Sobolev Embeddings

Introduction: A rule of thumb is that functions in L^q are "nicer" than functions in L^p for $q > p$. On this "niceness" scale, L^∞ is closer to the space of continuous functions than is L^1. The Sobolev embeddings say that when f and its first partial derivatives belong to L^p, then f automatically belongs to a "nicer" space, either L^q for some $q > p$, or a space of continuous functions. In one dimension, the embedding is immediate: if $f'(x)$ is merely in $L^1(R)$, then $f(x)$ is bounded and (absolutely) continuous. Embeddings in higher dimensions depend on the dimension in a somewhat surprising way.

We'll prove the two easiest and most important of the embedding theorems. The proofs require only Hölder's inequality and the Fundamental Theorem of Calculus.

These notes are organized into four sections: a (brief) review of Hölder's inequality and the embeddings it implies; estimates for embeddings into L^q; estimates for embeddings into C^λ, the (Hölder) continuous functions; and general remarks about embeddings and the problems caused by the boundary of a domain.

I have tried to emphasize the motivation behind the estimates below, to complement the density of the standard reference[1] in this subject.

Hölder's Inequality: A (real-valued, measurable) function belongs to the space $L^p(\Omega)$ if the norm,

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f(x)|^p \, dx \right)^{\frac{1}{p}},$$

is finite. We will always take functions f to be measurable and powers p to be ≥ 1.

One version of Hölder's inequality says that if p and p' are conjugate powers, meaning

$$\frac{1}{p} + \frac{1}{p'} = 1, \tag{H_1}$$

then

$$\int_\Omega |f(x)g(x)| \, dx \leq \|f\|_{L^p(\Omega)} \|g\|_{L^{p'}(\Omega)} \tag{H_2}$$

for all $f \in L^p(\Omega)$ and $g \in L^{p'}(\Omega)$. Conjugate powers are frequently written as p and q, but Sobolev's embedding uses p for the space the derivatives belong to, and q for the "nicer" space the function belongs to, so we'll use the p and p' notation when applying Hölder's inequality.

If $|\Omega|$, the measure (= volume) of Ω, is finite, then we may take $g(x) = 1$ to estimate

$$\|f\|_{L^1(\Omega)} = \int_\Omega |f(x)| \cdot 1 \, dx \leq \|f\|_{L^p(\Omega)} \|1\|_{L^{p'}(\Omega)} = \|f\|_{L^p(\Omega)} |\Omega|^{\frac{1}{p'}}. \tag{H_3}$$

One normed vector space X is **embedded** in another normed space Y if there is a constant K for which

$$\|f\|_Y \leq K \|f\|_X \quad \text{for all } f \in X.$$

Estimate (H_3) says that $L^p(\Omega)$ is embedded in $L^1(\Omega)$ with embedding constant $K = |\Omega|^{\frac{1}{p'}}$. As sets,

$$L^p(\Omega) \subset L^1(\Omega) \quad \text{if } |\Omega| < \infty.$$

In particular, L^p is more exclusive than L^1, and we think of functions in L^p as being "nicer" than those in L^1. This is an embedding theorem. We say "$L^p(\Omega)$ is embedded in $L^1(\Omega)$ when $|\Omega| < \infty$", and we write

$$L^p(\Omega) \hookrightarrow L^1(\Omega) \quad \text{if } |\Omega| < \infty.$$

[1] Adams, Robert A., *Sobolev Spaces*, Academic Press, New York, 1975.

Sobolev's embeddings ask what L^q-space or C^λ-space f belongs to if its *derivatives* belong to L^p. We therefore seek estimates of the form

$$\|f\|_{L^q} \text{ or } \|f\|_{C^\lambda} \leq K \sum_i \left\| \frac{\partial f}{\partial x_i} \right\|_{L^p}.$$

Our immediate goal, then, is to estimate the L^q and C^λ norms of f with the L^p norms of f's partial derivatives.

