L^p Spaces

Last modified: 2026-10-31 15:22:31 (PDT)

L^p spaces

Definition 0.1 (L^p norm (p-norm of a function))**.** Let $(S, \mathcal{S}, \mu)$ be a measure space. For any real number $p \geq 1$, the L^p **norm** of a measurable function $f : S \to \mathbb{R}$ is:

$$\|f\|_p \stackrel{\text{def}}{=} \left(\int_S |f|^p \, d\mu \right)^{1/p}.$$

Example 0.1 (The L^1 and L^2 norms of $x \mapsto x$ on the unit interval)**.** Let $S = [0, 1]$ with the Lebesgue measure μ, and let $f(x) = x$. By Definition 0.1,

$$\begin{aligned} \|f\|_1 &= \int_0^1 x \, dx \\ &= \frac{1}{2}, \end{aligned}$$

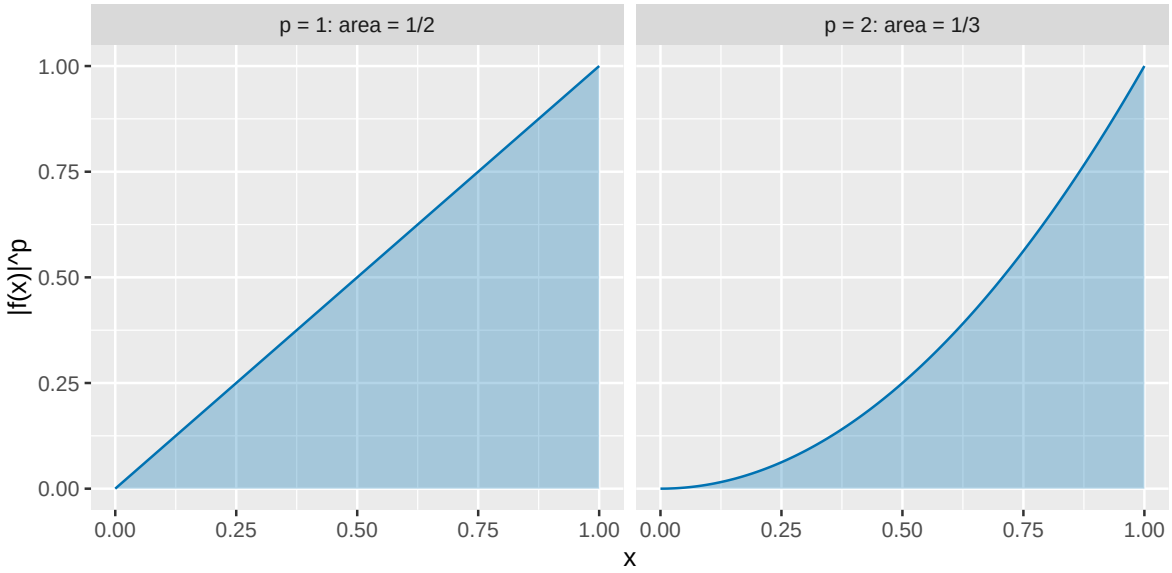
and

$$\begin{aligned} \|f\|_2 &= \left(\int_0^1 x^2 \, dx \right)^{1/2} \\ &= \left(\frac{1}{3} \right)^{1/2}. \end{aligned}$$

Definition 0.2 ($\mathcal{L}^p$ space ($\mathcal{L}^p(S, \mathcal{S}, \mu)$))**.** Let $(S, \mathcal{S}, \mu)$ be a measure space. Let $p \geq 1$ be a real number. The $\mathcal{L}^p$ **space** $\mathcal{L}^p(S, \mathcal{S}, \mu)$ is the set of all measurable functions $f : S \to \mathbb{R}$ whose L^p norm is less than ∞: $\|f\|_p < \infty$.

