Let $a = a_x \hat{i} + a_y \hat{j}$ be a point in 2D space.

$\Rightarrow$ Here a_x is the length of projection of a on x-axis & a_y is the projection on the y-axis.

We can express a_x & a_y as

$$a_x = \begin{bmatrix} a_x & a_y \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (i)$$

$$a_y = \begin{bmatrix} a_x & a_y \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (ii)$$

Combining (i) & (ii)

$$\begin{bmatrix} a_x & a_y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_x & a_y \end{bmatrix} \quad (iii)$$

$$= \begin{bmatrix} L_{\hat{i}}(a) & L_{\hat{j}}(a) \end{bmatrix}$$

where $L_{\hat{i}}(a)$ is the length of projection of a on $\hat{i}$ & similarly $L_{\hat{j}}(a)$ is for $\hat{j}$.

$\Rightarrow$ Extending the equation in (iii) to any arbitrary vector v_i

$$\begin{bmatrix} a_x & a_y \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 & \dots & v_n \end{bmatrix} = a \cdot V$$

$$\Rightarrow \begin{bmatrix} v_1 & v_2 & v_3 & \dots & v_n \end{bmatrix} = a \cdot V$$

$$\Rightarrow \begin{bmatrix} a \cdot v_1 & a \cdot v_2 & a \cdot v_3 & \dots & a \cdot v_n \end{bmatrix} = a \cdot V$$

$$\Rightarrow \begin{bmatrix} a \cdot v_{1x} & a \cdot v_{2x} & & a \cdot v_{nx} \\ a \cdot v_{1y} & a \cdot v_{2y} & \dots & a \cdot v_{ny} \\ a \cdot v_{1z} & a \cdot v_{2z} & & a \cdot v_{nz} \end{bmatrix} = a \cdot V$$

$$\Rightarrow \begin{bmatrix} L_{v_1}(a) & L_{v_2}(a) & \dots & L_{v_n}(a) \end{bmatrix} = a \cdot V \quad \text{(iv)}$$

If we have more points like a, let's say b, c, ... z, we can extend (iv) as:

$$\Rightarrow \begin{bmatrix} L_{v_1}(a) & L_{v_2}(a) & & L_{v_n}(a) \\ L_{v_1}(b) & L_{v_2}(b) & \dots & L_{v_n}(b) \\ \vdots & \vdots & & \vdots \\ L_{v_1}(z) & L_{v_2}(z) & & L_{v_n}(z) \end{bmatrix} = A \cdot V$$

or

$$L = A \cdot V \quad (v)$$

where the rows of A contain the points

a, b, c ... z.

⇒ Multiplying the equation in (v) by V^T on both sides, we get

$$L V^T = (A V) V^T \quad (vi)$$

If V was to be orthonormal, we could rewrite (vi) as

$$A = L V^T \quad (vii)$$

Now this starts to look like the typical SVD formula we see everywhere

→ If we divide the columns of L by their norms, we get

$$L = U \begin{bmatrix} \sigma_1 & 0 & \cdots & \cdot \\ 0 & \sigma_2 & & \\ \vdots & & \ddots & \\ & & & \sigma_n \end{bmatrix}$$

$$L = U\Sigma \qquad \text{(viii)}$$

→ Merging (vii) & (viii), we get

$$A = U\Sigma V^T \qquad \text{(ix)}$$

where U & V are orthogonal matrices, & Σ is a diagonal matrix.