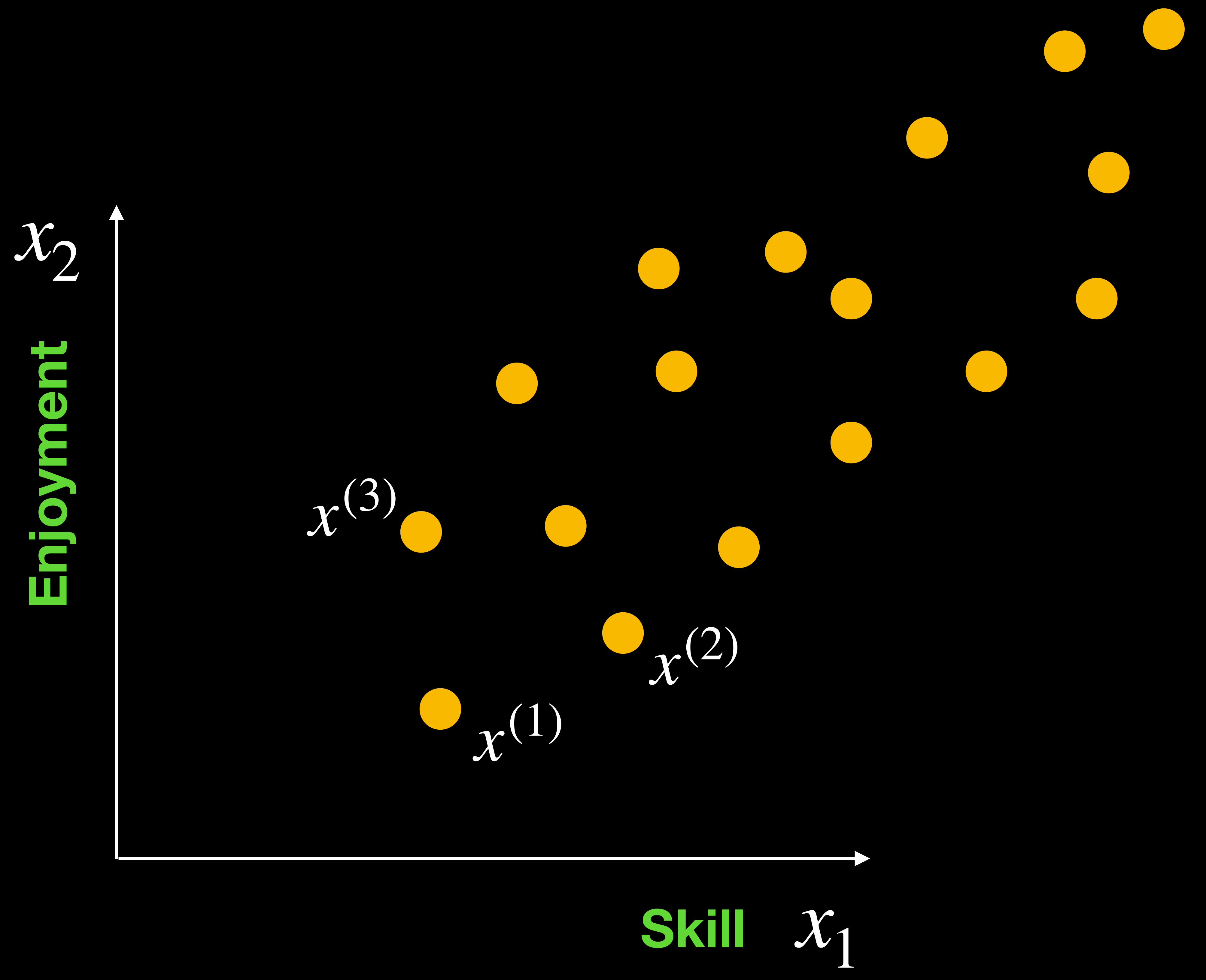
Principal Component Analysis

Principal Component Analysis

Dataset

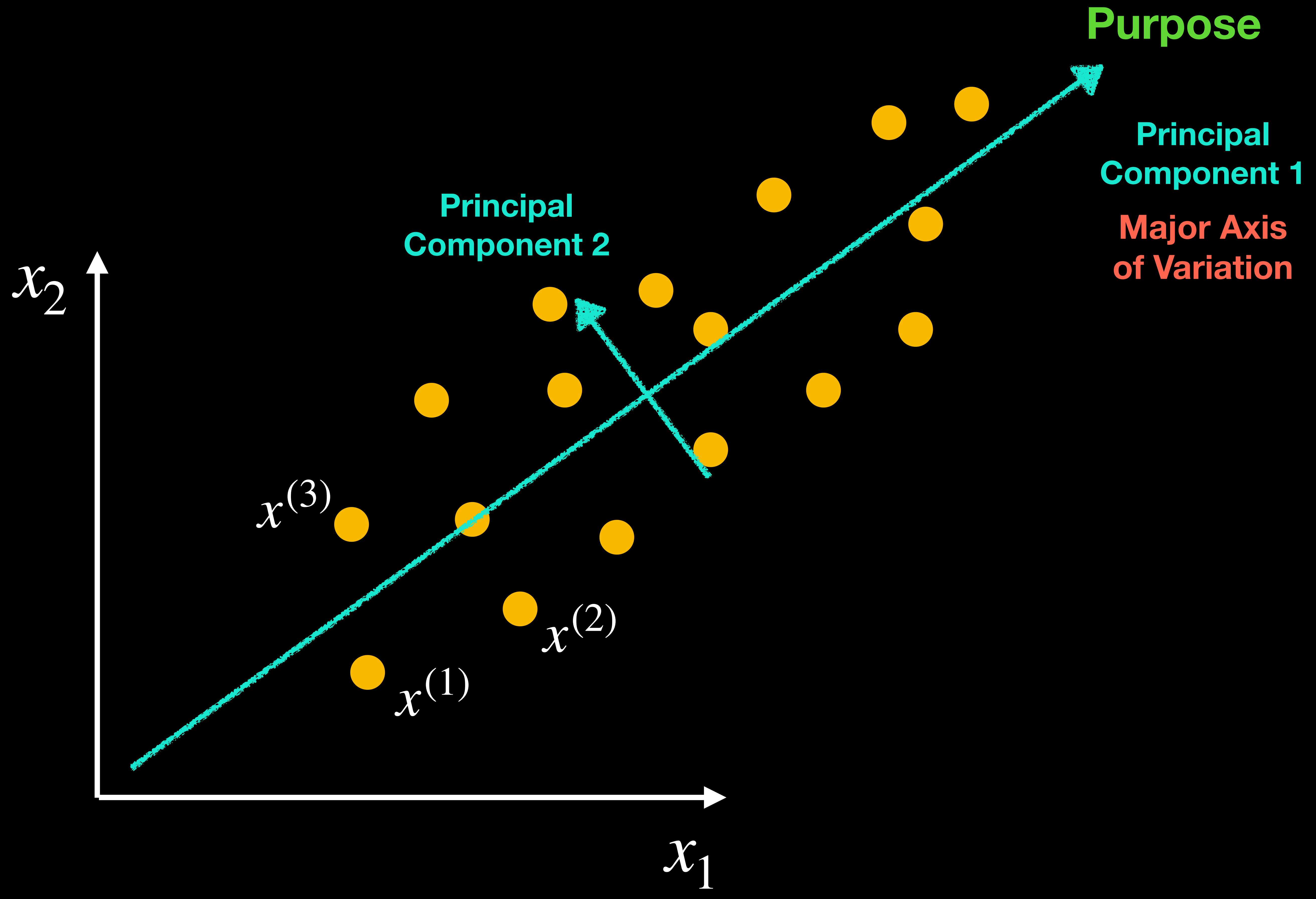
x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...



