

# Pipeline de Indicadores Macroeconômicos e Modelos de Séries Temporais (VAR, VECM, PCA–ARX e Markov-Switching)

Relatório Técnico

24 de setembro de 2025

## Sumário

<b>1</b>	<b>Visão geral</b>	<b>2</b>
<b>2</b>	<b>Atualizações do pipeline (versão atual)</b>	<b>2</b>
<b>3</b>	<b>Pipeline de dados (resumo do código)</b>	<b>2</b>
3.1	Janela temporal e agregação . . . . .	2
3.2	Fontes e robustez . . . . .	2
<b>4</b>	<b>Transformações para modelagem</b>	<b>3</b>
<b>5</b>	<b>Resultados empíricos mais recentes</b>	<b>3</b>
5.1	VAR (6 variáveis, seleção por AIC) . . . . .	3
5.2	VECM (USD/BRL e spread SELIC–Fed) . . . . .	5
5.3	PCA–ARX (retorno do Ibovespa) . . . . .	5
5.4	Markov-Switching (2 regimes) . . . . .	5
<b>6</b>	<b>Como executar</b>	<b>5</b>
<b>7</b>	<b>Diagnósticos e cuidados</b>	<b>6</b>
<b>8</b>	<b>Extensões</b>	<b>6</b>

# 1 Visão geral

Este relatório documenta:

- O **pipeline de coleta e preparação** de dados macroeconômicos (últimos 10 anos), que produz um Excel consolidado em:

`C:\Users\Lenovo\Desktop\Desktop\Mestrado FGV\IndicadoresMacro\indicadores_macro.xlsx`

- Quatro **modelos de séries temporais** aplicados aos dados salvos:
  1. VAR (previsão multivariada e IRFs);
  2. VECM (cointegração em níveis, ex.: câmbio e diferencial de juros);
  3. PCA-ARX (fatores dinâmicos para prever retorno do Ibovespa);
  4. Markov-Switching (mudança de regime de volatilidade/retornos).

## 2 Atualizações do pipeline (versão atual)

- **Treasury 10 anos (% a.a.)**: inclusão de *fix* automático de *unidade*. Se a mediana vier em fração (e.g.,  $0,04 = 4\%$ ), o script multiplica por 100; se vier  $\approx 0,4$  (deveria ser 4,0), multiplica por 10. Fallback permanece  $\hat{TNX}/10$ .
- **PIB EUA (% a/a)**: passa a ser baixado do FRED (A191RL1Q225SBEA), marcado no fim do trimestre e reamostrado para ME.
- **Crédito/Endividamento/Inadimplência**: mapeamento ampliado (PF, PJ, livres, total, %PIB e % renda).
- **Estrutura para Ipeadata/SIDRA/SGS extra**: dicionários de mapeamento (IPEA\_MAP, SIDRA\_MAP, SGS\_EXTRA) prontos para ativar novas séries sem mudar o pipeline.
- **Metadados e médias anuais**: aba adicional com médias anuais (média simples das observações mensais de cada ano).

## 3 Pipeline de dados (resumo do código)

### 3.1 Janela temporal e agregação

A amostra cobre os **últimos 10 anos** até o *último mês fechado*. Séries diárias são **reamostradas para ME** (*month-end*) por *last* ou *mean*.

### 3.2 Fontes e robustez

**Yahoo Finance** com *yfinance* (coluna *Adj Close* ou *Close*).

**PTAX (Olinda/BCB)** com fatiamento anual e **fallback** para `USDBRL=X`.

**SGS (python-bcb)** com janelas de 5 anos e concatenação.

**FRED** (DGS10, DFEDTARU/DFEDTARL, PIB EUA YoY). Fallback UST10 por  $\hat{\text{TNX}}/10$  quando necessário.

## 4 Transformações para modelagem

- Índices/preços:  $\Delta \ln x_t$  (retornos/crescimentos);
- Juros: `pct_change`;
- VAR em estacionários; VECM para níveis  $I(1)$  quando houver cointegração.

## 5 Resultados empíricos mais recentes

Esta seção resume **os resultados que você rodou** com o Excel gerado pela versão atual do pipeline.

### 5.1 VAR (6 variáveis, seleção por AIC)

**Especificação.**  $y_t = \{\Delta \ln \text{IBC}, \Delta \ln \text{USD/BRL}, \Delta \ln \text{Brent}, \Delta \ln \text{S\&P500}, \Delta \text{Selic}, \Delta \text{UST10}\}$ .

