projection

Authors

Alexander Jung[*], Department of Computer Science, Aalto University, Espoo, Finland
[*]Corresponding author: alex.jung@aalto.fi

Abstract

The projection of a vector onto a closed non-empty subset of a Euclidean space is a point in the subset that is closest to the vector in the Euclidean norm. If the subset is convex, this closest point is unique. If the subset is a subspace, the projection is a linear map, namely the orthogonal projection onto the subspace. In machine learning (ML), projections enforce constraints on model parameters during training: projected gradient descent (projected GD) alternates gradient steps with projections onto the constraint set, e.g., the ℓ_1-ball in an equivalent formulation of the least absolute shrinkage and selection operator (Lasso).

Definition

Consider a closed non-empty subset $\mathcal{W} \subseteq \mathbb{R}^d$ of the d-dimensional Euclidean space. The projection $P_{\mathcal{W}}(\mathbf{w})$ of a vector $\mathbf{w} \in \mathbb{R}^d$ onto $\mathcal{W}$ is defined as

$$P_{\mathcal{W}}(\mathbf{w}) = \underset{\mathbf{w}' \in \mathcal{W}}{\arg\min} \|\mathbf{w} - \mathbf{w}'\|_2.$$

In other words, $P_{\mathcal{W}}(\mathbf{w})$ is a vector in $\mathcal{W}$ that is closest to $\mathbf{w}$. The above minimum exists for every closed non-empty subset $\mathcal{W}$; if $\mathcal{W}$ is also convex, the closest point is unique (Boyd and Vandenberghe, 2004). If $\mathcal{W}$ is a subspace, the map $\mathbf{w} \mapsto P_{\mathcal{W}}(\mathbf{w})$ is linear: it is the orthogonal projection onto $\mathcal{W}$.

In machine learning (ML), projections are used to enforce constraints on model parameters during training. As a case in point, consider linear regression with the constraint $\|\mathbf{w}\|_1 \leq \tau$, which is an equivalent formulation of the least absolute shrinkage and selection operator (Lasso) (Tibshirani, 1996). This constrained problem can be solved by projected gradient descent (projected GD): each iteration takes a gradient step and then projects the result onto the ℓ_1-ball $\mathcal{W} = \{\mathbf{w}' \in \mathbb{R}^d : \|\mathbf{w}'\|_1 \leq \tau\}$. Since the ℓ_1-ball is not a subspace, this projection is not a linear map (see Fig. 1).

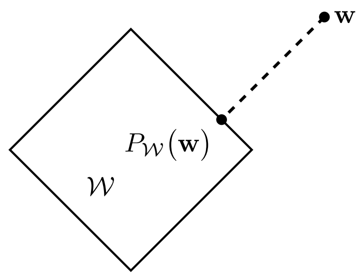

Figure 1: Projection of a vector $\mathbf{w} \in \mathbb{R}^2$ onto the ℓ_1-ball $\mathcal{W} = \{\mathbf{w}' \in \mathbb{R}^2 : \|\mathbf{w}'\|_1 \leq \tau\}$. The projection $P_{\mathcal{W}}(\mathbf{w})$ is the point of $\mathcal{W}$ closest to $\mathbf{w}$ in the Euclidean norm

Cross References

Euclidean space, vector, minimum, convex, orthogonal projection, projected GD

References

1. Boyd and Vandenberghe (2004). Convex Optimization. Cambridge Univ. Press, Cambridge, U.K.

2. Tibshirani (1996). Regression Shrinkage and Selection via the Lasso. J. Roy. Statist. Soc.: Ser. B (Methodological), 58(1), pp. 267–288.