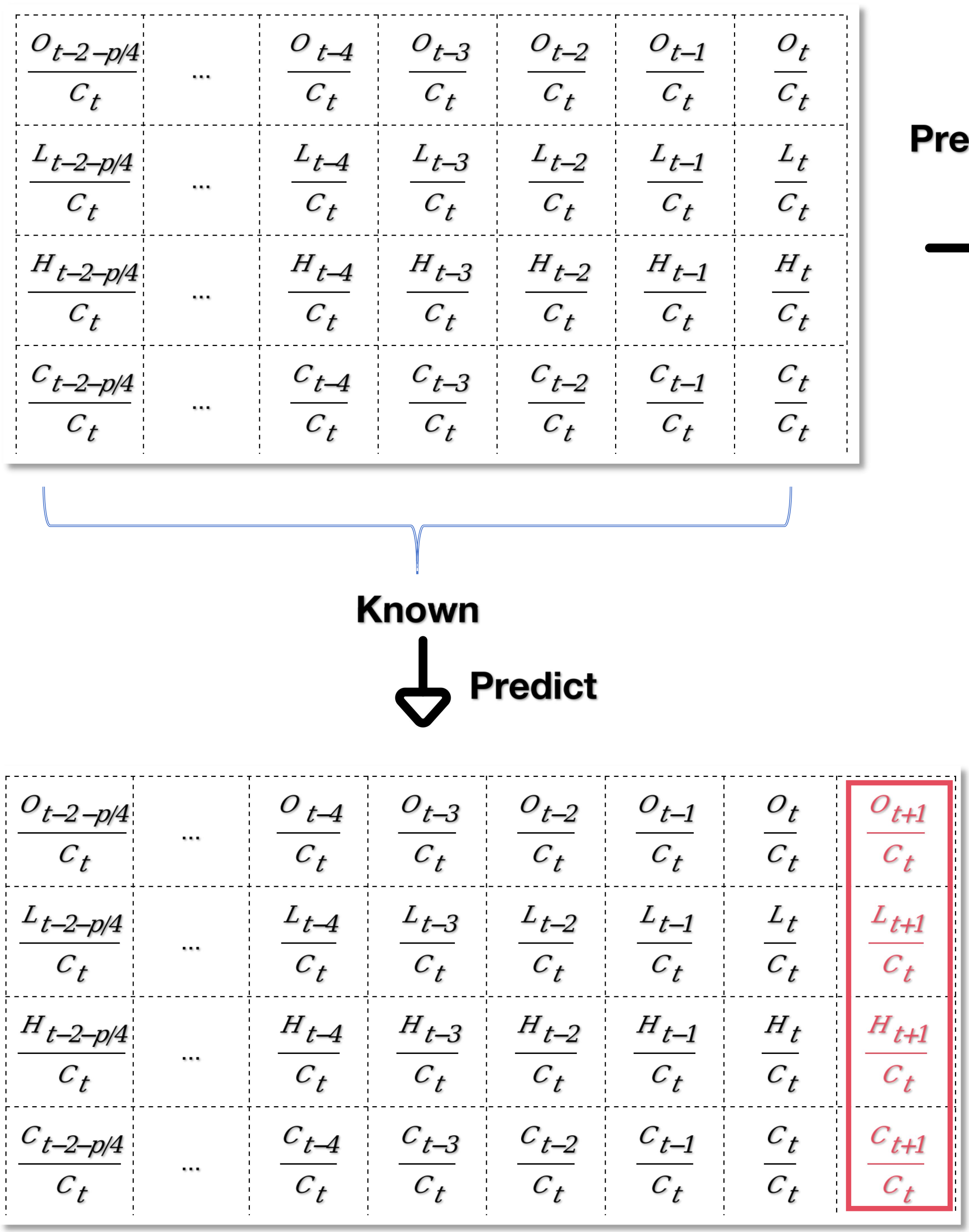
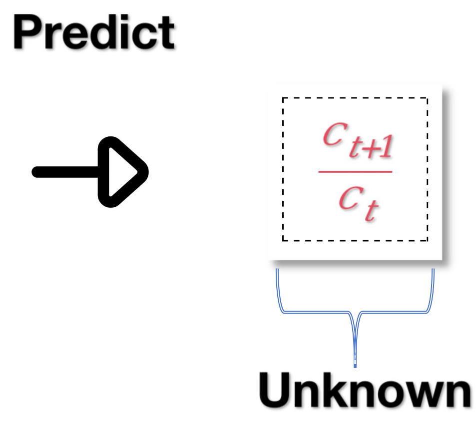
