocolo de Verificación	2
1.1.3	Justificación del Nivel de Significancia	3
1.2	Análisis de Cointegración	3
1.2.1	Enfoque Dual: Engle-Granger y Johansen	3
1.2.2	Justificación de la Regla “Either”	4
1.3	Modelo de Corrección de Errores (ECM)	4
1.3.1	Especificación del ECM Lineal	4
1.3.2	Selección Óptima de Rezagos	4
1.3.3	Test Unidireccional con Errores HAC	5
1.4	Extensión No-Lineal con MARS	5
1.4.1	Motivación Económica	5
1.4.2	Especificación del Modelo MARS-ECM	6
1.4.3	Configuración de Hiperparámetros	6
1.5	Validación Cruzada Temporal	6
1.5.1	Rolling-Origin con Ventana Deslizante	6
1.5.2	Validación Cruzada Anidada	7
1.6	Métricas de Evaluación	7
1.6.1	Métricas en Escala	7
1.6.2	Métricas Relativas	7
1.6.3	Métrica Explicativa	7
1.6.4	Descomposición de Theil	7
1.7	Criterios de Estabilidad Temporal	8

1.7.1	Métrica de Soporte	8
1.7.2	Umbrales de Validación	8
1.7.3	Métricas Ajustadas por Estabilidad	8
1.8	Implementación Computacional	8
1.8.1	Paralelización Multi-Nivel	8
1.8.2	Gestión de Progreso	8
1.8.3	Complejidad Computacional	9

1 Marco Econométrico para el Análisis de Cointegración y Corrección de Errores

El presente documento detalla la metodología implementada para evaluar relaciones de equilibrio de largo plazo y dinámicas de ajuste de corto plazo entre variables de producción y circulación mediante un enfoque híbrido que combina técnicas econométricas clásicas de cointegración con métodos modernos de aprendizaje estadístico no paramétrico. Esta metodología integra el rigor del análisis de cointegración con la flexibilidad de los Multivariate Adaptive Regression Splines (MARS) para capturar no-linealidades en el mecanismo de corrección de errores.

1.1 Determinación de Orden de Integración I(1)

1.1.1 Fundamento Teórico

El análisis de cointegración requiere que las series temporales involucradas sean integradas de orden uno, denotadas como $I(1)$. Una serie $I(1)$ es no estacionaria en niveles pero se vuelve estacionaria tras una diferenciación. Esta propiedad es fundamental porque solo procesos $I(1)$ pueden mantener relaciones de equilibrio de largo plazo sin divergir indefinidamente.

1.1.2 Protocolo de Verificación

Sea $z_t \in \{Y_t, X_t\}$ una serie temporal. Determinamos que $z_t \sim I(1)$ mediante el siguiente protocolo de dos etapas:

Etapas 1: No estacionariedad en niveles

Aplicamos el test Augmented Dickey-Fuller (ADF) con componentes determinísticos:

$$\Delta z_t = \mu + \beta t + \rho z_{t-1} + \sum_{i=1}^p \psi_i \Delta z_{t-i} + \varepsilon_t$$

donde testeamos $H_0: \rho = 0$ (raíz unitaria). Requerimos fallo en rechazar H_0 al 10% de significancia tanto con drift como con trend, indicando no estacionariedad robusta a la especificación de determinísticos.

Etap 2: Estacionariedad en primeras diferencias

Testeamos la primera diferencia sin componentes determinísticos:

$$\Delta^2 z_t = \rho' \Delta z_{t-1} + \sum_{i=1}^p \psi'_i \Delta^2 z_{t-i} + \eta_t$$

Requerimos rechazo de $H_0: \rho' = 0$ al 10%, confirmando estacionariedad tras diferenciación.

1.1.3 Justificación del Nivel de Significancia

Utilizamos 10% en las pruebas $I(1)$ como pre-filtro conservador para reducir el False Negative Rate (FNR)[^1]. Este umbral más permisivo en la etapa inicial se compensa con filtros más estrictos en etapas posteriores (cointegración al 5%, ECM unidireccional al 5%, validación out-of-sample).