**Coefficientes relevantes (p-valores).**

- **Atividade** ( $\Delta \ln \text{IBC}$ ): efeito negativo de  $\text{UST10}_{t-1}$  sobre o crescimento ( $p \approx 0,003$ ).
- **Brent**: `ret_spxt-1` positivo ( $p \approx 0,044$ ) e  $\text{UST10}_{t-1}$  positivo ( $p \approx 0,013$ ).
- **S&P 500**: inércia negativa em  $t-4$  ( $p \approx 0,029$ );  $\text{Selic}_{t-5}$  positiva ( $p \approx 0,009$ );  $\Delta \ln \text{Brent}_{t-8}$  negativa ( $p \approx 0,012$ );  $\text{Selic}_{t-9}$  positiva ( $p \approx 0,045$ ).
- **Selic**:  $\Delta \ln \text{IBC}_{t-2}$  positiva ( $p \approx 0,031$ ); `ret_spxt-2` positiva ( $p \approx 0,004$ );  $\text{UST10}_{t-3}$  positiva ( $p \approx 0,036$ );  $\Delta \ln \text{USD/BRL}_{t-7}$  positiva ( $p \approx 0,029$ ).
- **UST10**: persistência ( $\text{UST10}_{t-2}$ ,  $p \approx 0,025$ ); ligações com Brent (+ em  $t-6$ ,  $p \approx 0,014$ ; - em  $t-11$ ,  $p \approx 0,015$ ); e com `ret_spxt-10` ( $p \approx 0,029$ ).

**IRFs.** As funções resposta ao impulso (Figura 1) mostram impactos moderados e bandas largas; choques de **UST10** tendem a reduzir o crescimento do **IBC** nos meses seguintes; **Brent** e **S&P** se influenciam mutuamente.