Principal Component Analysis

Dataset

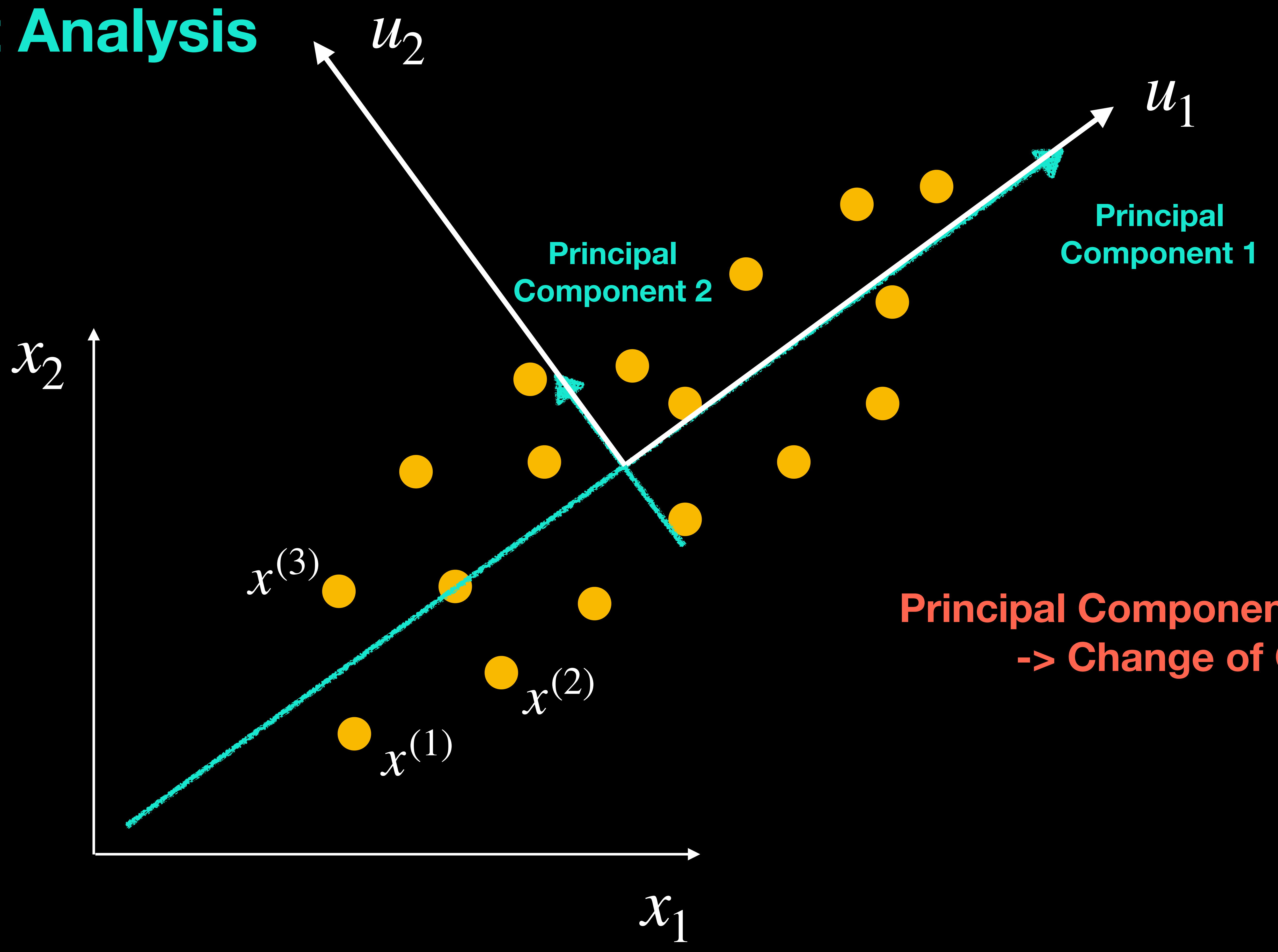
x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...



Principal Component Analysis

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...