1.2 Análisis de Cointegración

1.2.1 Enfoque Dual: Engle-Granger y Johansen

Implementamos dos metodologías complementarias de cointegración, reconociendo que cada una tiene fortalezas específicas:

1.2.1.1 Procedimiento Engle-Granger con Phillips-Ouliaris

Paso 1: Regresión de cointegración

Estimamos la relación de largo plazo:

$$Y_t = \alpha + \beta X_t + u_t$$

donde α y β son los parámetros del vector de cointegración $(1, -\alpha, -\beta)'$.

Paso 2: Test de estacionariedad en residuos

Aplicamos dos pruebas sobre \hat{u}_t :

- **ADF sin determinísticos:** Testea raíz unitaria en los residuos al nivel $p < 0.05$
- **Phillips-Ouliaris (PO):** Prueba robusta a endogeneidad y correlación serial

Paso 3: Regla de decisión “either”

Aceptamos cointegración si **cualquiera** de las pruebas (ADF o PO) valida al 5%. Esta regla reduce FNR en presencia de cambios estructurales, compensándose con validación posterior más estricta.

1.2.1.2 Test de Johansen con Traza

Especificamos un VAR de orden K (seleccionado por criterio BIC via VARselect) y lo transformamos a su representación VECM:

$$\Delta Z_t = \Pi Z_{t-1} + \sum_{i=1}^{K-1} \Gamma_i \Delta Z_{t-i} + \Psi D_t + \varepsilon_t$$

donde: - $Z_t = (Y_t, X_t)'$ es el vector de variables - $\Pi = \alpha\beta'$ es la matriz de impacto de largo plazo - D_t contiene los determinísticos (constante y/o tendencia)

Testeamos $H_0: r = 0$ (ningún vector de cointegración) mediante el estadístico de traza. Probamos con especificaciones {const,trend} y aceptamos cointegración si el estadístico excede el valor crítico al 5% para alguna especificación.

1.2.2 Justificación de la Regla “Either”

La regla “either” (EG o Johansen) en lugar de “both” (EG y Johansen) se justifica por:

1. **Robustez a cambios de régimen:** Diferentes pruebas pueden ser sensibles a distintos tipos de quiebres estructurales
2. **Compensación con filtros posteriores:** El ECM con $\lambda < 0$ y la validación out-of-sample filtran falsos positivos
3. **Evidencia empírica:** Tests de Monte Carlo muestran menor FNR sin inflación sustancial de FPR cuando se combina con validación predictiva

1.3 Modelo de Corrección de Errores (ECM)

1.3.1 Especificación del ECM Lineal

Dado el vector de cointegración (α, β) de Engle-Granger, construimos:

$$ECM1_t = Y_{t-1} - \alpha - \beta X_{t-1}$$

Este término captura el desvío del equilibrio de largo plazo en $t - 1$. El modelo ECM completo es:

$$\Delta Y_t = \lambda \cdot ECM1_t + \sum_{i=1}^L \phi_i \Delta Y_{t-i} + \sum_{i=1}^L \gamma_i \Delta X_{t-i} + \varepsilon_t$$

donde: - $\lambda < 0$ es la velocidad de ajuste hacia el equilibrio - L es el número de rezagos en diferencias - ϕ_i, γ_i capturan la dinámica de corto plazo

1.3.2 Selección Óptima de Rezagos

Implementamos un procedimiento de selección automática:

1. **Búsqueda sobre $L \in \{1, 2, \dots, L_{\max}\}$** con $L_{\max} = 4$ por defecto

2. **Criterio primario:** BIC para parsimonia
3. **Restricción de ruido blanco:** Si el modelo con menor BIC falla Ljung-Box ($p \leq 0.05$ con 12 rezagos), seleccionamos el modelo con menor BIC que pase la prueba
4. **Guía opcional:** $L \approx \max(1, K - 1)$ donde K es el orden VAR, aunque no es vinculante

1.3.3 Test Unidireccional con Errores HAC

1.3.3.1 Hipótesis y Justificación

Testeamos: - $H_0: \lambda \geq 0$ (no hay corrección o hay divergencia) - $H_1: \lambda < 0$ (existe corrección hacia el equilibrio)

La prueba unidireccional es fundamental porque $\lambda > 0$ implicaría divergencia del equilibrio, lo cual es económicamente incoherente y violaría la condición de estabilidad del sistema.