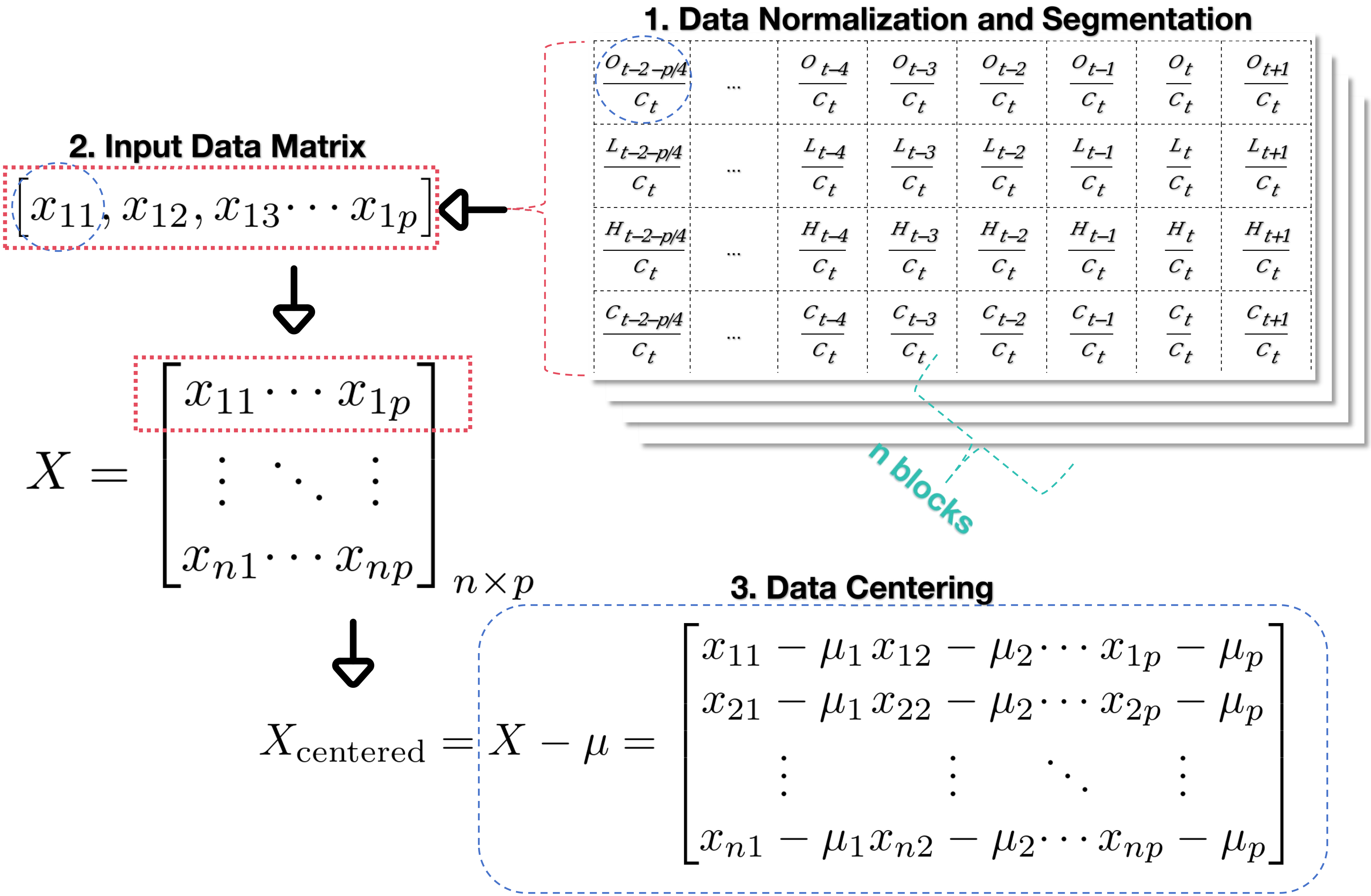