R

```r
lpna_x_pf <- seq(0, 1, length.out = 201)
lpna_p1_pf <- tibble::tibble(
  x = lpna_x_pf, y = lpna_x_pf, panel = "p = 1: area = 1/2"
)
lpna_sq_pf <- lpna_x_pf^2
lpna_p2_pf <- tibble::tibble(
  x = lpna_x_pf, y = lpna_sq_pf, panel = "p = 2: area = 1/3"
)
lpna_curves_pf <- dplyr::bind_rows(lpna_p1_pf, lpna_p2_pf)
lpna_aes_pf <- ggplot2::aes(x, y)
ggplot2::ggplot(lpna_curves_pf, lpna_aes_pf) +
  ggplot2::geom_area(fill = "#0072B2", alpha = 0.3) +
  ggplot2::geom_line(colour = "#0072B2") +
  ggplot2::facet_wrap(ggplot2::vars(panel)) +
  ggplot2::labs(x = "x", y = "|f(x)|^p")
```



Python

```python
import matplotlib.pyplot as plt
import numpy as np

lpna_x_pf = np.linspace(0, 1, 201)
fig, axes = plt.subplots(1, 2, figsize=(7, 3.5), sharey=True,
                         layout="constrained")
for ax, p, title in zip(axes, [1, 2], ["p = 1: area = 1/2", "p = 2: area = 1/3"]):
    ax.fill_between(lpna_x_pf, lpna_x_pf**p, color="#0072B2", alpha=0.3)
    ax.plot(lpna_x_pf, lpna_x_pf**p, color="#0072B2")
    ax.set_title(title)
    ax.set_xlabel("x")
axes[0].set_ylabel("|f(x)|^p")
plt.show()
```



Example 0.2 (An L^1 function not in L^2)**.** Consider the interval $(0,1)$ with the Borel σ-algebra (Definition 3 in Sigma-Algebras) and Lebesgue measure μ (Definition 6 in Measures and Probability Measures). Define $f : (0,1) \to \mathbb{R}$ by $f(x) = x^{-1/2}$.

1. For $p = 1$:
$$\begin{aligned}\|f\|_1 &= \int_0^1 x^{-1/2}\,dx \\ &= \left[2x^{1/2}\right]_0^1 \\ &= 2 \\ &< \infty,\end{aligned}$$
so $f \in \mathcal{L}^1((0,1),\mu)$.
2. For $p = 2$:
$$\begin{aligned}|f(x)|^2 &= \left(x^{-1/2}\right)^2 \\ &= \frac{1}{x}.\end{aligned}$$
The integral is:
$$\begin{aligned}\int_0^1 \frac{1}{x}\,dx &= \lim_{\varepsilon\to 0^+}\left[\ln x\right]_\varepsilon^1 \\ &= \lim_{\varepsilon\to 0^+}(-\ln\varepsilon) \\ &= \infty.\end{aligned}$$
So $\|f\|_2 = \infty$, so $f \notin \mathcal{L}^2((0,1),\mu)$.

This shows that $\mathcal{L}^2 \subset \mathcal{L}^1$ on sets of finite measure: a larger p imposes stricter requirements near singularities.

