

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_k X_{t-k} + u_t.$$

Esse modelo é conhecido como um **modelo de defasagem distribuída**. Note a pequena mudança na notação; utilizamos um subscrito nos nossos coeficientes β indicando o número de períodos em que a variável está defasada em relação ao seu valor atual; assim, o β para X_t é β_0 (porque $t - 0 = t$). Nessa configuração, o **impacto acumulativo** β de X em Y é

$$\beta = \beta_0 + \beta_1 + \beta_2 + \dots + \beta_k = \sum_{i=0}^k \beta_i.$$

Vale enfatizar que estamos interessados no impacto acumulativo de X em Y , não meramente no **efeito instantâneo** de X_t sobre Y_t representado pelo coeficiente β_0 .

Mas como podemos capturar os efeitos de X em Y sem estimar um modelo bizarro como o anterior? Como pontuamos, um modelo como esse certamente sofreria de multicolinearidade.

Se estamos dispostos a assumir que o efeito de X em Y é inicialmente maior e diminui geometricamente a cada período (finalmente, após muitos períodos, tornando-se 0), então o uso de alguma álgebra geraria o seguinte modelo matematicamente idêntico ao anterior¹¹. Esse novo modelo seria dado por

$$Y_t = \lambda Y_{t-1} + \alpha + \beta_0 X_t + v_t.$$

Esse processo é conhecido como **transformação Koyck** e é frequentemente referido como **modelo da variável dependente defasada**, por razões que esperamos que sejam óbvias. Compare a transformação Koyck com o modelo de defasagem distribuída anterior. Ambos têm a mesma variável dependente, Y_t , ambos têm uma variável que representa o impacto imediato de X_t em Y_t . Mas, enquanto o modelo de defasagem distribuída também possui uma enorme quantidade de coeficientes para representar todas as defasagens de 1 até k para X em Y_t , o modelo da variável dependente defasada contém uma única variável e coeficiente, λY_{t-1} . Como as duas configurações são equivalentes, significa que a variável dependente defasada não representa como Y_{t-1} de algum modo causa Y_t , mas, em vez disso, Y_{t-1} é um substituto para o efeito acumulado de todos os valores defasados de X (isto é, para as defasagens de 1 até k) em Y_t . Conseguimos estimar esse efeito a partir da estimação de um único coeficiente, em vez de um grande número de parâmetros.

O coeficiente λ , então, representa como os valores anteriores de X afetam o valor atual de Y , o que resolve satisfatoriamente o problema sublinhado no início desta seção. Normalmente, o valor de λ variará entre¹² 0 e 1. Você pode facilmente observar

¹¹ Sabemos que esse modelo não parece matematicamente idêntico, mas ele é. Para simplificar a apresentação, pulamos a álgebra necessária para demonstrar a equivalência.

¹² Na verdade, valores próximos de 1 e, especialmente, maiores que 1 indicam que existem problemas com o modelo, provavelmente relacionados à existência de tendências nos dados.