N / A / N / O / D / E / G / R / E / E

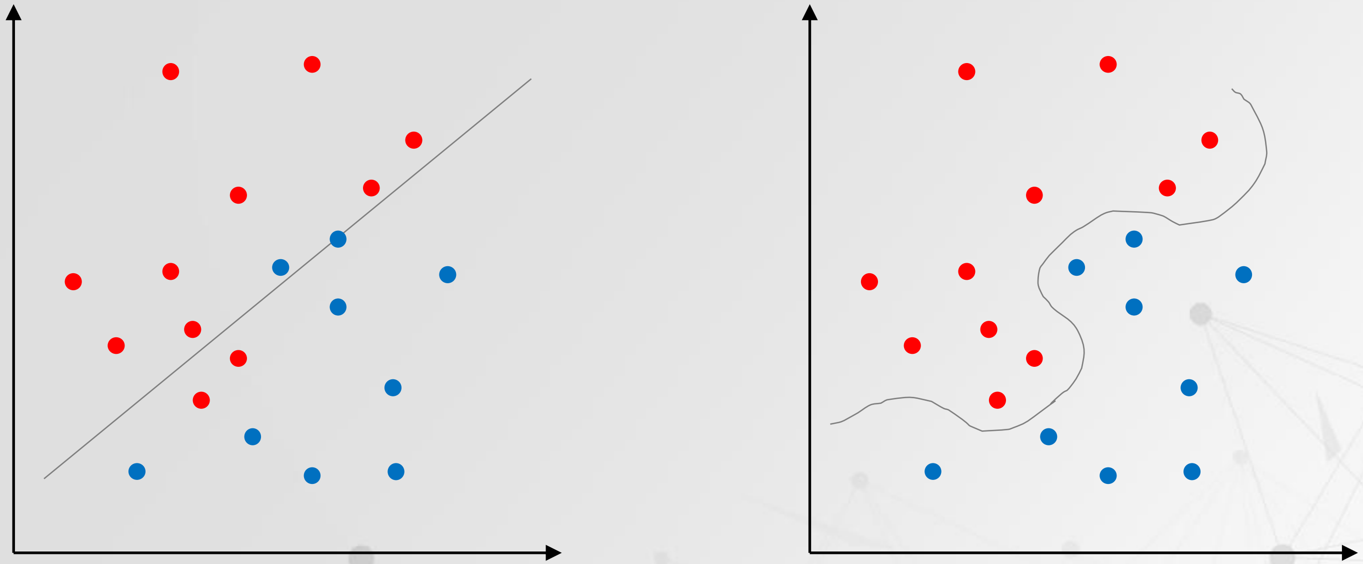
지도학습 알고리즘

11. Support Vector Machine IV

Kernel method and nonlinear SVM

Nonlinear SVM

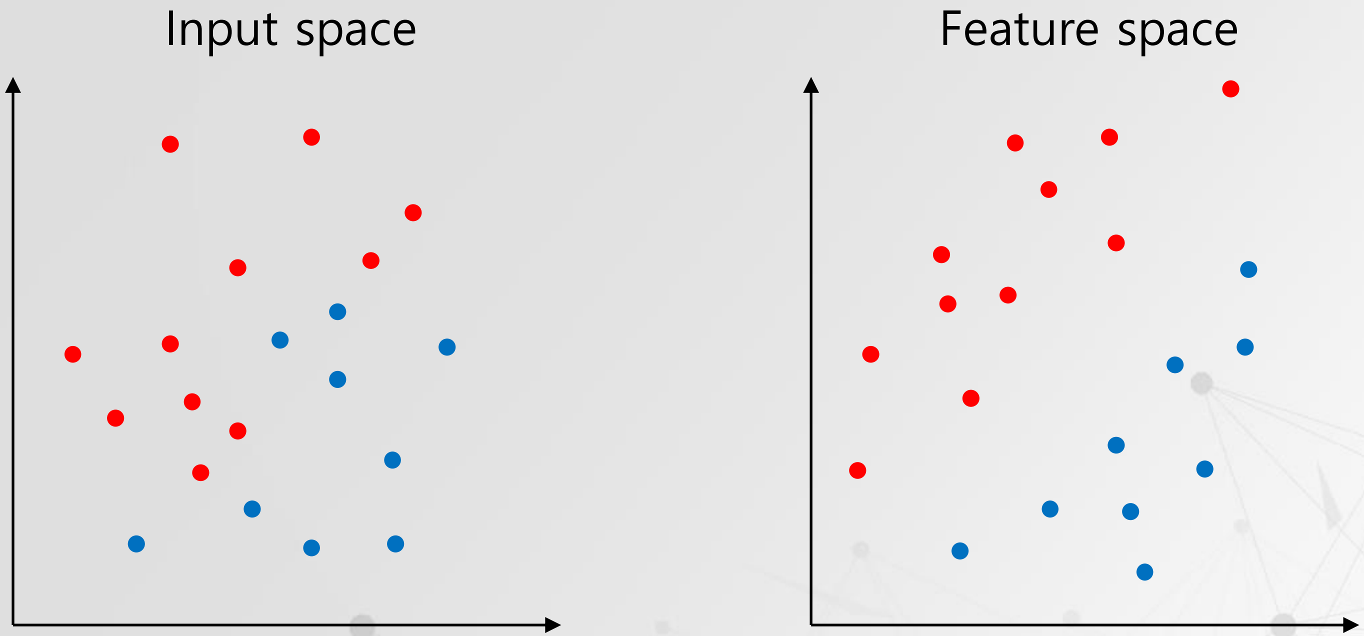
- Feature mapping: nonlinear and higher dimensional feature