Principal Components Are Orthogonal
-> Change of Coordinates

Principal Component Analysis

Mean

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} x_j^{(i)}$$

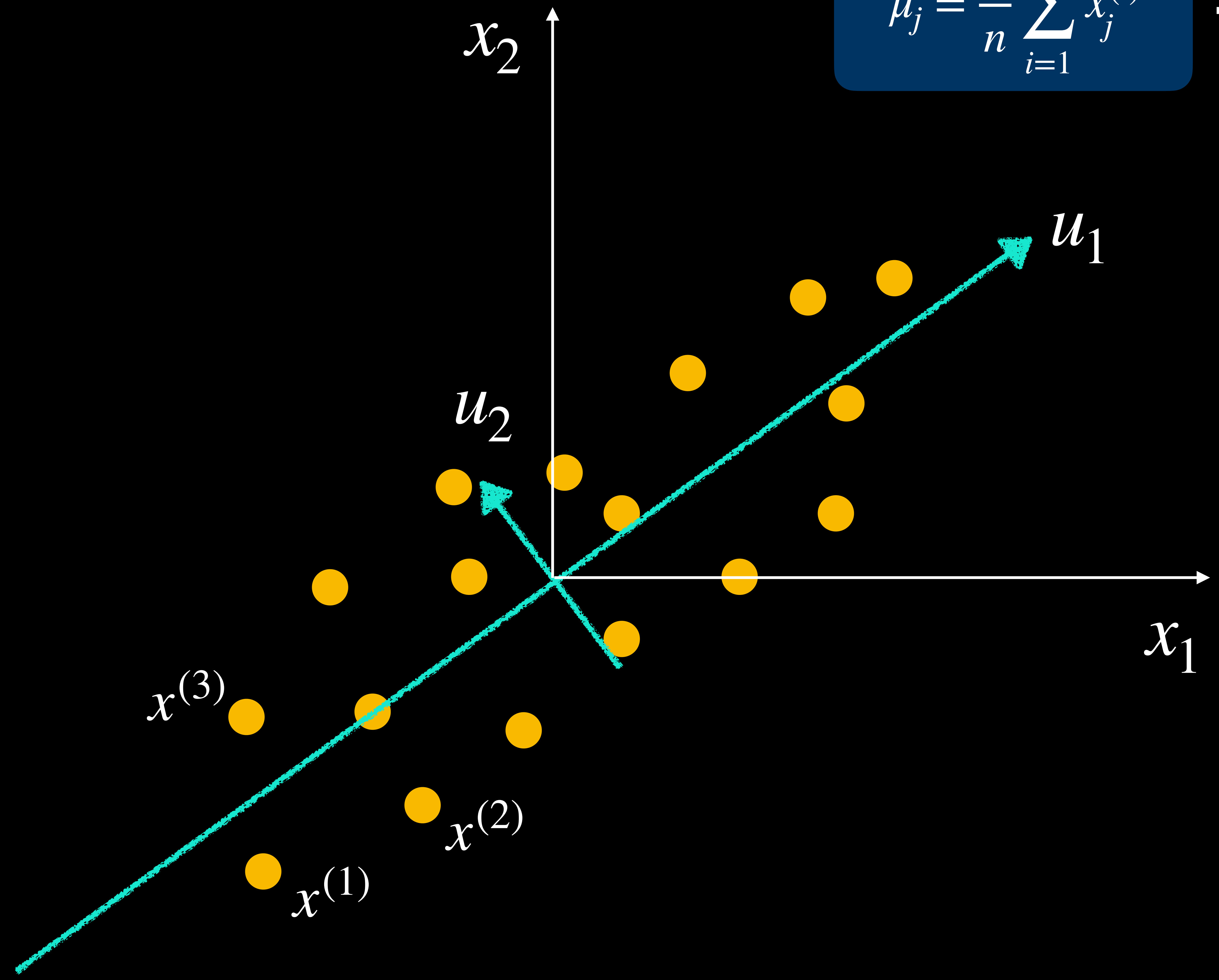
→

Center the data

$$x_j^{(i)} \leftarrow x_j^{(i)} - \mu_j$$

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
…	…



Principal Component Analysis

Mean

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} x_j^{(i)}$$

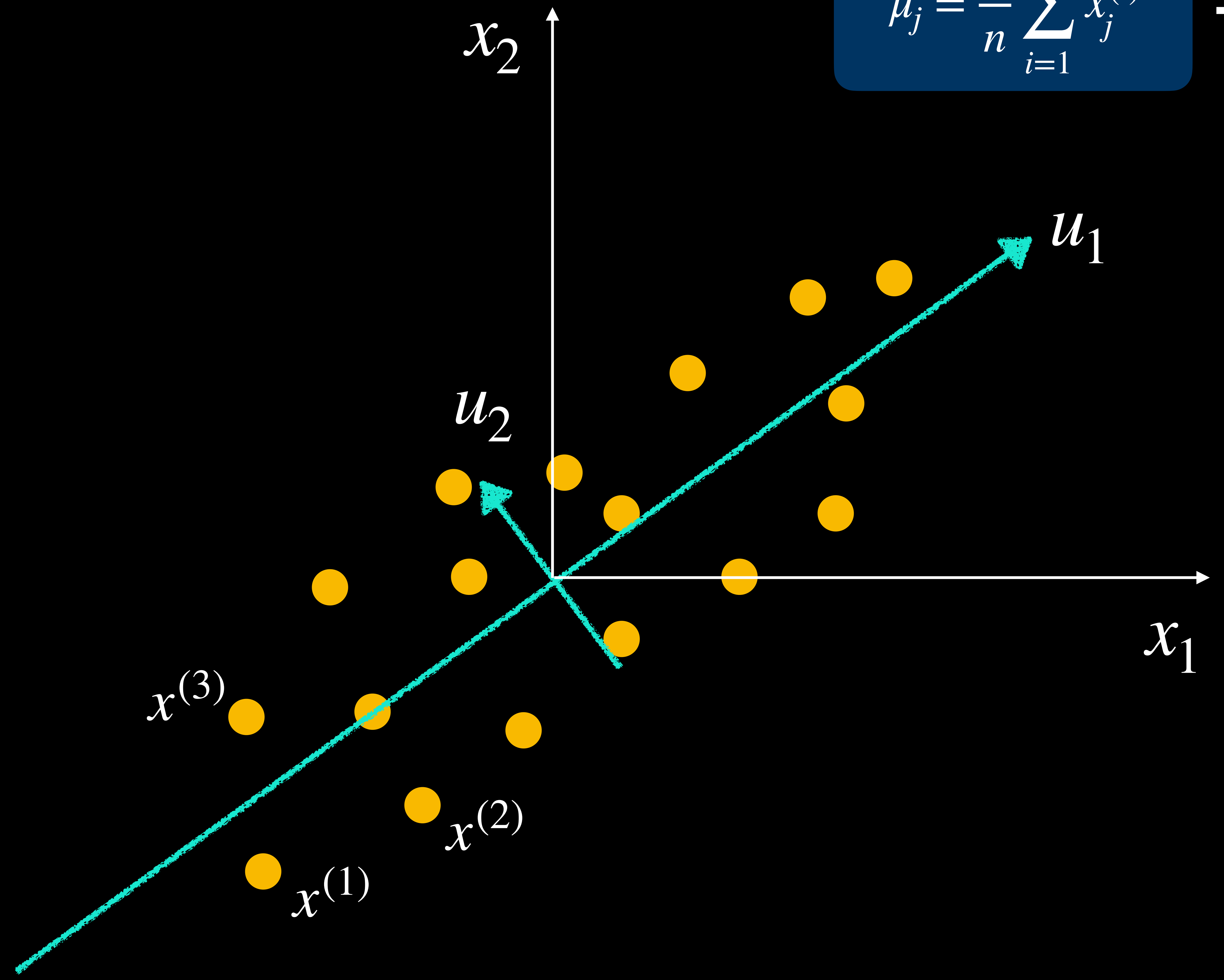
→

Center the data

$$x_j^{(i)} \leftarrow x_j^{(i)} - \mu_j$$

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...



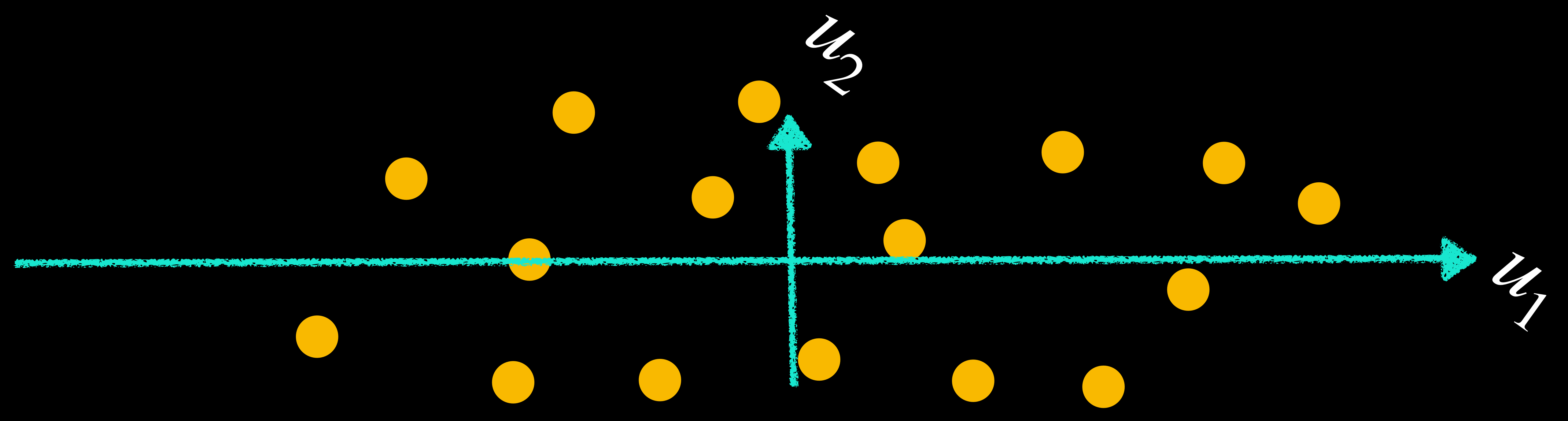
Principal Component Analysis

Data can be **projected**

On axis of highest variation: u_1

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...