The embeddings in R^1 are immediate consequences of Hölder's inequality: if $f'(x) \in L^p(R^1)$ for $p > 1$, then

$$\begin{aligned} |f(y) - f(x)| = \left| \int_x^y f'(t)\, dt \right| &\leq \int_x^y |f'(t)|\, dt \\ &\leq \left(\int_x^y |f'(t)|^p\, dt \right)^{\frac{1}{p}} \left(\int_x^y 1^{p'}\, dt \right)^{\frac{1}{p'}} \qquad (H_4) \\ &\leq \|f'(t)\|_{L^p(R^1)} |y - x|^{\frac{1}{p'}}. \end{aligned}$$

In other words, if the derivative $f' \in L^p(R^1)$, then $f \in C^{\frac{1}{p'}}(R^1)$, the space of Hölder continuous functions with Hölder exponent $\frac{1}{p'}$ and norm

$$\|f\|_{C^\lambda} = \sup_{x \neq y} \frac{|f(y) - f(x)|}{|y - x|^{\frac{1}{p'}}}. \qquad (H_5)$$

If $p = \infty$ in Estimate (H_4), then $p' = 1$ and we may replace "Hölder continuous" with "Lipschitz continuous". The case $p = 1$ (so $p' = \infty$) is more complicated, and we only get to replace "Hölder continuous" with "absolutely continuous". It is common for the "boundary cases" $p = 1$ and $p = \infty$ to require special care, and the embeddings may not hold for them.

The proofs of Sobolev's embeddings are deceptively easy when the function's domain $\Omega \subset R^1$. The proofs are more challenging when the function's domain $\Omega \subset R^n$ for $n > 1$. Consequently(?), the embeddings appear more frequently in the study of partial differential equations than in ordinary differential equations.

We note here (for future use) that Hölder's inequality applies to the product of more than two functions, as well. If

$$\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_k} = 1, \qquad (H_6)$$

then

$$\int_\Omega |f_1(x) f_2(x) \cdots f_k(x)|\, dx \leq \|f_1\|_{L^{p_1}(\Omega)} \|f_2\|_{L^{p_2}(\Omega)} \cdots \|f_k\|_{L^{p_k}(\Omega)}. \qquad (H_7)$$

(We are sticking to our convention that the p_i are all ≥ 1, of course.)

Embeddings into L^q: It's easy to get lost in the subscripts, powers, and dimensions in the theorem. If we study functions defined in R^3 to start with, then we can use the familiar x, y, and z instead of $x_1, \cdots, x_n$.

(First) Estimates in R^3: To estimate a function in terms of its derivatives, we use the Fundamental Theorem of Calculus. Suppose f and its first derivatives are, say, continuous in R^3, and that $f(x, y, z)$ tends to zero as $\sqrt{x^2 + y^2 + z^2} \to \infty$. (In symbols, $f \in C_0^1(R^3)$.) Then the Fundamental Theorem of Calculus says

$$f(x, y, z) = \int_{-\infty}^x D_1 f(t, y, z)\, dt,$$

where $D_1 = \frac{\partial}{\partial x}$, the partial derivative with respect to the first variable. The equality holds for arbitrary values of y and z.

The Fundamental Theorem of Calculus also says

$$f(x, y, z) = -\int_x^\infty D_1 f(t, y, z)\, dt.$$

The average of the absolute values of the two integrals gives the estimate

$$
\begin{aligned}
|f(x,y,z)| &= \frac{1}{2}\left|\int_{-\infty}^{x} D_1 f(t,y,z)\,dt\right| + \frac{1}{2}\left|\int_{x}^{\infty} D_1 f(t,y,z)\,dt\right| \\
&\le \frac{1}{2}\int_{-\infty}^{\infty} |D_1 f(t,y,z)|\,dt \\
&\equiv g_1(y,z).
\end{aligned} \tag{S_1}
$$

The important point is that $g_1(y,z)$ is a function of 2 variables only. It is independent of x but still dominates $f(x,y,z)$ pointwise.