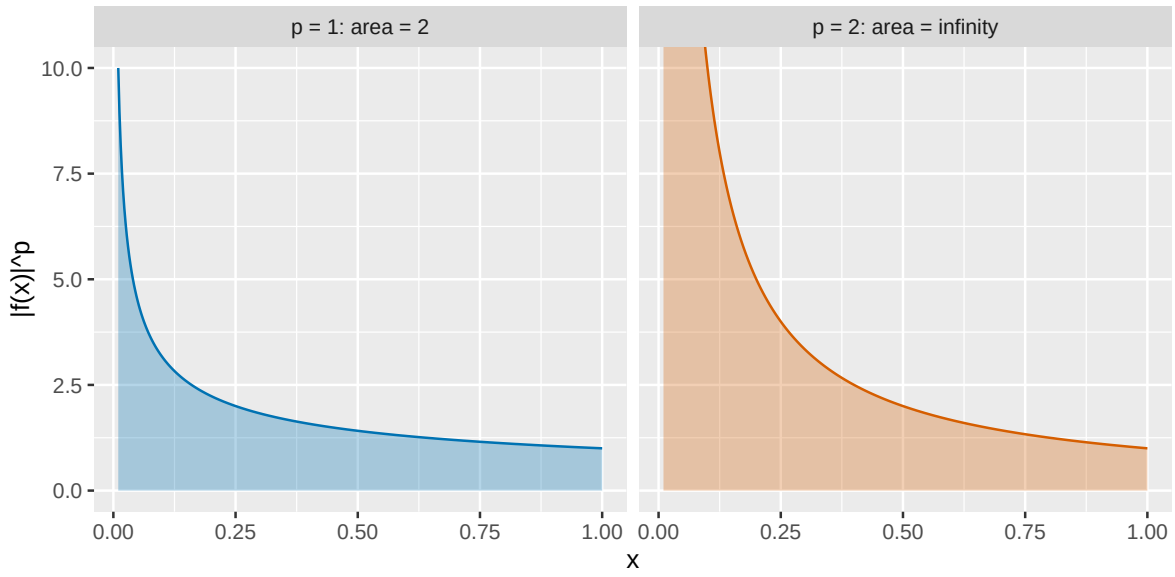
Definition 0.3 (Equal almost everywhere ($f = g$ μ-a.e.))**.** Let $(S,\mathcal{S},\mu)$ be a measure space. Let $f, g : S \to \mathbb{R}$ be measurable. The functions f and g are **equal almost everywhere**, written $f = g$ μ-a.e., if $f(s) = g(s)$ holds almost everywhere.

Example 0.3 (A function that differs from 0 at one point)**.** Let $S = \mathbb{R}$ with the Lebesgue measure μ. Let $f(s) = 1$ if $s = 0$, and $f(s) = 0$ otherwise. The set $\{s : f(s) \neq 0\} = \{0\}$ has measure zero, so $f = 0$ almost everywhere by Definition 0.3.

Definition 0.4 (L^p space ($L^p(S,\mathcal{S},\mu)$, $L^p(\mu)$))**.** Let $(S,\mathcal{S},\mu)$ be a measure space. Let $p \geq 1$ be a real number. For each f in the $\mathcal{L}^p$ space, let $[f]$ be the set of all functions in $\mathcal{L}^p(S,\mathcal{S},\mu)$ that are equal almost everywhere to f. The L^p **space** $L^p(S,\mathcal{S},\mu)$ (or simply

R

```r
lpss_x_pf <- seq(0.01, 1, length.out = 400)
lpss_root_pf <- lpss_x_pf^(-1 / 2)
lpss_p1_pf <- tibble::tibble(
  x = lpss_x_pf, y = lpss_root_pf, panel = "p = 1: area = 2"
)
lpss_inv_pf <- 1 / lpss_x_pf
lpss_p2_pf <- tibble::tibble(
  x = lpss_x_pf, y = lpss_inv_pf, panel = "p = 2: area = infinity"
)
lpss_curves_pf <- dplyr::bind_rows(lpss_p1_pf, lpss_p2_pf)
lpss_aes_pf <- ggplot2::aes(x, y, fill = panel, colour = panel)
lpss_colours_pf <- c("#0072B2", "#D55E00")
lpss_ylim_pf <- c(0, 10)
ggplot2::ggplot(lpss_curves_pf, lpss_aes_pf) +
  ggplot2::geom_area(alpha = 0.3, colour = NA) +
  ggplot2::geom_line() +
  ggplot2::facet_wrap(ggplot2::vars(panel)) +
  ggplot2::scale_fill_manual(values = lpss_colours_pf) +
  ggplot2::scale_colour_manual(values = lpss_colours_pf) +
  ggplot2::coord_cartesian(ylim = lpss_ylim_pf) +
  ggplot2::guides(fill = "none", colour = "none") +
  ggplot2::labs(x = "x", y = "|f(x)|^p")
```



Python

```python
import matplotlib.pyplot as plt
import numpy as np

lpss_x_pf = np.linspace(0.01, 1, 400)
lpss_curves_pf = [
    ("p = 1: area = 2", lpss_x_pf ** (-1 / 2), "#0072B2"),
    ("p = 2: area = infinity", 1 / lpss_x_pf, "#D55E00"),
]
fig, axes = plt.subplots(1, 2, figsize=(7, 3.5), sharey=True,
```

$L^p(\mu)$) is the set of all such sets $[f]$, with the norm $\|[f]\|_p \stackrel{\text{def}}{=} \|f\|_p$ (L^p norm).

