We will demonstrate 2 ways to derive a pair trading strategy.

First let us define some terms (Engle and Granger 1987)

1. A series with no deterministic component which has a stationary, invertible, ARMA representation after differencing $d$ times, is said to be integrated of order $d$, denoted $x_t = I(d)$ where $x_t \in \mathbb{R}^k$
2. The components of the vector $x_t$ are said to be *co-integrated of order d,b*, denoted $x_t \sim CI(d, b)$, if (i) all components of $x_t$ are $I(d)$; (ii) there exists a vector $\alpha(\neq 0)$ so that $z_t = \alpha^T x_t \sim I(d-b), b > 0$, where $\alpha$ is called the co-integrating vector

We consider d=b=1 case.

Method 1: Vector Error Correction Model (VECM)

Suppose each components of $x_t$ are $I(1)$. Then $\Delta x_t = x_t - x_{t-1}$ is stationary. The idea is to find a stationary linear combination of the components of $x_t$, i.e. $p_t = u^T x_t$

Consider a VAR(1) model, we have that

$$x_t = \Phi x_{t-1} + a_t \Rightarrow \Delta x_t = -(I - \Phi)x_{t-1} + a_t \tag{1}$$

In other words,

$$(I - \Phi)x_{t-1} = -\Delta x_t + a_t \quad \text{ is trend stationary} \tag{2}$$

For any $u^T$, $u^T(I - \phi)x_{t-1} = -\Delta x_t + a_t$ is a stationary portfolio.

If $u^T(I - \Phi) = v^T \neq 0$, then $v^T$ is a cointegration vector.

Applying to a VAR(2) model, we have that

$$x_t = \phi_0 + \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + a_t \Rightarrow \Delta x_t = \phi_0 + \Pi x_{t-1} - \Phi_2 \Delta r_{t-1} + a_t \tag{3}$$

where the error correction form is defined as

$$\Delta x_t = \phi_0 + \Pi x_{t-1} - \Phi_2 \Delta r_{t-1} + a_t \tag{4}$$

with $\Pi = \Phi_1 + \Phi_2 - I$

Recall our objective is to find a stationary linear combination of components of $x_t$, i.e. $p_t = u^T x_t$. Suppose there exist a constant scalar $\lambda$ s.t. $\lambda u^T = u^T \Pi$, i.e. $u^T$ is the left eigenvector of the VECM matrix $\Pi$ with associated eigenvalue $\lambda, |\lambda| < 1 \text{ and } \lambda \neq 0$. Then $u^T$ is a co-integration vector, since

$$\lambda p_{t-1} = \lambda u^T x_{t-1} = u^T \Pi x_{t-1} = u^T \Delta x_t - u^T \phi_0 + u^T \Phi_2 \Delta r_{t-1} - u^T a_t \tag{5}$$

which implies that $p_{t-1}$ is stationary since

$$\lambda p_{t-1} = \text{ stationary terms} \Rightarrow p_{t-1} = \frac{1}{\lambda} \text{ stationary terms} \tag{6}$$

Hence we can find a co-integration vector $u^T$ by finding the left eigenvectors of $\Pi$ with eigenvalues $|\lambda| < 1$ and obtain a trading strategy $p_t = u^T x_t$.

To check for the existence of co-integration, we check the rank of $\Pi$:

1. Rank is 0: no co-integration

2. Rank is k: a stationary process, no integration
3. Rank is 0<m<k: co-integration, with $k - m$ unit root. Each left eigenvector of $\Pi$ with nonzero eigenvalue signals a linearly independent portolfio trading strategy.

Method 2: Engle-Granger (regression method)

Suppose each component of $x_t$ is integrated of order 1, i.e. $x_t \sim I(1)$, when the existence of co-integration implies that a linear combination of each time series is stationary,

$$w_t = v_1 x_{1,t} + v_2 x_{2,t} + \ldots + v_k x_{k,t} \quad \text{is stationary} \tag{7}$$

Or equivalently,

$$x_{1,t} = -(v_1^{-1} v_2 x_{2,t} + \ldots + v_1^{-1} v_k x_{k,t}) + v_1^{-1} w_t \tag{8}$$

which is a linear regression model ($x_{1,t} \sim x_{2,t}, \ldots, x_{k,t}$) and can be obtained through OLS. $w_t$ is not i.i.d so regress with time series error.

The co-integration vector can be obtained by inverting the regression coefficients.