There is nothing special about the variable x; estimates independent of y and z hold as well:

$$
\begin{aligned}
|f(x,y,z)| &\le \frac{1}{2}\int_{-\infty}^{\infty} |D_2 f(x,t,z)|\,dt \equiv g_2(x,z), \qquad \text{and} \\
|f(x,y,z)| &\le \frac{1}{2}\int_{-\infty}^{\infty} |D_3 f(x,y,t)|\,dt \equiv g_3(x,y),
\end{aligned} \tag{S_2}
$$

By themselves, none of the g_i are integrable over R^3. For example, the first integral in the iterated integral $\int\int\int g_1(y,z)\,dx\,dy\,dz$ is

$$
\int g_1(y,z)\,dx = \begin{cases} \infty, & \text{if } g_1(y,z) > 0 \\ 0, & \text{if } g_1(y,z) = 0 \\ -\infty, & \text{if } g_1(y,z) < 0 \end{cases},
$$

which has no chance of being integrable over y and z. The trouble is that g_1 does not die off sufficiently rapidly at infinity. Indeed, it doesn't die off in the x-direction at all.

The point is, we can not hope to make progress by trying to estimate an L^q-norm using

$$
|f(x,y,z)| \le \frac{g_1(y,z) + g_2(x,z) + g_3(x,y)}{3},
$$

for example. The right side simply won't cooperate with integration.

A smarter move — one that guarantees die-off at infinity — is to multiply the g_i:

$$
|f(x,y,z)|^3 \le g_1(y,z)g_2(x,z)g_3(x,y). \tag{S_3}
$$

At least the right side has a chance of having a finite integral

$$
\|f\|^3_{L^3(R^3)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(x,y,z)|^3\,dx\,dy\,dz \le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(y,z)g_2(x,z)g_3(x,y)\,dx\,dy\,dz. \tag{S_4}
$$

We estimate the right side using Hölder's inequality. Fubini's Theorem permits us to treat the integral as an iterated integral, so we only need Hölder's inequality one variable at a time. We begin with the inner-most integral — the integral with respect to x. Since $g_1(y,z)$ is constant in x,

$$
\int_{-\infty}^{\infty} g_1(y,z)g_2(x,z)g_3(x,y)\,dx = g_1(y,z)\int_{-\infty}^{\infty} g_2(x,z)g_3(x,y)\,dx.
$$

The integrand on the right is the product of 2 functions, which suggests we use Hölder's inequality (H_2) with $p = 2 = p'$:

$$
\int_{-\infty}^{\infty} g_1(y,z)g_2(x,z)g_3(x,y)\,dx \le g_1(y,z)\sqrt{\int_{-\infty}^{\infty} g_2(x,z)^2\,dx}\sqrt{\int_{-\infty}^{\infty} g_3(x,y)^2\,dx}.
$$

(All of the $g_i \geq 0$, so we don't need any extra absolute values around g_1.)

Next, we integrate with respect to y. Since $\sqrt{\int_{-\infty}^{\infty} g_2(x,z)^2\,dx}$ is constant in y, it factors out of the integral, and we apply Hölder's inequality again:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(y,z)g_2(x,z)g_3(x,y)\,dx\,dy \leq \sqrt{\int_{-\infty}^{\infty} g_2(x,z)^2\,dx}\int_{-\infty}^{\infty}\left[g_1(y,z)\sqrt{\int_{-\infty}^{\infty} g_3(x,y)^2\,dx,}\right]\,dy.$$

$$\leq \sqrt{\int_{-\infty}^{\infty} g_2(x,z)^2\,dx}\sqrt{\int_{-\infty}^{\infty} g_1(y,z)^2\,dy}\sqrt{\int_{-\infty}^{\infty}\left(\sqrt{\int_{-\infty}^{\infty} g_3(x,y)^2\,dx}\right)^2 dy}$$

$$= \sqrt{\int_{-\infty}^{\infty} g_2(x,z)^2\,dx}\sqrt{\int_{-\infty}^{\infty} g_1(y,z)^2\,dy}\sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_3(x,y)^2\,dx\,dy}.$$

Notice how nicely the square and square root "cooperate" in the last integral.