Example 0.4 (The zero element of L^p)**.** Let f be the function in Example 0.3, which is 1 at $s = 0$ and 0 elsewhere. Then $f = 0$ almost everywhere, so $[f] = [0]$. Also $\|f\|_p = 0$ although $f \neq 0$, so $\|f\|_p = 0$ can hold for a nonzero element of $\mathcal{L}^p$; on $L^p(\mu)$, $\|[f]\|_p = 0$ only for $[f] = [0]$, as a norm requires.

Definition 0.5 (Essential supremum norm (L^∞ norm))**.** Let $(S, \mathcal{S}, \mu)$ be a measure space. The **essential supremum norm** (or L^∞ **norm**) of a measurable function $f : S \to \mathbb{R}$ is:

$$\|f\|_\infty \stackrel{\text{def}}{=} \inf\{C \geq 0 : \mu(\{s \in S : |f(s)| > C\}) = 0\}.$$

Example 0.5 (The essential supremum norm ignores a null set)**.** Let $S = [0, 1]$ with the Lebesgue measure μ. Let $f(s) = 5$ if $s = 0$ and $f(s) = s$ otherwise. Then $\sup_s |f(s)| = 5$. For $C \geq 1$, the set $\{s : |f(s)| > C\}$ is empty or $\{0\}$, so it has measure zero, and for $C < 1$ it has positive measure. By Definition 0.5, $\|f\|_\infty = 1$.

Definition 0.6 ($\mathcal{L}^\infty$ space ($\mathcal{L}^\infty(S, \mathcal{S}, \mu)$))**.** Let $(S, \mathcal{S}, \mu)$ be a measure space. The $\mathcal{L}^\infty$ **space** $\mathcal{L}^\infty(S, \mathcal{S}, \mu)$ is the set of all measurable functions $f : S \to \mathbb{R}$ whose essential supremum norm is less than ∞: $\|f\|_\infty < \infty$.

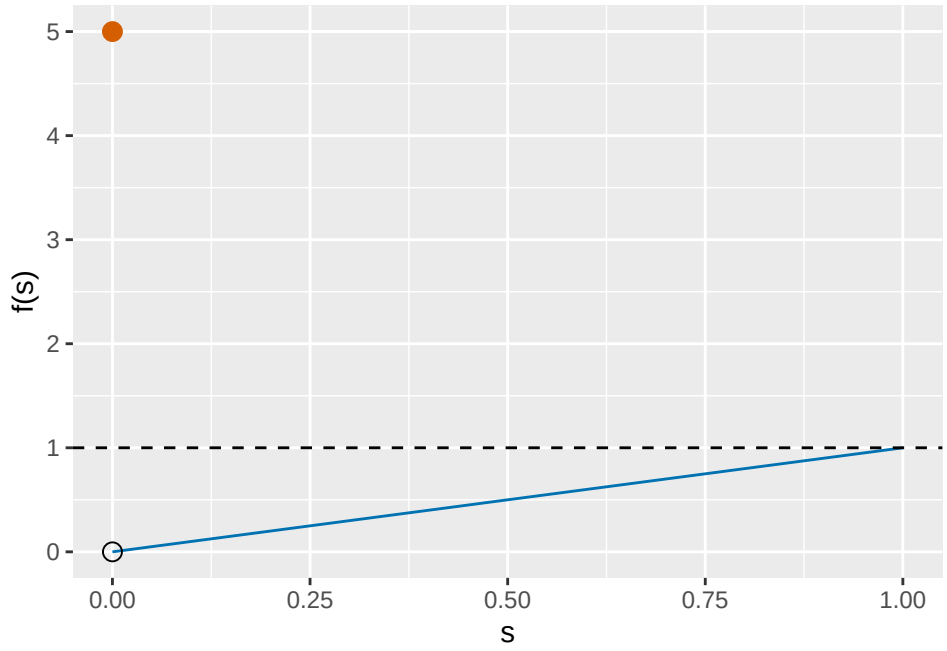
Example 0.6 (Bounded and unbounded functions)**.** Let $S = (0, 1)$ with the Lebesgue measure μ. The function $f(s) = s$ has $\|f\|_\infty = 1 < \infty$, so $f \in \mathcal{L}^\infty(S, \mathcal{S}, \mu)$ by Definition 0.6. The function $g(s) = 1/s$ exceeds every C on the set $(0, 1/C)$ of positive measure, so $\|g\|_\infty = \infty$ and $g \notin \mathcal{L}^\infty(S, \mathcal{S}, \mu)$.