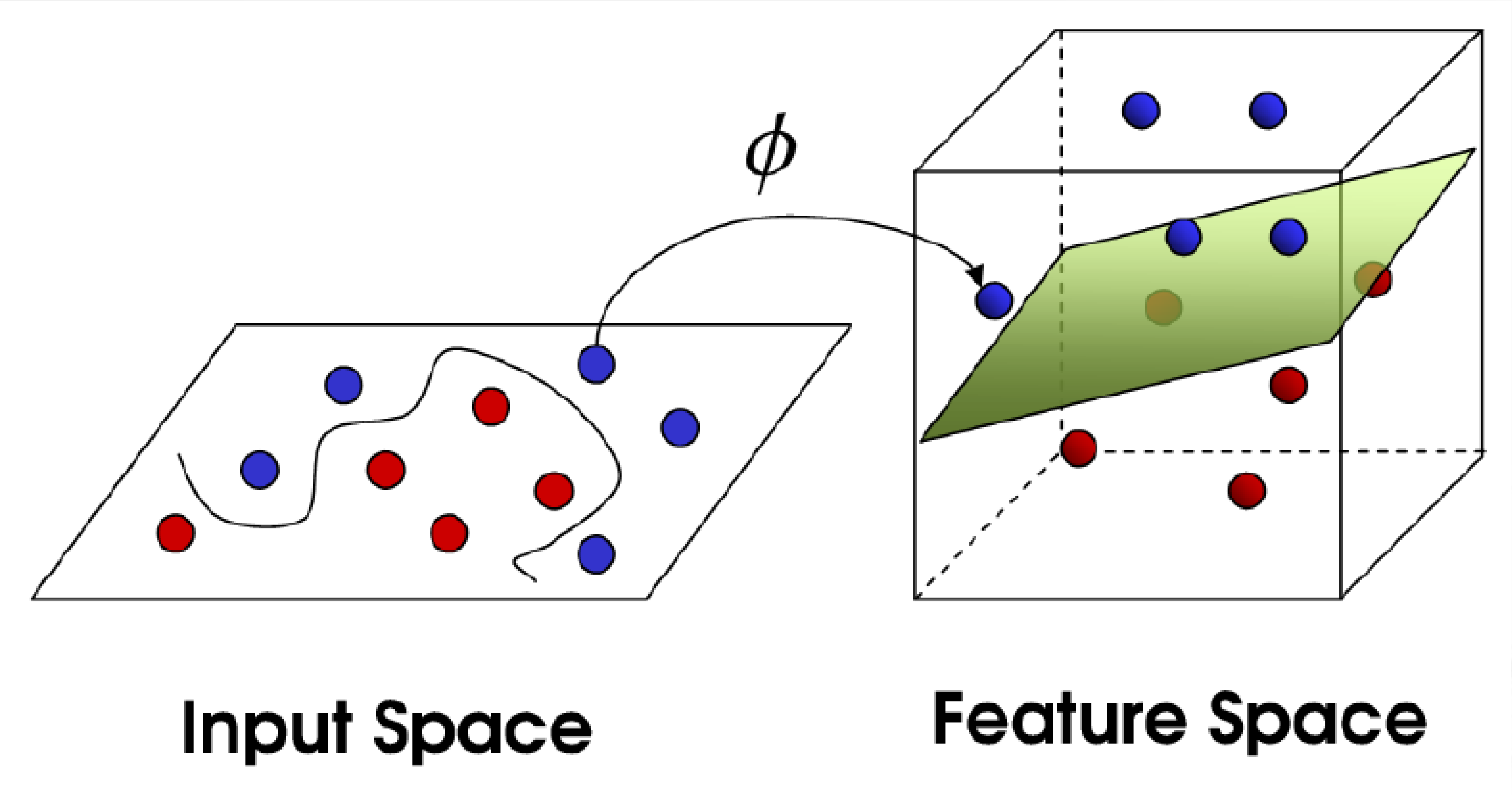
Feature mapping

- Feature mapping from input space to feature space

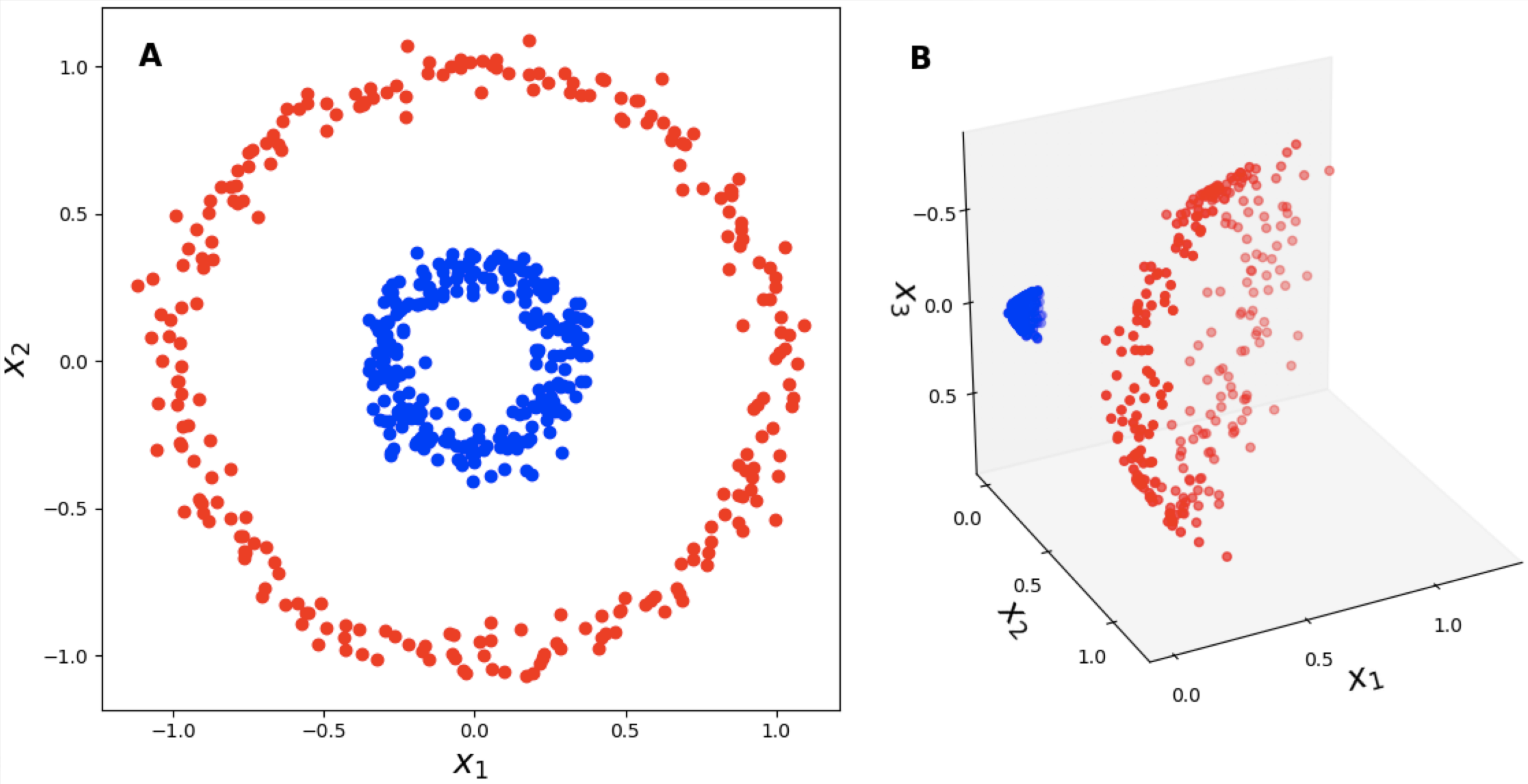
$$\phi : \Re^n \rightarrow \Re^m$$



Feature mapping



Feature mapping



Kernel function

- Kernel function은 어떤 feature space에서, feature간의 inner product로 정의됨

$$k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \phi(\mathbf{x}')$$

- Kernel function은 두 벡터 사이의 similarity로 해석할 수 있음
- 또한 feature mapping 이후에 inner product를 취하는 것보다, kernel 함수 자체로 계산하는 것이 훨씬 간편함

$$k(\mathbf{x}, \mathbf{x}') = (1 + \mathbf{x}^T \mathbf{x}')^2$$

$$\boldsymbol{\phi}(\mathbf{x}) = [\ 1 \quad x_1^2 \quad \sqrt{2}x_1x_2 \quad x_2^2 \quad \sqrt{2}x_1 \quad \sqrt{2}x_2\ ]$$