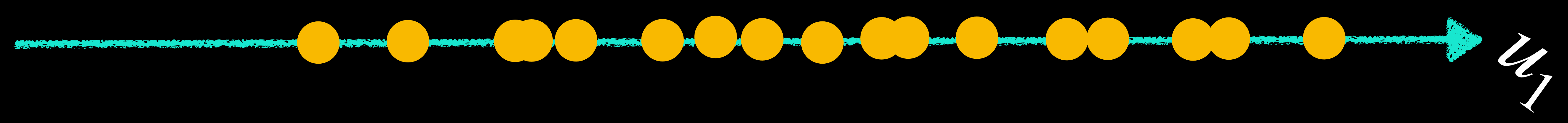
Principal Component Analysis

Data can be **projected**

On axis of highest variation: u_1

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

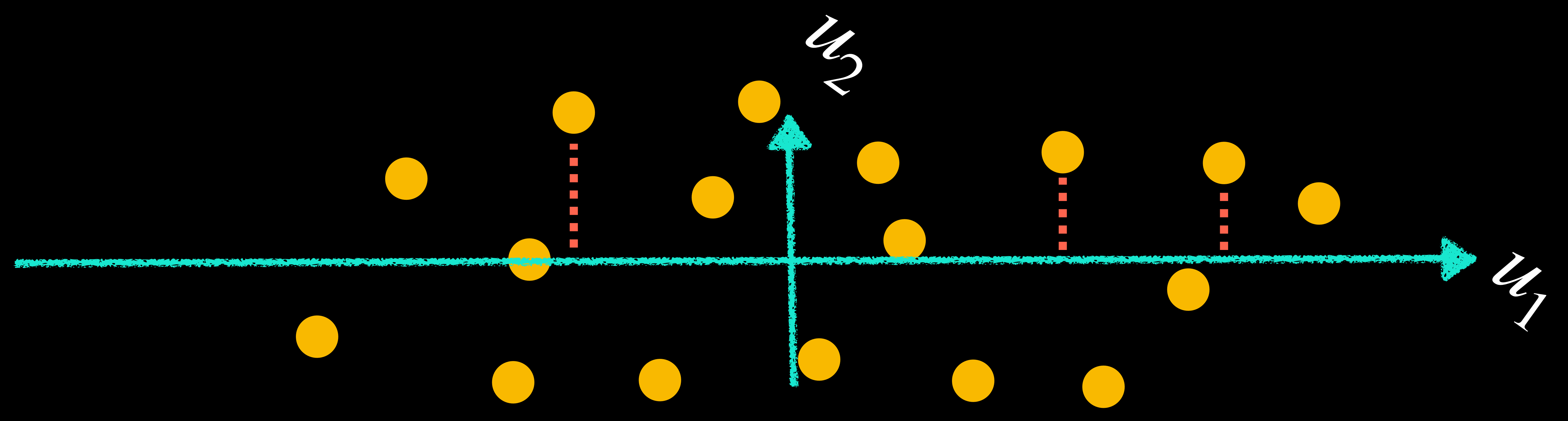


Principal Component Analysis

Error can be computed from distances in direction of u_2

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
…	…



How do we find u_1 and u_2?

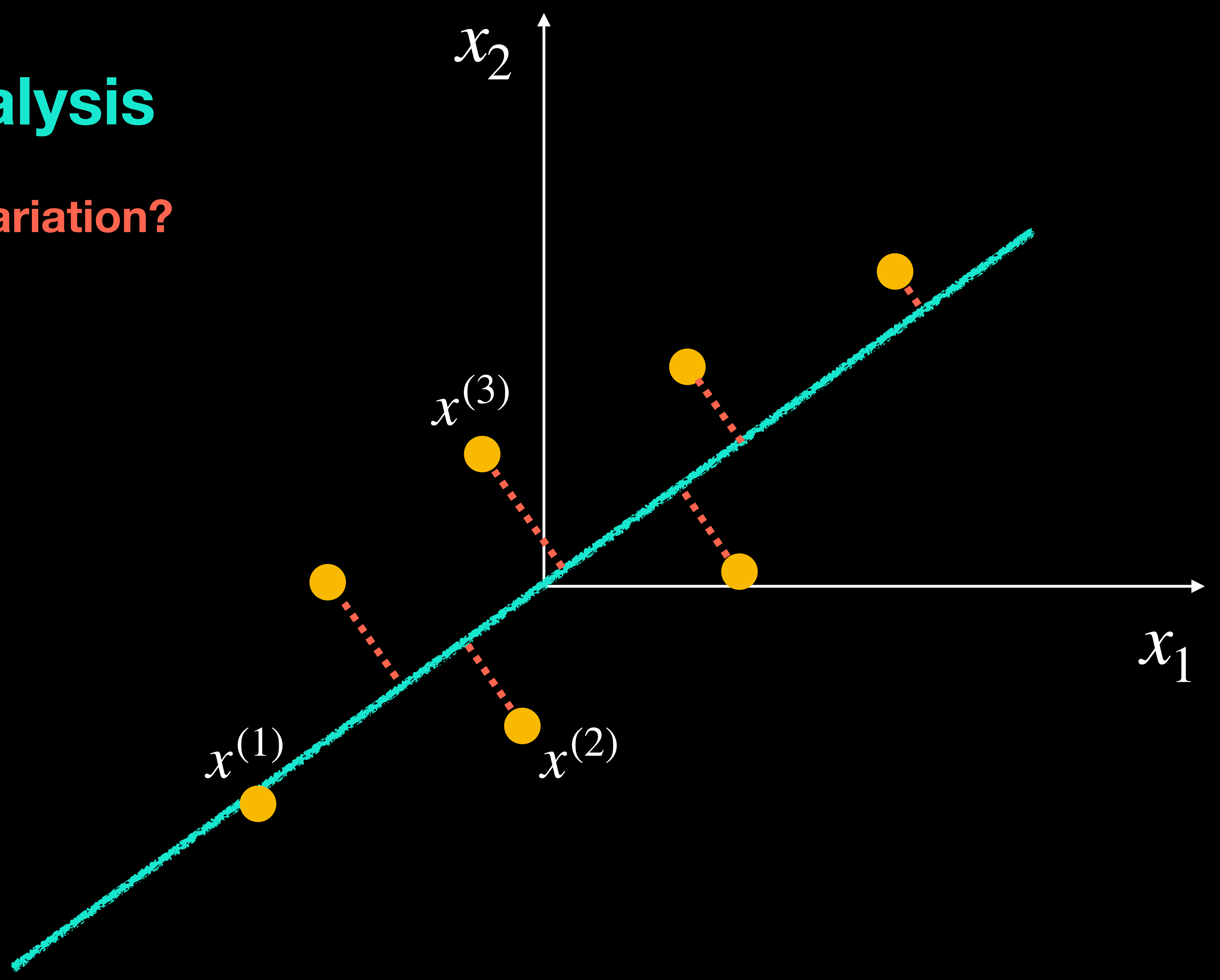
```
U, S, Vt = np.linalg.svd(data_centered)
```

Principal Component Analysis

What's the Major Axis (Direction) of Variation?