Change the approach: shift from predicting individual future prices to forecasting an integrated information entity encompassing both past and future prices.

We need to integrate future price information with past price information into a unified framework.

Method: PCA-**Principal Component Analysis**


1. Data Normalization and Segmentation
2. Input Data Matrix
n blocks
3. Data Centering
$[x_{11}, x_{12}, x_{13} \cdots x_{1p}]$
$X = \begin{bmatrix} x_{11} \cdots x_{1p} \\ \vdots \ddots \vdots \\ x_{n1} \cdots x_{np} \end{bmatrix}_{n \times p}$
$X_{\text{centered}} = X - \mu = \begin{bmatrix} x_{11} - \mu_1 & x_{12} - \mu_2 & \cdots & x_{1p} - \mu_p \\ x_{21} - \mu_1 & x_{22} - \mu_2 & \cdots & x_{2p} - \mu_p \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} - \mu_1 & x_{n2} - \mu_2 & \cdots & x_{np} - \mu_p \end{bmatrix}$

4. Covariance Matrix

$$C = \frac{1}{n} X_{\text{centered}}^{\top} X_{\text{centered}} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix}$$

5. Eigenvalue Computation

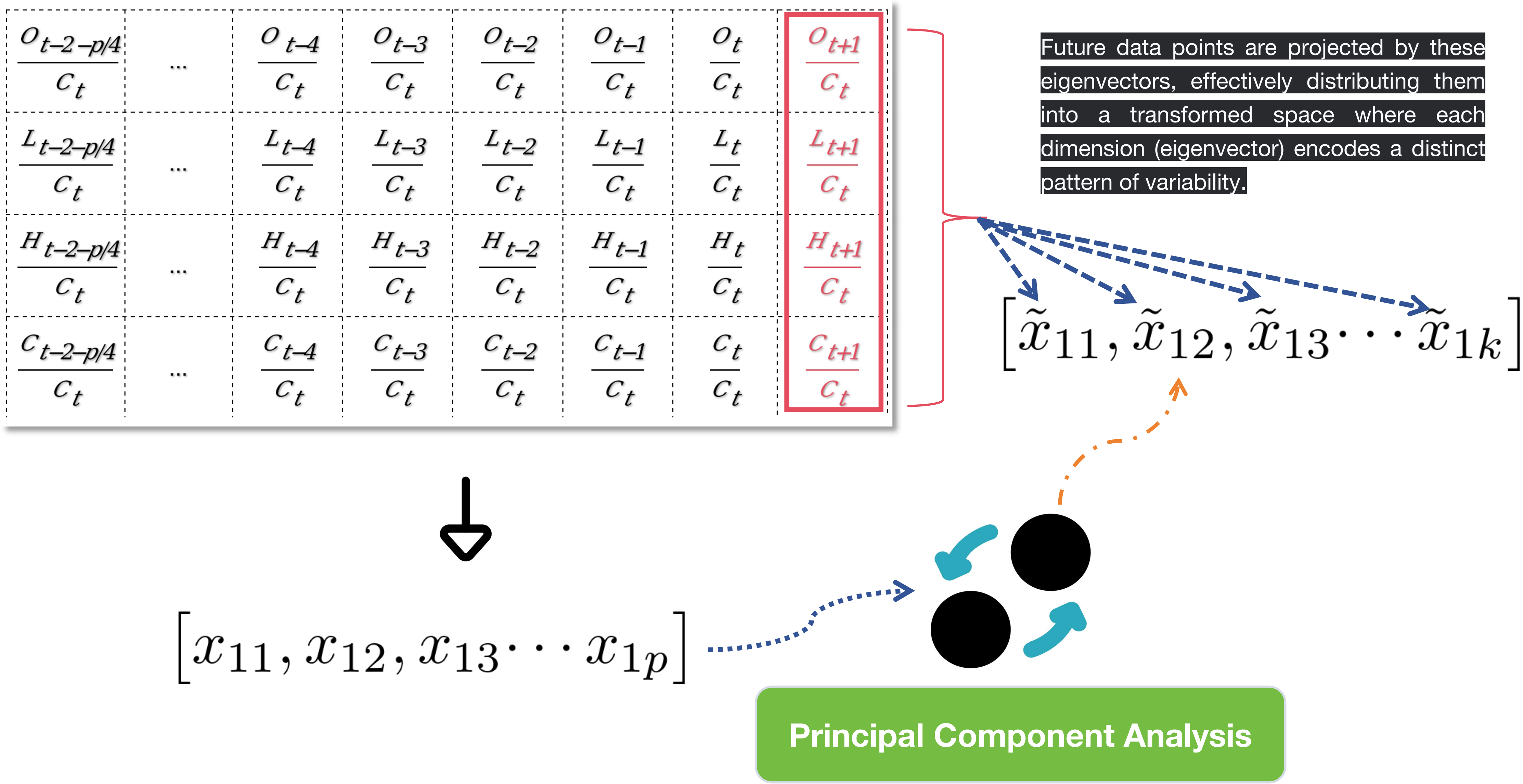
$$\det(C - \lambda I) = \begin{vmatrix} c_{11} - \lambda & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} - \lambda & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} - \lambda \end{vmatrix} = 0$$

