$$(AB)_{ij} = \sum_{n=1}^{N} A_{in} B_{nj}, \tag{B.37}$$

Transpose

and the *transpose* $(A^\top)_{ij} := A_{ji}$.

Inverse and Pseudo-Inverse

The *inverse* of a matrix A is written A^{-1} and satisfies $AA^{-1} = A^{-1}A = \mathbf{1}$. The *pseudo-inverse* $A^\dagger$ satisfies $AA^\dagger A = A$. While every matrix has a pseudo-inverse, not all have an inverse. Those which do are called *invertible* or *nonsingular*, and their inverse coincides with the pseudo-inverse. Sometimes, we simply use the notation A^{-1}, and it is understood that we mean the pseudo-inverse whenever A is not invertible.

B.2.2 Norms and Dot Products

Thus far, we have explained the linear structure of spaces such as the feature space induced by a kernel. We now move on to the *metric* structure. To this end, we introduce concepts of length and angles.

Definition B.5 (Norm) *A function* $\|\cdot\| : \mathcal{H} \to \mathbb{R}_0^+$ *that for all* $\mathbf{x}, \mathbf{x}' \in \mathcal{H}$ *and* $\lambda \in \mathbb{R}$ *satisfies*

$$\|\mathbf{x} + \mathbf{x}'\| \leq \|\mathbf{x}\| + \|\mathbf{x}'\|, \tag{B.38}$$

$$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|, \tag{B.39}$$

$$\|\mathbf{x}\| > 0 \text{ if } \mathbf{x} \neq 0, \tag{B.40}$$

Norm

is called a norm on $\mathcal{H}$. *If we replace the "$>$" in (B.40) by "$\geq$," we are left with what is called a* semi-norm.

Metric

Any norm defines a *metric* d via

$$d(\mathbf{x}, \mathbf{x}') := \|\mathbf{x} - \mathbf{x}'\|; \tag{B.41}$$

likewise, any semi-norm defines a *semi-metric*. The (semi-)metric inherits certain properties from the (semi-)norm, in particular the triangle inequality (B.38) and positivity (B.40).

While every norm gives rise to a metric, the converse is not the case. In this sense, the concept of the norm is stronger. Similarly, every *dot product* (to be introduced next) gives rise to a norm, but not vice versa.

Before describing the dot product, we start with a more general concept.

Definition B.6 (Bilinear Form) *A* bilinear form *on a vector space* $\mathcal{H}$ *is a function*

$$Q : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$$

$$(\mathbf{x}, \mathbf{x}') \to Q(\mathbf{x}, \mathbf{x}') \tag{B.42}$$

with the property that for all $\mathbf{x}, \mathbf{x}', \mathbf{x}'' \in \mathcal{H}$ *and all* $\lambda, \lambda' \in \mathbb{R}$*, we have*

$$Q((\lambda\mathbf{x} + \lambda'\mathbf{x}'), \mathbf{x}'') = \lambda Q(\mathbf{x}, \mathbf{x}'') + \lambda' Q(\mathbf{x}', \mathbf{x}''), \tag{B.43}$$

$$Q(\mathbf{x}'', (\lambda\mathbf{x} + \lambda'\mathbf{x}')) = \lambda Q(\mathbf{x}'', \mathbf{x}) + \lambda' Q(\mathbf{x}'', \mathbf{x}'). \tag{B.44}$$