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

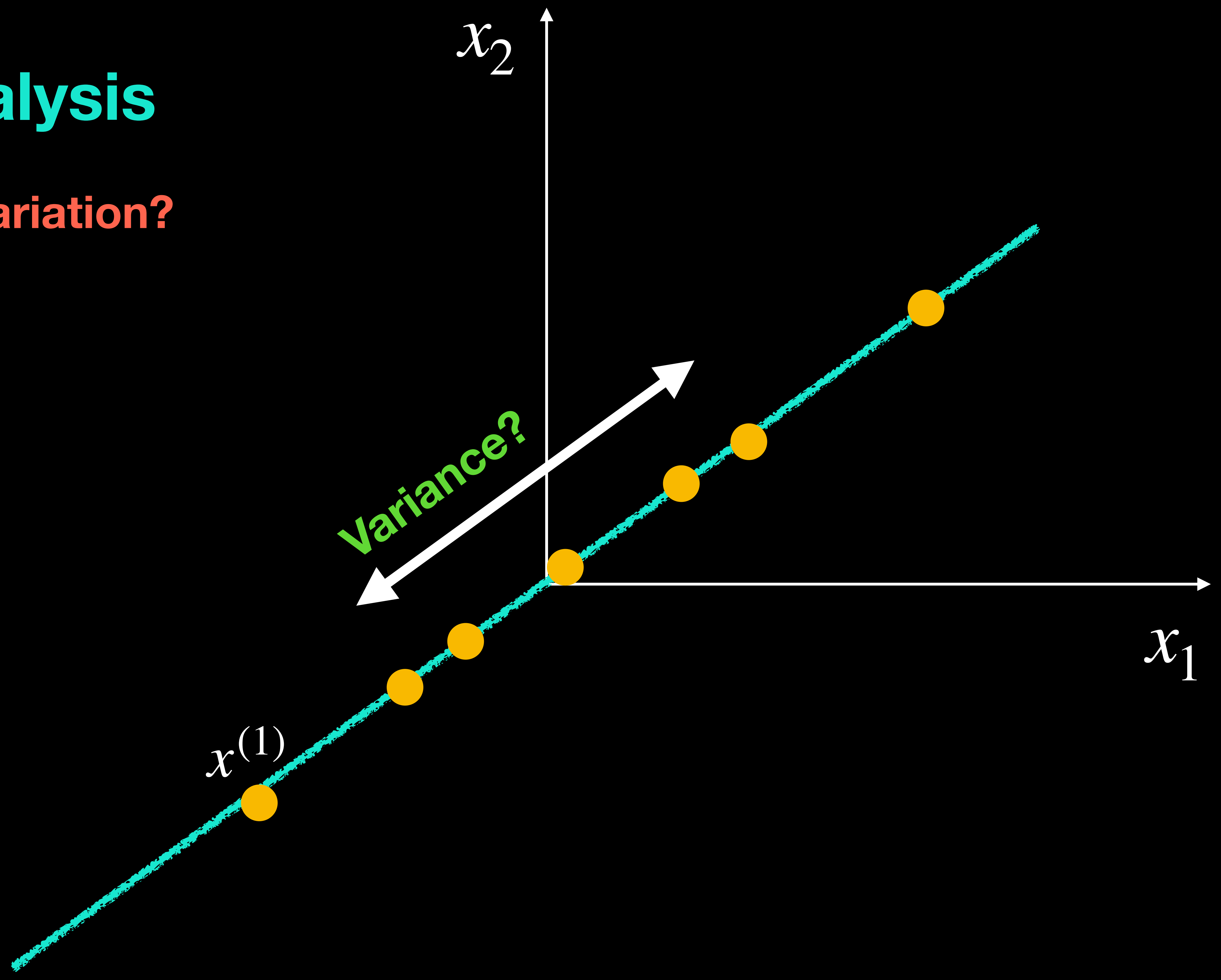


Principal Component Analysis

What's the Major Axis (Direction) of Variation?

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

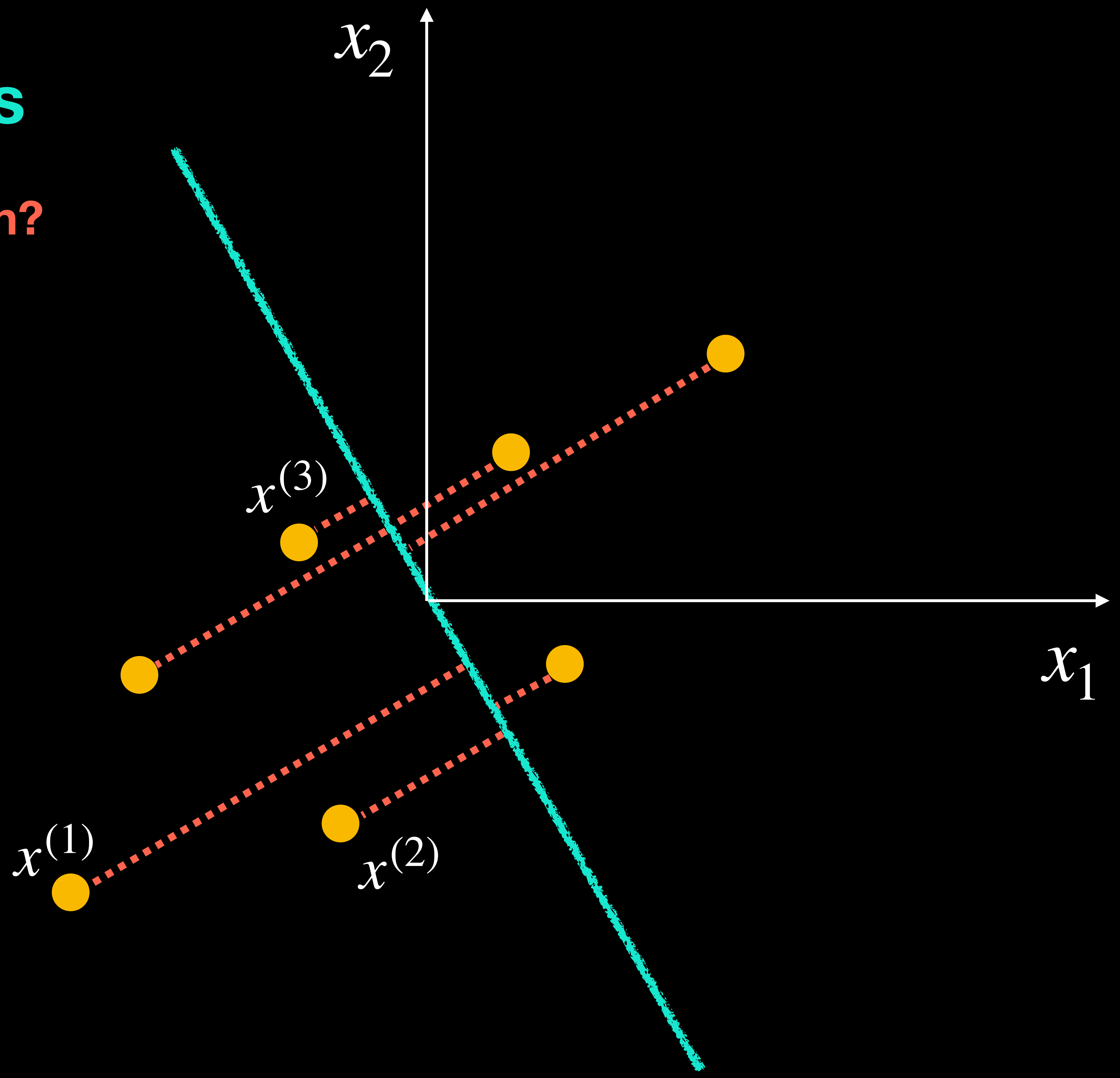


Principal Component Analysis

What's the Major Axis (Direction) of Variation?

