

inverse

Author and editors

Alexander Jung*, Department of Computer Science, Aalto University, Espoo, Finland (ORCID 0000-0001-7538-0990)

Editors: Konstantina Oliumtsevit (ORCID 0000-0003-2697-8659), Ekkehard Schnoor (ORCID 0000-0001-8252-1499), Aalto University, Espoo, Finland

*Corresponding author: alex.jung@aalto.fi

Abstract

An inverse of a function is a second function that undoes it. For $f : \mathcal{U} \rightarrow \mathcal{V}$ it is the function $g : \mathcal{V} \rightarrow \mathcal{U}$ with $g(f(u)) = u$ and $f(g(v)) = v$, and it exists precisely when f is injective and every element of $\mathcal{V}$ has a nonempty preimage. The inverse of a linear map of $\mathbb{R}^n$ is again a linear map, represented by the inverse matrix $\mathbf{A}^{-1}$ with $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$, which exists exactly when the square matrix $\mathbf{A}$ is of full-rank, equivalently when its determinant is nonzero. Where no inverse exists, a left or right inverse or the Moore–Penrose pseudoinverse takes its place. In machine learning (ML), inverses appear in the closed-form solution of linear regression, in the pair formed by the sigmoid function and the logit, which turns a score into a probability and back, and in drawing a realization through the inverse of a cumulative distribution function (cdf).

Definition

A frost warning for Krems an der Donau delivers a real-valued score for the day, and the sigmoid function turns that score into the probability of frost: the score 0 becomes the probability 0.5, the score 2.2 becomes 0.9. Reading the score back off a reported probability undoes this step, and the function that does so is the logit $\ln(p/(1-p))$. The logit is the inverse of the sigmoid function.

An inverse of a function $f : \mathcal{U} \rightarrow \mathcal{V}$ is a function $g : \mathcal{V} \rightarrow \mathcal{U}$ that undoes it in both directions: $g(f(u)) = u$ for every $u \in \mathcal{U}$, and $f(g(v)) = v$ for every $v \in \mathcal{V}$. Such a g exists precisely when f is injective and every $v \in \mathcal{V}$ has a nonempty preimage. It is then unique and written f^{-1} . If only the first identity holds, g is a left inverse, and if only the second holds, a right inverse. Fig. 1 shows both cases for a function on a domain of four shapes.

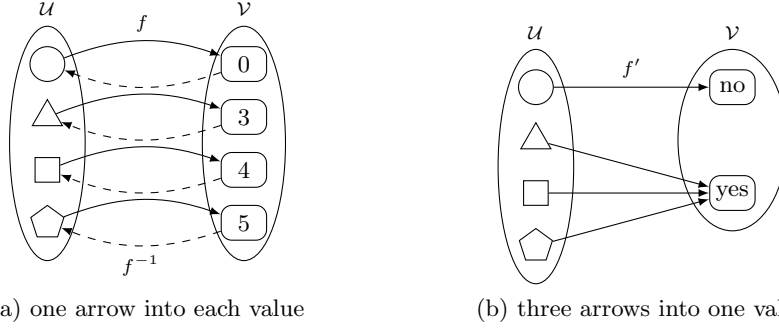


Figure 1: Two functions on the same domain $\mathcal{U}$ of four shapes, with $\mathcal{U}$ and the set $\mathcal{V}$ of values each drawn as a thin ellipse. In (a) the function f sends a shape to its number of corners, and every value of $\mathcal{V} = \{0, 3, 4, 5\}$ receives exactly one solid arrow, so following the arrows backwards defines the inverse f^{-1} (dashed arrows), which returns the shape with that number of corners. In (b) the function f' records only whether a shape has a straight edge. Three shapes share the value yes, so no function undoes f' : the value yes does not say which shape produced it

The best known instance is the inverse of a linear map. A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ represents a linear map of $\mathbb{R}^n$ into itself, and that map has an inverse exactly when $\mathbf{A}$ is of full-rank, meaning its columns are linearly independent. Such a matrix is said to be invertible, and the matrix representing the inverse linear map is the inverse matrix $\mathbf{A}^{-1}$, which satisfies

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

A square matrix is invertible if and only if its determinant is nonzero (Horn and Johnson, 1991; G. Strang, 2007). Fig. 2 shows the two linear maps acting on the standard basis of $\mathbb{R}^2$.

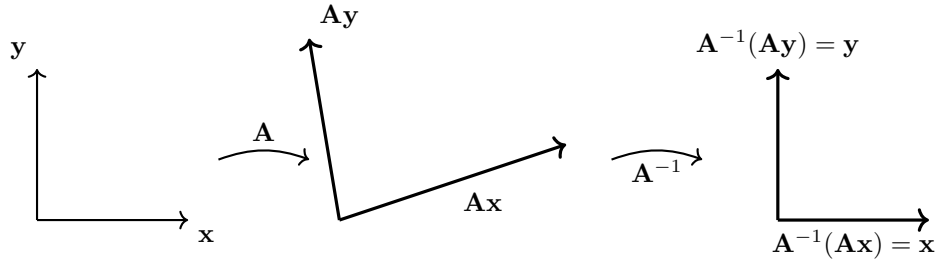


Figure 2: Matrix $\mathbf{A}$ representing a linear transformation of $\mathbb{R}^2$; inverse matrix $\mathbf{A}^{-1}$ representing the inverse transformation

For a matrix that is not square or not of full-rank, the one-sided inverses are what remains: a left inverse $\mathbf{B}$ satisfies $\mathbf{B}\mathbf{A} = \mathbf{I}$, and a right inverse $\mathbf{C}$ satisfies

$\mathbf{A}\mathbf{C} = \mathbf{I}$. The Moore–Penrose pseudoinverse $\mathbf{A}^+$ selects one generalized inverse for every matrix (Golub and Van Loan, 2013).

Inverses are used throughout machine learning (ML). Solving a system of linear equations is one task, and ridge regression is another: its closed-form solution inverts the matrix $\mathbf{X}^\top \mathbf{X} + \alpha \mathbf{I}$, which is invertible for every positive regularization strength α . The sigmoid function and the logit of the opening paragraph are an invertible pair, which allows translating between the score of a linear classifier and a probability value. A feature transformation that is invertible loses no information, since the raw feature vector can be recovered from the transformed one. The inverse of the cumulative distribution function (cdf) of a random variable (RV) allows generating realizations of that RV from a realization of another, simpler RV, one uniformly distributed on $[0, 1]$ (Ross, 2014).

A counterfactual explanation for a learned hypothesis can be constructed from a right inverse of it, which returns, for a given prediction, one feature vector of its preimage. A learned hypothesis usually maps many feature vectors to the same prediction, so it has no inverse and no left inverse; a right inverse is what remains, and choosing which element of the preimage it returns is what makes one counterfactual more useful than another.

Cross References

function, injective, preimage, invertible, matrix, linear map, determinant, pseudoinverse, linear regression, sigmoid function, logit

References

1. Golub and Van Loan (2013). Matrix Computations. 4th ed., The Johns Hopkins Univ. Press, Baltimore, MD, USA.
2. Horn and Johnson (1991). Topics in Matrix Analysis. Cambridge Univ. Press, Cambridge, U.K.
3. G. Strang (2007). Computational Science and Engineering. Wellesley-Cambridge Press, Wellesley, MA, USA.
4. Ross (2014). A First Course in Probability. 9th ed., Pearson Education, Boston, MA, USA.