Kernel functions: exmaple

- Linear kernel

$$k(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}'$$

- Polynomial kernel

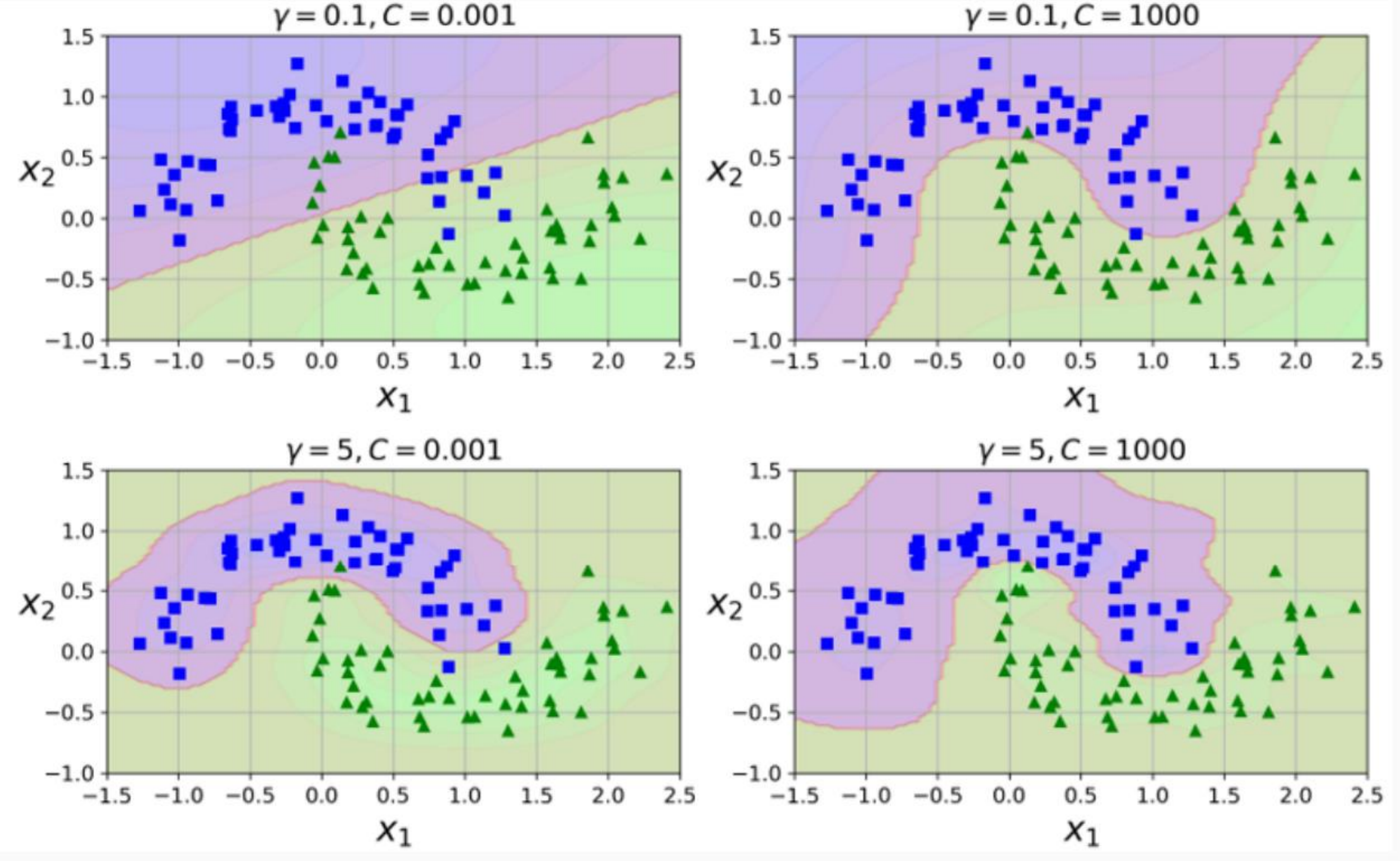
$$k(\mathbf{x}, \mathbf{x}') = (a\mathbf{x}^T \mathbf{x}' + b)^d$$

- Gaussian kernel (radial basis function kernel, RBF kernel)

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right) = \exp(-\gamma\|\mathbf{x} - \mathbf{x}'\|^2)$$

  - Kernel value는 x'가 landmark x와 가까울수록 큰 값
  - γ값으로 영향력의 범위를 조절할 수 있음
  - γ 가 클수록: 멀어질수록 kernel 값이 급격하게 줄어듦 = 영향력의 범위가 좁아짐
  - γ 가 작을수록: 멀어질수록 줄어드는 kernel 값의 폭이 작음 = 영향력의 범위가 큼

Soft SVM and Gaussian kernel



Construction of kernel

- 주어진 임의의 두 벡터에 대한 함수가 kernel function이 성립하게 하는 feature mapping이 존재하는지 검사하는 것은 까다로움

- Mercer's theorem
  Every positive semi-definite symmetric function is a kernel
  - 다음과 같이 정의되는 Gram matrix가 positive semi-definite symmetric matrix면, 주어진 함수 k는 kernel function임

$$\mathbf{K} = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & \cdots & k(\mathbf{x}_1, \mathbf{x}_N) \\ k(\mathbf{x}_2, \mathbf{x}_1) & k(\mathbf{x}_2, \mathbf{x}_2) & \cdots & k(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{x}_N, \mathbf{x}_1) & k(\mathbf{x}_N, \mathbf{x}_2) & \cdots & k(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix}$$

Positive semi-definite matrix

Definition

$$\mathbf{K} \in \Re^{n\times n} \text{ is positive semi-definite} \Longleftrightarrow \mathbf{x}^T\mathbf{K}\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \in \Re^n$$

Eigenvalues

- Positive semi-definite if and only if all of its eigenvalues are non-negative