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

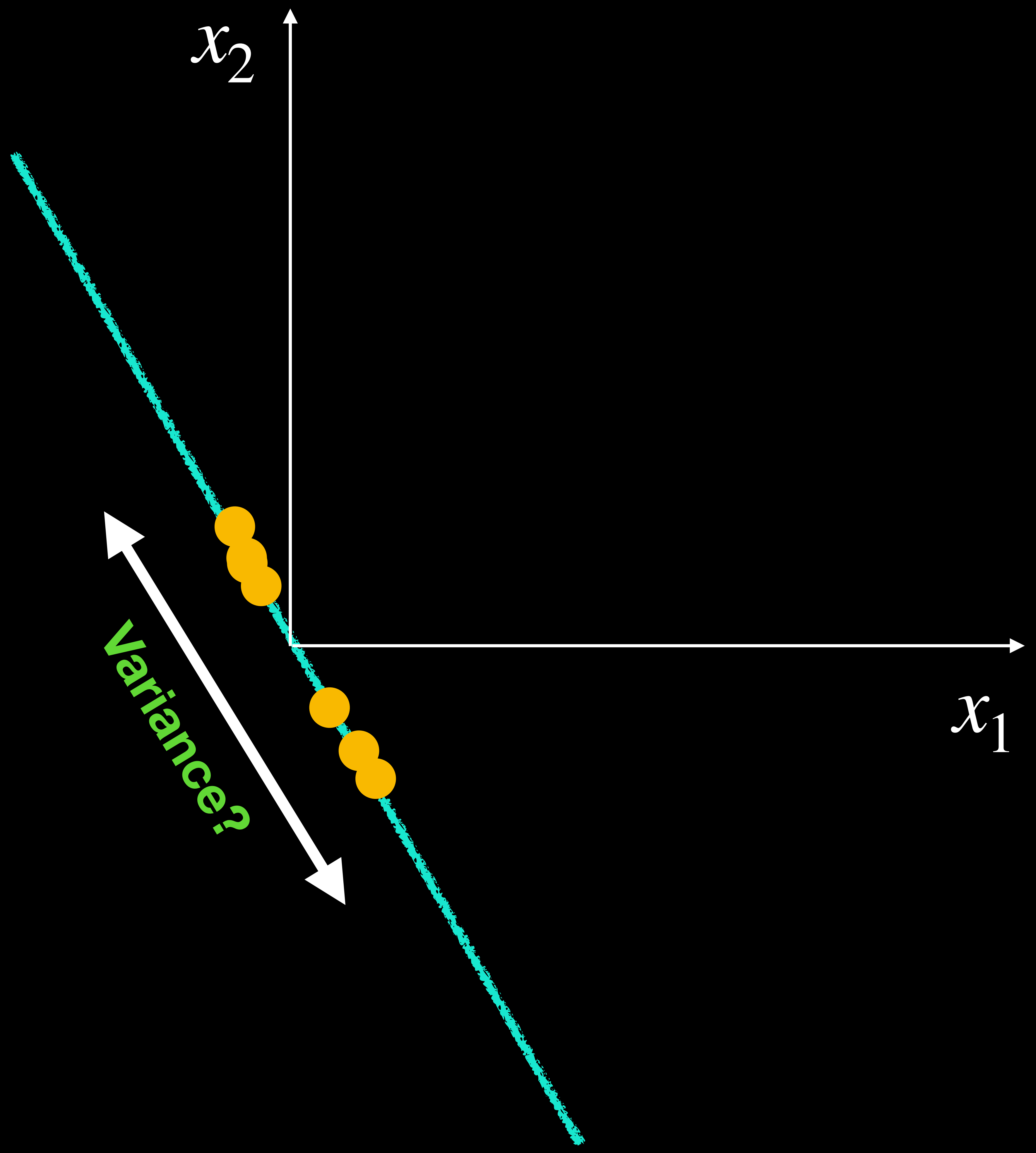


Principal Component Analysis

What's the Major Axis (Direction) of Variation?

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

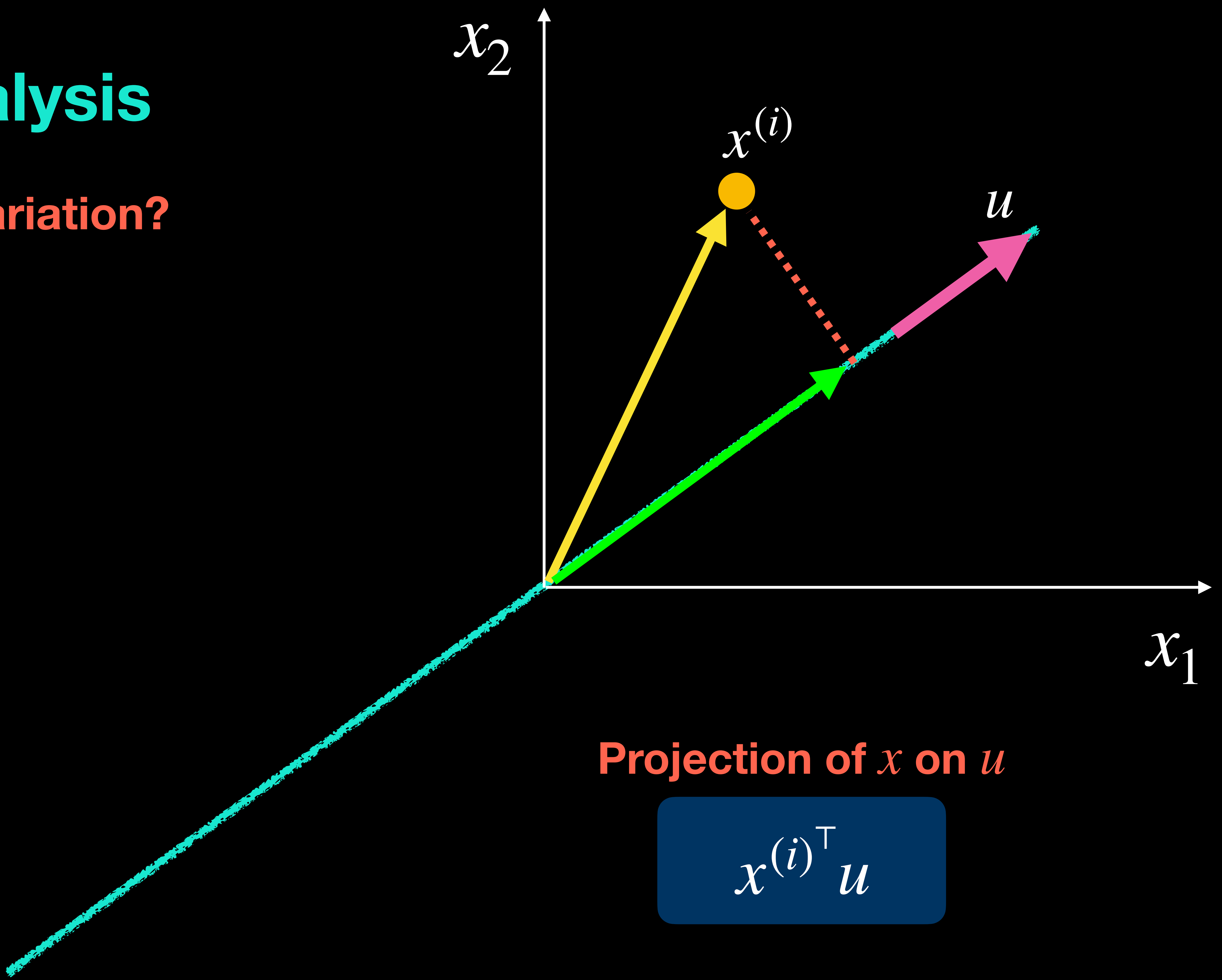


Principal Component Analysis

What's the Major Axis (Direction) of Variation?

Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
…	…



Projection of x on u

$$x^{(i)^\top} u$$

Principal Component Analysis

Maximize Variance of Projections

Maximize $\displaystyle \frac{1}{n}\sum_{i=1}^{n}\left(x^{(i)^\top}u\right)^2$

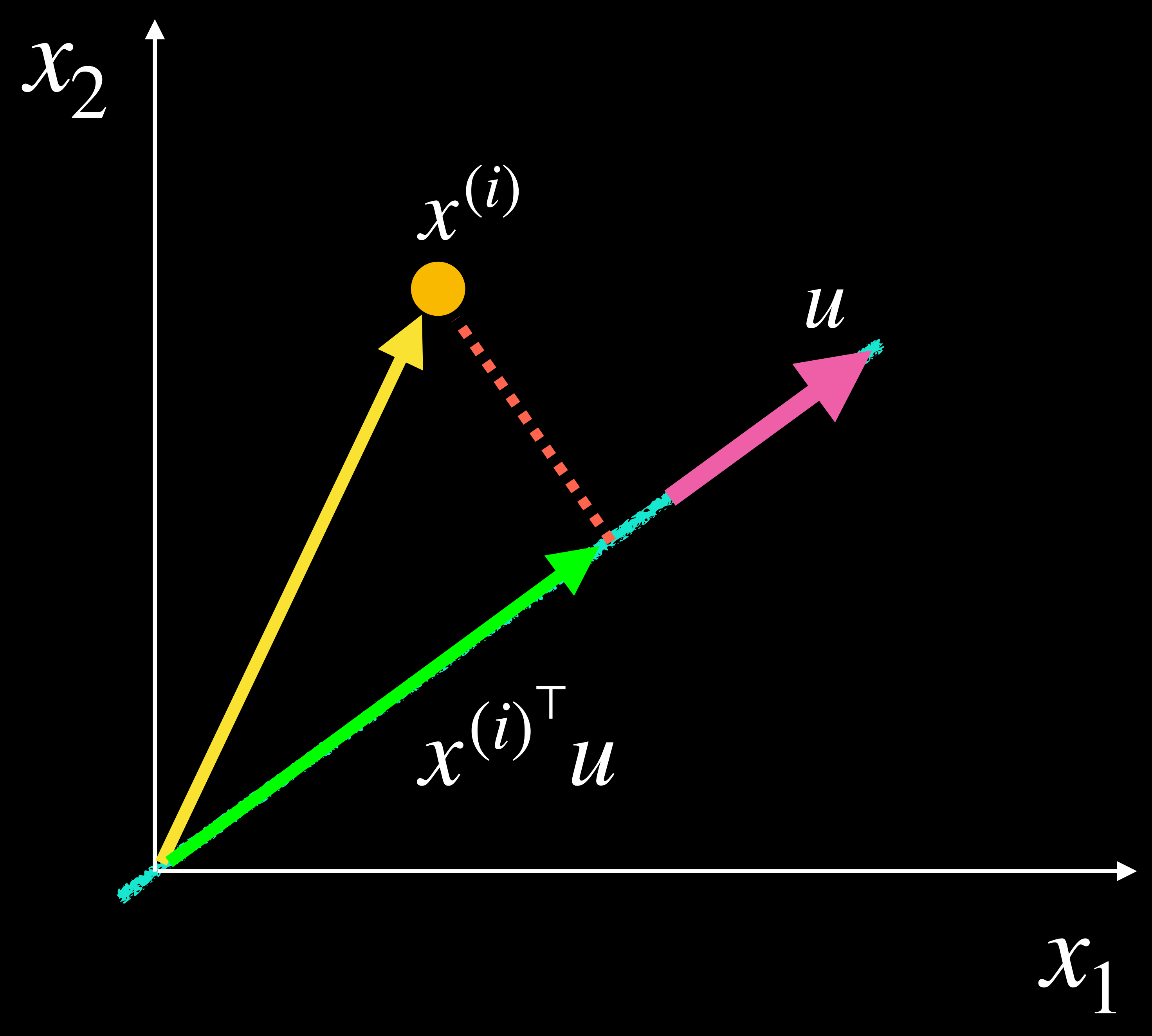
s.t. $\|u\|_2 = 1$

$$= \frac{1}{n}\sum_{i=1}^{n} u^T x^{(i)} x^{(i)T} u$$

$$= u^T\left(\frac{1}{n}\sum_{i=1}^{n} x^{(i)} x^{(i)T}\right) u\,.$$

Covariance Matrix

$$\max_{\|u\|=1} u^T \Sigma u$$



Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

Lagrange Multipliers

$$L(u, \lambda) = u^T \Sigma u - \lambda(u^T u - 1)$$

$$\frac{\partial L}{\partial u} = 2\Sigma u - 2\lambda u = 0$$

$$\Sigma u = \lambda u$$

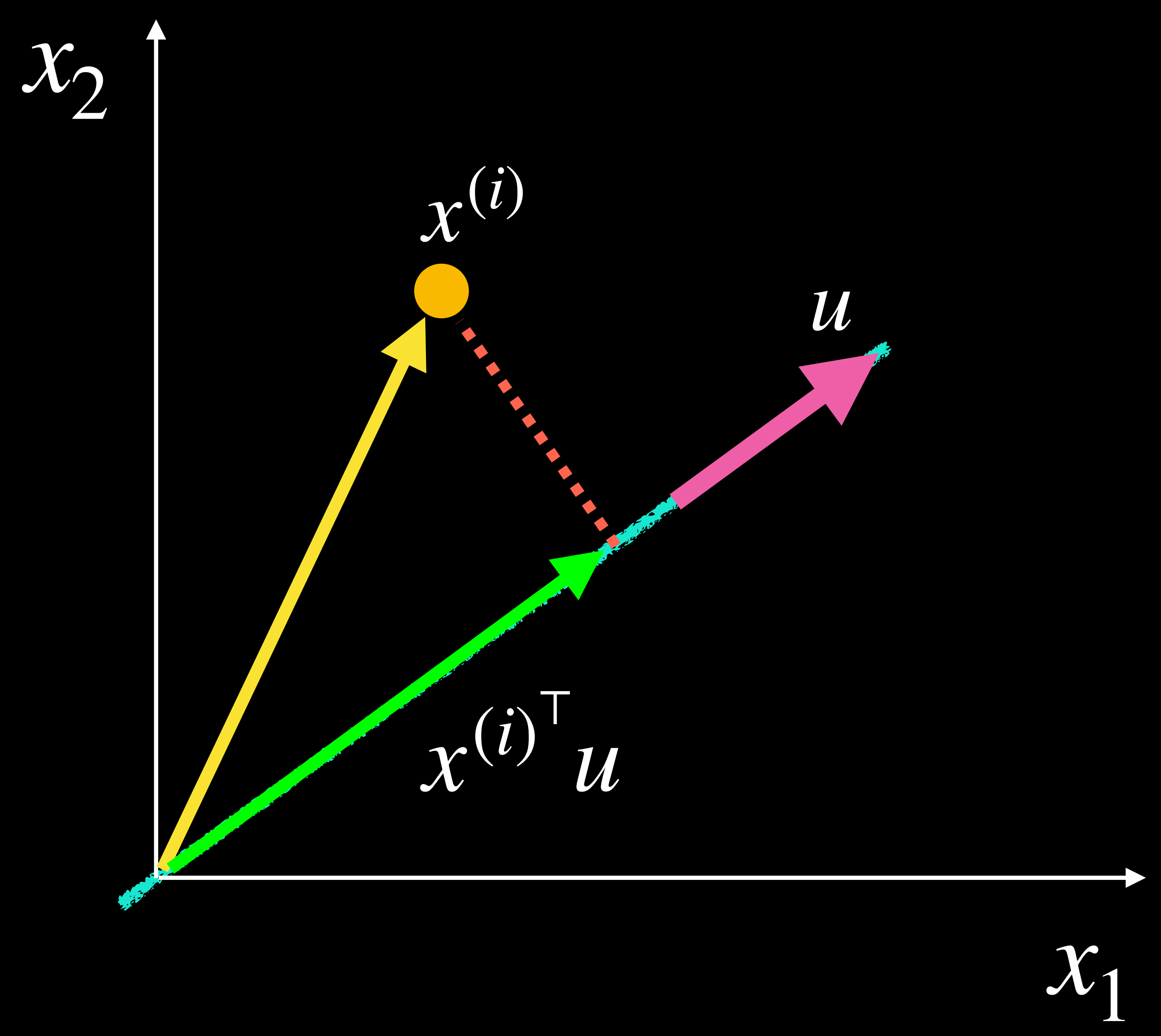
Principal Component Analysis

To Maximize Variance of Projections

