

Without loss of generality let $\pi_{n \rightarrow m} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{m \times m}$ with $m < n$ be the operator that projects a $n \times n$ matrix onto a $m \times m$ matrix, by simply taking the first m rows and columns.¹

Let $L_n : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ be the Leontief operator, that maps a matrix A to it's Leontief matrix, so

$$L_n A = (I - A)^{-1} = \sum_{i=0}^{\infty} A^i.$$

The series is often called **Neumann series**.

Let $A \in \mathbb{R}^{n \times n}$ be our technical coefficient matrix arising from empirical data. Dependency shares are defined as followed in the appendix and are also calculated like this in the replication code

$$DS = \frac{v_s(\pi_{n \rightarrow s}(L_n A))(y_{n \rightarrow s} + z_{n \rightarrow s})}{\|v_s\|_1}.$$

The issue lies in the way the Leontief matrix is chosen and the role of intermediate demand, more concretely

$$\pi_{n \rightarrow s}(L_n A) \neq L_s(\pi_{n \rightarrow s} A)$$

This fact can be easily seen by looking at the **Neumann Series**.

The problem is, that $\pi_{n \rightarrow s} L_n A$ also captures indirect demand from sectors $j > s$. This can be seen by looking at the final demand.

Let

$$\tilde{\pi}_{n \rightarrow s} : \mathbb{R}^n \rightarrow \mathbb{R}^s$$

be the projection that takes the first s entries of a vector, Then

$$\tilde{\pi}_{n \rightarrow s}((L_n A)_{ns} \tilde{\pi}_{n \rightarrow s} y) = (\pi_{n_s}(L_n A))(\tilde{\pi}_{n \rightarrow s} y).$$

This can be easily verified by looking at the definition of matrix multiplication.

So it does not matter if the full leontief matrix is chosen or just a submatrix to get the souths demand, when first calculating the Leontief with the full technological coefficient matrix. This means all information about intermediates is already in the leontief coefficients and intermediates from the north are already captured.

Instead the following approach has to be chosen.

Total production is given as

$$y_i = \sum_{j=1}^s a_{i \rightarrow j} y_j + \sum_{j=s+1}^n x_{i \rightarrow j} + c_i$$

for all $i < n$, or:

$$\begin{aligned} y &= A_{ss} y + z + c \\ z + c &= \underbrace{(I - A_{ss})}_{L \pi_{n \rightarrow s} A} y. \\ y &= (I - A_{ss})^{-1} (z + c) \end{aligned}$$

Thus the correct definition of the total demand shares should be either

¹This is essentially what is done when doing `eora_mario.w.loc[to_region_idx, to_region_idx]`

$$\text{DS} = \frac{v_s(\pi_{n \rightarrow s}(\boldsymbol{L}_n A))(y_{n \rightarrow s})}{\|v_s\|_1}.$$

or

$$\text{DS} = \frac{v_s((\boldsymbol{L}_s \pi_{n \rightarrow s} A))(y_{n \rightarrow s} + z_{n \rightarrow s})}{\|v_s\|_1}.$$