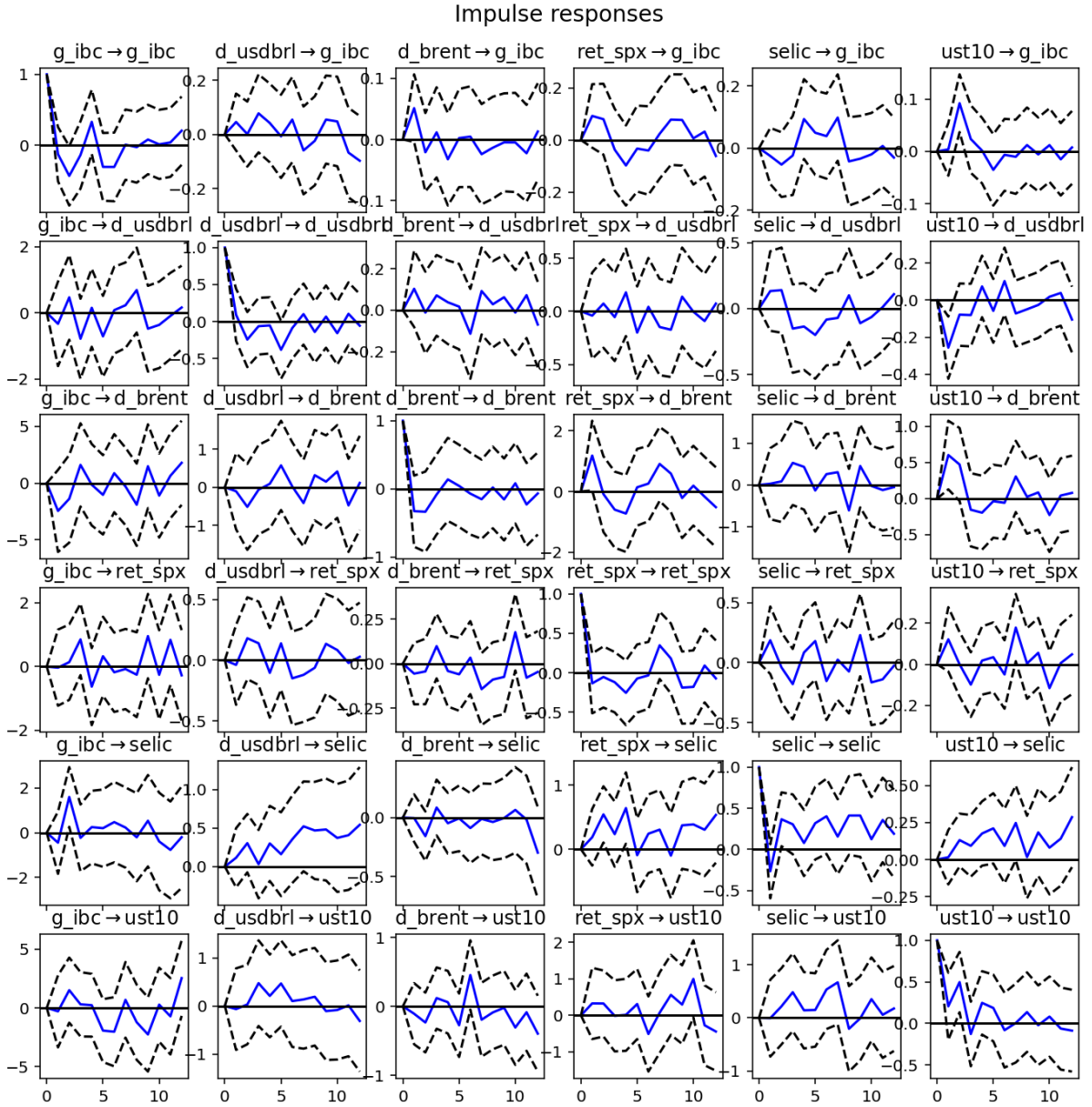


Figura 1: IRFs do VAR (horizonte de 12 meses).

**Previsão (12 meses).** As variações mensais previstas são pequenas (reversão à média). Primeiras 5 linhas (ilustrativo):

Data	$\Delta \ln \text{IBC}$	$\Delta \ln \text{USD}$	$\Delta \ln \text{Brent}$	$\Delta \ln \text{S\&P}$	$\Delta \text{Selic}$	$\Delta \text{UST10}$
2025-08-31	0,027	-0,028	0,033	0,024	0,021	-0,076
2025-09-30	-0,010	0,042	-0,112	0,012	0,001	0,068
2025-10-31	-0,015	0,014	-0,025	-0,030	0,055	-0,176
2025-11-30	-0,015	0,066	-0,203	-0,056	0,001	-0,006
2025-12-31	-0,008	-0,002	-0,032	0,036	-0,027	-0,079

## 5.2 VECM (USD/BRL e spread SELIC–Fed)

**Teste de Johansen.** Estatística traço = 13,38 para  $r = 0$  e críticos de 95% = 15,49; para  $r = 1$ , traço = 4,72 e crítico = 3,84. **Conclusão:** não há cointegração a 95% na amostra corrente (spread com *proxy* UST10 em vez do Fed Funds real).

**Implicação.** VECM não é indicado nesta configuração; use VAR em diferenças ou reestime com **Fed Funds** via FRED (média da banda) e/ou amostra mais longa.

## 5.3 PCA–ARX (retorno do Ibovespa)

**Ajuste.**  $R^2 \approx 0,7\%$ ; coeficientes dos fatores não significativos; intercepto quase significativo ( $p \approx 0,052$ ).

**Sinal de 1 passo** (mês seguinte):  $\hat{r}_{t+1} \approx 0,009$  (cerca de 0,9%).

**Leitura.** Em base mensal e usando apenas 2 fatores macro/mercado, o poder preditivo é baixo — condizente com literatura. Sugere-se testar horizontes mais longos, mais fatores e validação fora da amostra.

## 5.4 Markov-Switching (2 regimes)

**Ajuste atual.** AIC  $-398,4$ ; porém, as probabilidades suavizadas colapsaram para  $\approx 1,0$  em sequência e houve `ConvergenceWarning` e matriz quase singular.

**Diagnóstico.** *Colapso de regime* (um regime com variância quase zero). Provável causa: escala dos dados ou especificação muito restrita.

**Correções sugeridas.**

- Usar **retornos padronizados**  $(r - \bar{r})/\sigma$ ;
- Estimar com `switching_mean=True` além de `switching_variance=True`;
- Fornecer chutes/EM (`em_iter=10`) e aumentar `maxiter`;
- Se persistir, testar 3 regimes.

## 6 Como executar

### Pré-requisitos de Python

- pandas, numpy, requests, yfinance, statsmodels, scikit-learn, python-dateutil;
- FRED: definir `FRED_API_KEY` no ambiente (para Fed Funds & PIB EUA);
- SGS: python-bcb.

## Fluxo

1. Rodar o **pipeline** (`indicadores_macro.xlsx`);
2. Rodar os blocos dos modelos (1 a 4).

## 7 Diagnósticos e cuidados

- Conferir **metadados** quando houver *fallbacks*;
- Em **pandas**, preferir frequência **ME** (aviso de depreciação para “M”);
- VAR: verificar *estabilidade* e autocorrelação dos resíduos; FEVD e Granger ajudam na interpretação;
- Estratégias: incorporar custos e gestão de risco.

## 8 Extensões

- Ativar Ipeadata/SIDRA/SGS extra (IPCA, confiança, fiscal/dívida, desemprego);
- DFM/Kalman e TVP-VAR; combinação de sinais condicionada ao regime do Markov.