Master of Technology

U2/6: Computational Intelligence I

To clarify slide 25

Slide 25 in lecture note

Input space $\vec{x_i} = (x_{i1}, x_{i2})$ Feature space $\vec{X_i} = (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2})$

Inner product $\vec{X_i} \cdot \vec{X_j} = X_{i1}X_{j1} + X_{i2}X_{j2} + X_{i3}X_{j3}$

Without kernel trick

To get the new first dimension: 1 multiplication
Second dimension: 1 multiplication
Third dimension: 2 multiplications
In all, 1+1+2 = 4 multiplications.
Multiplications: 8 (for the projections) + 3 (in the dot product) = 11 multiplications
Additions: 2 (in the dot product)
Total: 11 + 2 = 13 operations.

With kernel trick

$$\begin{aligned} K(\vec{x_i}, \vec{x_j}) &= (\vec{x_i} \cdot \vec{x_j})^2 \\ &= (x_{i1}x_{j1} + x_{i2}x_{j2})^2 \\ &= x_{i1}^2 x_{j1}^2 + x_{i2}^2 x_{j2}^2 + 2x_{i1}x_{i2}x_{j1}x_{j2} \\ &= (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}) \cdot (x_{j1}^2, x_{j2}^2, \sqrt{2}x_{j1}x_{j2}) \end{aligned}$$

Multiplications: 2 (for the dot product in the original space) + 1 (for squaring the result) = 3 multiplications
Additions: 1 (for the dot product in the original space)
Total: 3 + 1 = 4 operations.

Source: https://blog.statsbot.co/support-vector-machines-tutorial-c1618e635e93

Recall: trained SVM model (Slide 22)

Decision plane $d(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$

Provided in sample data

Parameters associated with input vectors in sample data, $\alpha_i \neq 0$ for support vectors

$$d(\mathbf{x}) = \sum_{i=1}^{l} y_i \alpha_i \langle \phi(\mathbf{x}), \phi(\mathbf{x}_i) \rangle + b$$

All the information the learning algorithm needs is the **inner products** between data points in the feature space, where $\mathbf{x}$, $\mathbf{x}_i$ $(i = 1, \ldots, l) \in \mathbf{X}$, the input space

Dot product in (old) 2-D input space

Input space $\vec{x_i} = (x_{i1}, x_{i2})$

$$K(\vec{x_i}, \vec{x_j}) = (\vec{x_i} \cdot \vec{x_j})$$
$$= (x_{i1}x_{j1} + x_{i2}x_{j2})$$

Dot product in (new) 3-D feature space

Input space $\vec{x_i} = (x_{i1}, x_{i2})$ Feature space $\vec{X_i} = (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2})$

$$
\begin{aligned}
K(\vec{x_i}, \vec{x_j}) &= (\vec{x_i} \cdot \vec{x_j})^2 \\
&= (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}) \cdot (x_{j1}^2, x_{j2}^2, \sqrt{2}x_{j1}x_{j2}) \\
&= x_{i1}^2 x_{j1}^2 + x_{i2}^2 x_{j2}^2 + 2x_{i1}x_{i2}x_{j1}x_{j2}
\end{aligned}
$$

Solution 1: Transform data, calculate dot product

Input space $\vec{x_i} = (x_{i1}, x_{i2})$ Feature space $\vec{X_i} = (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2})$

Inner product $\vec{X_i} \cdot \vec{X_j} = X_{i1}X_{j1} + X_{i2}X_{j2} + X_{i3}X_{j3}$

To map EACH data point from old space to new space
To get the new first dimension: 1 multiplication
Second dimension: 1 multiplication
Third dimension: 2 multiplications
In all, 1+1+2 = 4 multiplications.

To calculate the dot product in (new) feature space
Multiplications: 8 (for the projections) + 3 (in the dot product) = 11 multiplications
Additions: 2 (in the dot product)
Total: 11 + 2 = 13 operations.

Solution 2: Trick

Dot product in new space

$$(x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}) \cdot (x_{j1}^2, x_{j2}^2, \sqrt{2}x_{j1}x_{j2})$$

$$= x_{i1}^2 x_{j1}^2 + x_{i2}^2 x_{j2}^2 + 2x_{i1}x_{i2}x_{j1}x_{j2} \qquad = (x_{i1}x_{j1} + x_{i2}x_{j2})^2$$

Dot product in old space

Solution 2: Transform dot product

Input space $\vec{x_i} = (x_{i1}, x_{i2})$ Feature space $\vec{X_i} = (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2})$

Inner product $\vec{X_i} \cdot \vec{X_j} = X_{i1}X_{j1} + X_{i2}X_{j2} + X_{i3}X_{j3}$

$$(x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}) \cdot (x_{j1}^2, x_{j2}^2, \sqrt{2}x_{j1}x_{j2})$$

$$= x_{i1}^2 x_{j1}^2 + x_{i2}^2 x_{j2}^2 + 2x_{i1}x_{i2}x_{j1}x_{j2} \qquad = (x_{i1}x_{j1} + x_{i2}x_{j2})^2$$

- 2 multiplications (for the dot product in the original space)
- 1 addition (for the dot product in the original space)
- 1 (for squaring the result)

Total: 3 + 1 = 4 operations.

Source: https://blog.statsbot.co/support-vector-machines-tutorial-c1618e635e93

Comparison between two solutions

Input space $\vec{x_i} = (x_{i1}, x_{i2})$ Feature space $\vec{X_i} = (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2})$

Inner product $\vec{X_i} \cdot \vec{X_j} = X_{i1}X_{j1} + X_{i2}X_{j2} + X_{i3}X_{j3}$

Solution 1

To get the new first dimension: 1 multiplication
Second dimension: 1 multiplication
Third dimension: 2 multiplications
In all, 1+1+2 = 4 multiplications.
Multiplications: 8 (for the projections) + 3 (in the dot product) = 11 multiplications
Additions: 2 (in the dot product)
Total: 11 + 2 = 13 operations.

Solution 2

$$\begin{aligned} K(\vec{x_i}, \vec{x_j}) &= (\vec{x_i} \cdot \vec{x_j})^2 \\ &= (x_{i1}x_{j1} + x_{i2}x_{j2})^2 \\ &= x_{i1}^2 x_{j1}^2 + x_{i2}^2 x_{j2}^2 + 2x_{i1}x_{i2}x_{j1}x_{j2} \\ &= (x_{i1}^2, x_{i2}^2, \sqrt{2}x_{i1}x_{i2}) \cdot (x_{j1}^2, x_{j2}^2, \sqrt{2}x_{j1}x_{j2}) \end{aligned}$$

Multiplications: 2 (for the dot product in the original space) + 1 (for squaring the result) = 3 multiplications
Additions: 1 (for the dot product in the original space)
Total: 3 + 1 = 4 operations.

Source: https://blog.statsbot.co/support-vector-machines-tutorial-c1618e635e93

Examples of Kernel functions

- Polynomial kernel with degree *d*

$$K(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y} + 1)^d$$

- Radial basis function kernel with width σ

$$K(\mathbf{x}, \mathbf{y}) = \exp(-||\mathbf{x} - \mathbf{y}||^2/(2\sigma^2))$$