Solve

$$\Sigma u = \lambda u$$

Derive The SVD decomposition of Σ



Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

Principal Component Analysis

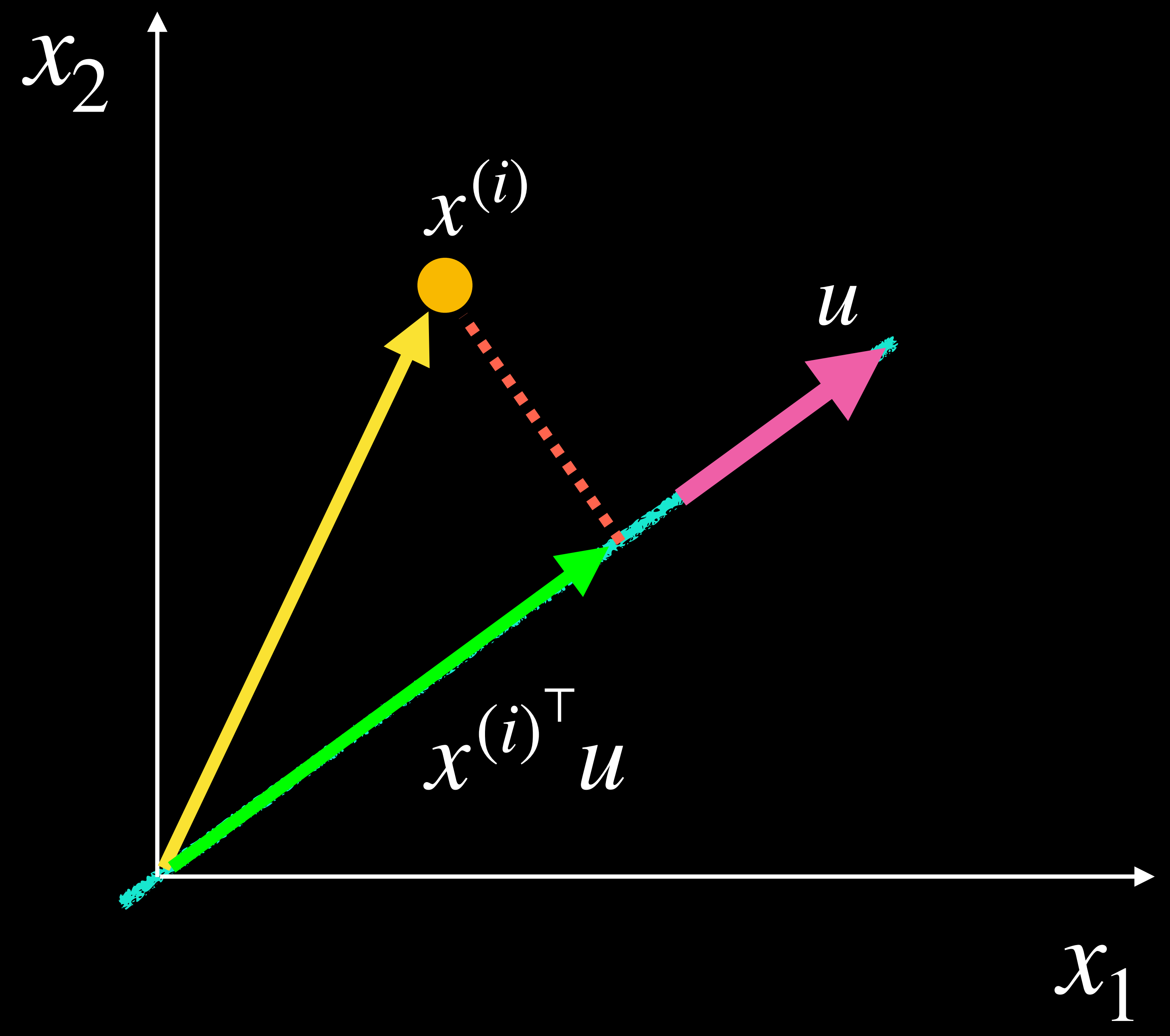
To Maximize Variance of Projections

Solve

$$\Sigma u = \lambda u$$

Given $x \in R^d$**,**
project on lower dimensions R^k

$$y^{(i)} = \begin{bmatrix} u_1^T x^{(i)} \\ u_2^T x^{(i)} \\ \vdots \\ u_k^T x^{(i)} \end{bmatrix} \in \mathbb{R}^k$$



Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...

Principal Component Analysis

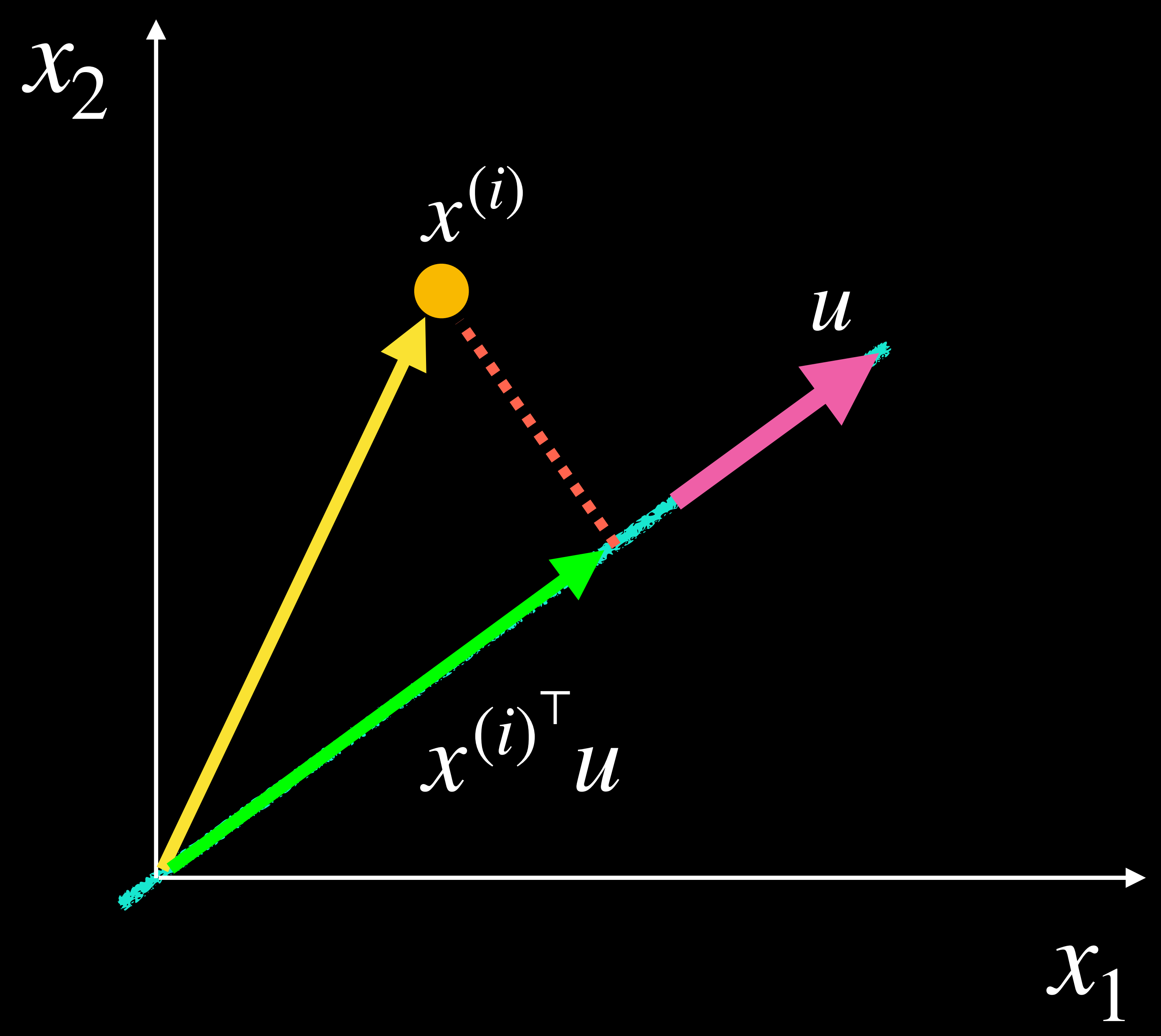
To Maximize Variance of Projections

Solve

$$\Sigma u = \lambda u$$

Percentage of Variance Preserved

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{d} \lambda_i} \times 100$$



Dataset

x_1	x_2
1.2	1.2
3.2	5.4
4.3	6.4
3.2	5.4
...	...