$$\begin{bmatrix} c_{11} - \lambda_i & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{p1} & \cdots & c_{pp} - \lambda_i \end{bmatrix} \begin{bmatrix} v_{i1} \\ \vdots \\ v_{ip} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow W = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & & | \end{bmatrix}_{p \times k}$$

$$X_{\text{pca}} = X_{\text{centered}} W \in \mathbb{R}^{n \times k}$$

$$\Downarrow$$

$$\boxed{\begin{bmatrix} \tilde{x}_{11} & \cdots & \tilde{x}_{1k} \\ \vdots & \ddots & \vdots \\ \tilde{x}_{n1} & \cdots & \tilde{x}_{nk} \end{bmatrix}} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} v_{11} & \cdots & v_{1k} \\ \vdots & \ddots & \vdots \\ v_{p1} & \cdots & v_{pk} \end{bmatrix}$$


Future data points are projected by these eigenvectors, effectively distributing them into a transformed space where each dimension (eigenvector) encodes a distinct pattern of variability.
$[\tilde{x}_{11}, \tilde{x}_{12}, \tilde{x}_{13} \cdots \tilde{x}_{1k}]$
$[x_{11}, x_{12}, x_{13} \cdots x_{1p}]$
Principal Component Analysis

$$\tilde{X} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \cdots & \tilde{x}_{1k} \\ \tilde{x}_{21} & \tilde{x}_{22} & \cdots & \tilde{x}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{n1} & \tilde{x}_{n2} & \cdots & \tilde{x}_{nk} \end{bmatrix}_{n \times k}$$

1. Calculate Row-wise Mean

$$m = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix}_{n \times 1}$$

$$m_i = \frac{1}{k} \sum_{j=1}^{k} \sin(\tilde{x}_{ij})$$

2. Compute Row-wise L2 Norm

$$n = \begin{bmatrix} \|\tilde{x}_1\|_2 \\ \|\tilde{x}_2\|_2 \\ \vdots \\ \|\tilde{x}_n\|_2 \end{bmatrix}_{n \times 1}$$

$$\|\tilde{x}_i\|_2 = \sqrt{\sum_{j=1}^{k} \tilde{x}_{ij}^2}$$

3.Dimensionality Reduction

$$r = m \odot n = \begin{bmatrix} m_1 \times \|\tilde{x}_1\|_2 \\ m_2 \times \|\tilde{x}_2\|_2 \\ \vdots \\ m_n \times \|\tilde{x}_n\|_2 \end{bmatrix}_{n \times 1} \rightarrow r = \begin{bmatrix} \boxed{r_1} \\ r_2 \\ \vdots \\ r_n \end{bmatrix}_{n \times 1}$$

$\frac{O_{t-2-p/4}}{C_t}$	...	$\frac{O_{t-4}}{C_t}$	$\frac{O_{t-3}}{C_t}$	$\frac{O_{t-2}}{C_t}$	$\frac{O_{t-1}}{C_t}$	$\frac{O_t}{C_t}$	$\frac{O_{t+1}}{C_t}$
$\frac{L_{t-2-p/4}}{C_t}$	...	$\frac{L_{t-4}}{C_t}$	$\frac{L_{t-3}}{C_t}$	$\frac{L_{t-2}}{C_t}$	$\frac{L_{t-1}}{C_t}$	$\frac{L_t}{C_t}$	$\frac{L_{t+1}}{C_t}$
$\frac{H_{t-2-p/4}}{C_t}$	...	$\frac{H_{t-4}}{C_t}$	$\frac{H_{t-3}}{C_t}$	$\frac{H_{t-2}}{C_t}$	$\frac{H_{t-1}}{C_t}$	$\frac{H_t}{C_t}$	$\frac{H_{t+1}}{C_t}$
$\frac{C_{t-2-p/4}}{C_t}$	...	$\frac{C_{t-4}}{C_t}$	$\frac{C_{t-3}}{C_t}$	$\frac{C_{t-2}}{C_t}$	$\frac{C_{t-1}}{C_t}$	$\frac{C_t}{C_t}$	$\frac{C_{t+1}}{C_t}$

Clustering