1.3.3.2 Inferencia Robusta

Empleamos errores estándar HAC (Heteroskedasticity and Autocorrelation Consistent) mediante el estimador Newey-West:

$$\hat{V}_{NW} = \hat{\Omega}_0 + \sum_{j=1}^m w_j (\hat{\Omega}_j + \hat{\Omega}'_j)$$

donde $w_j = 1 - j/(m + 1)$ son los pesos de Bartlett y m se selecciona automáticamente. El estadístico t robusto es:

$$t = \frac{\hat{\lambda}}{\sqrt{\hat{V}_{NW, \lambda \lambda}}}$$

con p-valor unidireccional $p = P(T \leq t | H_0)$. Rechazamos si $p < 0.05$ y $\hat{\lambda} < 0$.

1.4 Extensión No-Lineal con MARS

1.4.1 Motivación Económica

Las relaciones económicas frecuentemente exhiben no-linealidades: - **Efectos umbral:** Respuestas diferentes según el nivel de las variables - **Asimetrías:** Ajustes diferentes para desvíos positivos vs negativos - **Cambios de régimen:** Parámetros variables según el contexto económico

MARS captura estas características mediante funciones base adaptativas sin requerir especificación *a priori* de la forma funcional.

1.4.2 Especificación del Modelo MARS-ECM

El modelo no-lineal se especifica como:

$$\Delta Y_t = f(\text{ECM1}_t, \Delta X_t, \Delta Y_{t-1}, \Delta X_{t-1}, \Delta Y_{t-2}) + \eta_t$$

donde $f(\cdot)$ es aproximada por MARS como:

$$f(\mathbf{x}) = \beta_0 + \sum_{m=1}^M \beta_m \prod_{k=1}^{K_m} h_{km}(x_{v(k,m)})$$

con: - h_{km} son funciones bisagra: $\max(0, x - c)$ o $\max(0, c - x)$ - M es el número de funciones base (controlado por nk) - K_m es el grado de interacción (controlado por degree)

1.4.3 Configuración de Hiperparámetros

La grilla de búsqueda específica: - **degree** $\in \{1, 2\}$: Controla interacciones (1 = aditivo, 2 = permite interacciones) - **nk** $\in \{15, 25, 35, 50, 65\}$: Número máximo de términos antes de poda

Esta grilla balancea flexibilidad con riesgo de sobreajuste, siendo expandible con más datos históricos.

1.5 Validación Cruzada Temporal

1.5.1 Rolling-Origin con Ventana Deslizante

Implementamos validación cruzada temporal respetando la causalidad mediante rolling-origin con ventana deslizante (*sliding window*):

1.5.1.1 Parámetros de Configuración

- **Tamaño inicial:** $\max(40, 0.80 \times n)$ observaciones
- **Horizonte de test:** 12 meses (pronóstico anual)
- **Paso entre orígenes:** 12 meses (evita solapamiento)
- **Tipo de ventana:** Sliding (tamaño constante) vs Expanding (acumulativa)

1.5.1.2 Justificación de Sliding Window

La ventana deslizante mantiene “memoria comparable” entre folds y es más sensible a cambios de régimen que la expansiva. Esto es crucial para series económicas con parámetros no constantes en sentido amplio. Evidencia empírica muestra que sliding:

1. Preserva estabilidad global (mismo número de modelos robustos)
2. Aumenta sensibilidad local (detecta más relaciones régimen-dependientes)
3. Mejora adaptación a dinámicas recientes

1.5.2 Validación Cruzada Anidada

Para selección de hiperparámetros sin contaminar la evaluación:

1. **Nivel externo:** Rolling-origin para evaluación de desempeño
2. **Nivel interno:** Dentro de cada train externo, rolling-origin adicional con:
 - Inicial: 60% del train externo
 - Test interno: 6 meses
 - Paso interno: 3 meses

Esta estructura evita *data snooping*[^2] y proporciona estimaciones no sesgadas del error de generalización.