Definition 0.7 (L^∞ space ($L^\infty(S, \mathcal{S}, \mu)$, $L^\infty(\mu)$))**.** Let $(S, \mathcal{S}, \mu)$ be a measure space. For each f in the $\mathcal{L}^\infty$ space, let $[f]$ be the set of all functions in $\mathcal{L}^\infty(S, \mathcal{S}, \mu)$ that are equal almost everywhere to f. The L^∞ **space** $L^\infty(S, \mathcal{S}, \mu)$ (or simply $L^\infty(\mu)$) is the set of all such sets $[f]$, with the norm $\|[f]\|_\infty \stackrel{\text{def}}{=} \|f\|_\infty$ (essential supremum norm).

Example 0.7 (Bounded functions equal almost everywhere)**.** Let f be the function in Example 0.5 and $h(s) = s$ on $[0, 1]$. Then $f = h$ almost everywhere (Definition 0.3), so $[f] = [h]$ in $L^\infty(\mu)$, and $\|[f]\|_\infty = \|h\|_\infty = 1$.

R

```r
esn_x_pf <- seq(0, 1, length.out = 101)
esn_line_pf <- tibble::tibble(x = esn_x_pf, y = esn_x_pf)
esn_open_pf <- tibble::tibble(x = 0, y = 0)
esn_spike_pf <- tibble::tibble(x = 0, y = 5)
esn_aes_pf <- ggplot2::aes(x, y)
ggplot2::ggplot(esn_line_pf, esn_aes_pf) +
  ggplot2::geom_line(colour = "#0072B2") +
  ggplot2::geom_point(data = esn_open_pf, shape = 1, size = 3) +
  ggplot2::geom_point(data = esn_spike_pf, colour = "#D55E00", size = 3) +
  ggplot2::geom_hline(yintercept = 1, linetype = "dashed") +
  ggplot2::labs(x = "s", y = "f(s)")
```



Python

```python
import matplotlib.pyplot as plt
import numpy as np

esn_x_pf = np.linspace(0, 1, 101)
fig, ax = plt.subplots(figsize=(5, 3.5), layout="constrained")
ax.plot(esn_x_pf, esn_x_pf, color="#0072B2")
ax.plot(0, 0, marker="o", markerfacecolor="white", color="black")
ax.plot(0, 5, marker="o", color="#D55E00")
ax.axhline(1, linestyle="dashed", color="black")
ax.set_xlabel("s")
ax.set_ylabel("f(s)")
plt.show()
```




Theorem 0.1 (Holder's inequality)**.** *Let $(S, \mathcal{S}, \mu)$ be a measure space, and let $p, q \in [1, \infty]$ be conjugate exponents satisfying $\frac{1}{p} + \frac{1}{q} = 1$ (with the convention that $q = \infty$ when $p = 1$, and $q = 1$ when $p = \infty$). If $f \in L^p(\mu)$ and $g \in L^q(\mu)$ (Definition 0.4), then their product fg is absolutely integrable ($fg \in L^1(\mu)$) and satisfies:*

$$\|fg\|_1 = \int_S |fg| \, d\mu \leq \|f\|_p \, \|g\|_q.$$

When

$$\begin{aligned} p &= q \\ &= 2, \end{aligned}$$

this inequality is the Cauchy-Schwarz inequality for integrals.

