

Proof Writing

Last modified: 2026-09-28 23:45:45 (PDT)

This page collects general advice on how to write proofs and derivations. The goal of a proof is not just to convince yourself that a result is true; it is to convince a *reader*, and to show them *why* it is true. Each principle on this page serves that goal. The symbols for logical entailment are listed under [Proofs](#) on the Notation page.

Don't skip steps

Show every step of the derivation. Each line should follow from the previous line by a single, identifiable move:

- applying a definition,
- substituting a known result, or
- performing one algebraic operation.

When you skip steps, you force the reader to reconstruct your reasoning, which is exactly the work the proof was supposed to do *for* them. A step that feels “obvious” while you are writing it is often the step a reader gets stuck on. Phrases like “it can be shown that” or “clearly” usually mark a skipped step; show the step instead.

Annotate each step

For each step, note the definition, theorem, or algebraic rule that justifies it. A brief annotation to the right of the line is usually enough. Annotations make the proof checkable, and they teach the reader which tool to reach for in similar situations.

For example, here is a derivation that the [hat matrix](#) $\mathbf{H} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ is idempotent, $\mathbf{H}^2 = \mathbf{H}$, with each step annotated:

$$\begin{aligned}
\mathbf{H}^2 &= \mathbf{H}\mathbf{H} && \text{(definition of a matrix power)} \\
&= \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top && \text{(definition of } \mathbf{H} \text{)} \\
&= \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} (\mathbf{X}^\top \mathbf{X}) (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top && \text{(regroup; matrix multiplication is associative)} \\
&= \mathbf{X} \mathbf{I}_p (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top && \text{(a matrix inverse times the matrix is } \mathbf{I}_p \text{)} \\
&= \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top && \text{(the identity matrix changes nothing)} \\
&= \mathbf{H} && \text{(definition of } \mathbf{H} \text{)}
\end{aligned}$$

The [Linear Algebra](#) page uses this derivation to show that $\mathbf{H}$ is a projection matrix. The probability notes derive the variance identity $\text{Var}(X) = \text{E}[X^2] - (\text{E}[X])^2$ [in the same annotated style](#).

Follow the golden rule

In general, follow the golden rule: treat your readers the way you want to be treated as a reader.

When you read someone else's proof, you want to be able to follow every step without guessing, to know which result is being used at each line, and to never be left wondering where a quantity came from. Write your own proofs to meet that same standard.

References