1.6 Métricas de Evaluación

1.6.1 Métricas en Escala

- **RMSE:** $\sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}$
- **MAE:** $\frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$

1.6.2 Métricas Relativas

- **MAPE:** $\frac{100}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$ (protegido para $|Y_t| > \epsilon$)
- **sMAPE:** $\frac{100}{n} \sum_{t=1}^n \frac{2|Y_t - \hat{Y}_t|}{|Y_t| + |\hat{Y}_t|}$
- **Theil's U:** $\frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n Y_t^2}}$

1.6.3 Métrica Explicativa

- R^2 protegido:

$$R^2 = \begin{cases} 1 - \frac{\sum (Y_t - \hat{Y}_t)^2}{\sum (Y_t - \bar{Y})^2} & \text{si } SST > \epsilon \\ \text{NA} & \text{si } SST \leq \epsilon \end{cases}$$

1.6.4 Descomposición de Theil

El MSE se descompone en componentes interpretables:

$$\text{MSE} = \underbrace{(\bar{\hat{Y}} - \bar{Y})^2}_{\text{Sesgo}^2} + \underbrace{(\sigma_{\hat{Y}} - \sigma_Y)^2}_{\text{Var. diferencial}} + \underbrace{2\sigma_{\hat{Y}}\sigma_Y(1 - \rho_{\hat{Y},Y})}_{\text{Covarianza imperfecta}}$$

Las proporciones (bias_prop, var_prop, cov_prop) diagnostican la fuente principal del error predictivo.

1.7 Criterios de Estabilidad Temporal

1.7.1 Métrica de Soporte

Definimos el soporte como:

$$\text{support} = \frac{\text{folds_proceed}}{\text{folds}}$$

donde `folds_proceed` cuenta los folds que pasan todos los filtros econométricos y tienen desempeño predictivo aceptable.

1.7.2 Umbrales de Validación

- **Umbral estricto:** $\text{support} \geq 0.75$ y $\text{folds_proceed} \geq 5$
- **Umbral moderado:** $\text{support} \geq 0.60$ y $\text{folds_proceed} \geq 3$

El requisito de mínimo absoluto previene “falsos robustos” por denominador pequeño.

1.7.3 Métricas Ajustadas por Estabilidad

- **R^2 estable:** $R^2_{\text{stab}} = R^2 \times \text{support}$
- **U estable:** $U_{\text{stab}} = \frac{U}{\text{support}}$ (penaliza inestabilidad)

Estas métricas integran desempeño predictivo con consistencia temporal, privilegiando modelos “buenos y constantes” sobre “excelentes a veces”.

1.8 Implementación Computacional

1.8.1 Paralelización Multi-Nivel

La arquitectura de paralelización opera en dos niveles:

1. **Nivel de pares:** Cada combinación ($X \rightarrow Y$) se procesa en un worker independiente
2. **Control de BLAS:** Se fija `blas_set_num_threads(1)` para evitar sobresuscripción del CPU

Los workers son procesos R independientes (no threads) coordinados por `future::multisession`, cada uno con memoria propia. La semilla `future.seed=TRUE` garantiza reproducibilidad en paralelo.

1.8.2 Gestión de Progreso

El paquete `progressr` proporciona retroalimentación en tiempo real sin interferir con la paralelización, crucial para ejecuciones largas (84 pares × múltiples folds × validación anidada).

1.8.3 Complejidad Computacional

La complejidad total es:

$$\mathcal{O}(N_{\text{pares}} \times F_{\text{ext}} \times F_{\text{int}} \times G \times C_{\text{modelo}})$$

donde: - $N_{\text{pares}} = 84$ (6 circulación \times 7 producción \times 2 direcciones) - F_{ext} = número de folds externos (típicamente 8-15) - F_{int} = folds internos por fold externo (típicamente 3-5) - G = tamaño de