Theorem 0.2 (Minkowski's inequality (L^p triangle inequality))**.** *Let $(S, \mathcal{S}, \mu)$ be a measure space and let $1 \leq p \leq \infty$. For any functions $f, g \in L^p(\mu)$ (Definition 0.4), their sum $f + g$ is in $L^p(\mu)$ and satisfies:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Together with absolute homogeneity $\|cf\|_p = |c| \, \|f\|_p$ and definiteness, Minkowski's inequality establishes that $L^p(\mu)$ satisfies all axioms of a normed vector space.

Example 0.8 (Applying Holder's inequality to finite sums and expectations)**.**

1. **Finite vectors (Cauchy-Schwarz and taxicab bounds)**: Equipping a finite set with the counting measure, Holder's inequality for conjugate exponents p, q gives:

   $$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q \right)^{1/q}.$$

   Setting $p = 1$, $q = \infty$ gives $\sum_{i=1}^{n} |x_i y_i| \leq \|\underset{\sim}{x}\|_1 \, \|\underset{\sim}{y}\|_\infty$. Setting

   $$\begin{aligned} p &= q \\ &= 2 \end{aligned}$$

   gives the standard vector Cauchy-Schwarz inequality.

2. **Covariance bound for random variables**: On a probability space with measure $\mathbb{P}$ (Definition 12 in Measures and Probability Measures), taking

   $$\begin{aligned} p &= q \\ &= 2 \end{aligned}$$

gives $\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$, guaranteeing that the covariance between any two square-integrable random variables is finite.

i Remark (L^p spaces and norms in probability and machine learning)

Remark 0.1 (L^p spaces and norms in probability and machine learning). In probability, statistics, and machine learning, L^p spaces structure the geometry of random variables, estimators, and loss functions:

1. **Random variables, moments, and inclusion of L^p spaces**: When $\mu = \mathbb{P}$ is a probability measure (Definition 12 in Measures and Probability Measures) with total measure $\mathbb{P}(S) = 1$, an L^p function is a random variable X with finite p-th moment:
$$\|X\|_p = (\mathbb{E}[|X|^p])^{1/p}.$$
By Jensen's inequality, if $1 \leq p \leq q \leq \infty$, then $\|X\|_p \leq \|X\|_q$, so $L^q(P) \subseteq L^p(P)$. For instance, every random variable with finite variance ($X \in L^2$) has finite expectation ($X \in L^1$).

2. **The unique Hilbert space L^2 and conditional expectation**: Among all L^p spaces, only $L^2(\mu)$ is an inner product space, equipped with the inner product $\langle f, g\rangle \overset{\text{def}}{=} \int_S fg\, d\mu$ (all other L^p spaces fail the parallelogram law). The completeness of L^2 makes it a Hilbert space. In regression, the conditional expectation $\mathbb{E}[Y \mid X]$ is the orthogonal projection of $Y \in L^2$ onto the closed subspace of L^2 functions measurable with respect to X, minimizing the population mean squared error (pMSE) $\|Y - g(X)\|_2^2$.

3. **Loss functions, estimation, and robustness**: The choice of p in L^p norms defines loss functions for optimization and regression:
   - **L^2 loss (population mean squared error (pMSE), or sample mean squared error (sMSE) on a sample)**: penalizes large errors quadratically, yielding the conditional mean; sensitive to outliers.
   - **L^1 loss (population mean absolute error (pMAE), or sample mean absolute error (sMAE) on a sample)**: penalizes errors linearly, yielding the conditional median; robust to heavy tails and outliers.
   - **L^∞ loss (Chebyshev criterion)**: minimizes the worst-case maximum error across all instances.