One more integral, this time with respect to z, and we'll be done. Since $\sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_3(x,y)^2\,dx\,dy}$ is independent of z, it factors out of this final integral. The squares and square roots from Hölder's inequality cooperate nicely again, leaving

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(y,z)g_2(x,z)g_3(x,y)\,dx\,dy\,dz$$

$$\leq \sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_3(x,y)^2\,dx\,dy}\int_{-\infty}^{\infty}\left[\sqrt{\int_{-\infty}^{\infty} g_2(x,z)^2\,dx}\sqrt{\int_{-\infty}^{\infty} g_1(y,z)^2\,dy}\right]\,dz$$

$$\leq \sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_3(x,y)^2\,dx\,dy}\sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_2(x,z)^2\,dx\,dz}\sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_1(x,y)^2\,dy\,dz}$$

$$= \|g_3\|_{L^2(R^2)}\,\|g_2\|_{L^2(R^2)}\,\|g_1\|_{L^2(R^2)}\,.$$

In the norm notation, Estimate (S_4) becomes

$$\|f\|^3_{L^3(R^3)} \leq \|g_3\|_{L^2(R^2)}\,\|g_2\|_{L^2(R^2)}\,\|g_1\|_{L^2(R^2)}$$

so

$$\|f\|_{L^3(R^3)} \leq \left(\|g_1\|_{L^2(R^2)}\,\|g_2\|_{L^2(R^2)}\,\|g_3\|_{L^2(R^2)}\right)^{\frac{1}{3}}\,. \qquad (S_5)$$

In English: The L^3-norm of f over all of the 3-dimensional space R^3 is dominated by the geometric mean of the L^2-norms of the g_i over the 2-dimensional subspaces R^2. Estimate (S_5) holds as long as each of the g_i is independent of the i^{th} variable and dominates f pointwise. It is simply the correct application of Hölder's inequality to functions of a special type.

All that remains is to decipher what the norms of the g_i mean in terms of the partial derivatives of f. For example,

$$\|g_1\|_{L^2(R^2)} = \sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\frac{1}{2}\int_{-\infty}^{\infty} D_1 f(x,y,z)\,dx\right)^2 dy\,dz}.$$

Hmmm. This time, the square does not cooperate with Hölder's inequality. The norm is logically correct, but awkward — the x-coordinate is treated differently than the y- and z-coordinates. Is there some way to (re)move the square?

Estimate (S_3) is a pointwise estimate, and remains valid upon taking square roots of both sides:

$$|f(x,y,z)|^{\frac{3}{2}} \leq g_1(y,z)^{\frac{1}{2}}g_2(x,z)^{\frac{1}{2}}g_3(x,y)^{\frac{1}{2}}. \qquad (S_6)$$

Repeat the integrations and Hölder's inequalities from above, and Estimate (S_5) becomes

$$\|f\|^{\frac{3}{2}}_{L^{\frac{3}{2}}(R^3)} \leq \|\sqrt{g_1}\|_{L^2(R^2)}\,\|\sqrt{g_2}\|_{L^2(R^2)}\,\|\sqrt{g_3}\|_{L^2(R^2)}\,.$$

The advantage of this formulation is that the troublesome square goes away:

$$\begin{aligned}\|\sqrt{g_1}\|_{L^2(R^2)} &= \sqrt{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sqrt{\frac{1}{2}\int_{-\infty}^{\infty}|D_1 f(x,y,z)|\,dx}\right)^2 dy\,dz} \\ &= \sqrt{\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D_1 f(x,y,z)\,dx\,dy\,dz} \\ &= \sqrt{\frac{1}{2}\,\|D_1 f\|_{L^1(R^3)}}.\end{aligned}$$

The elegant estimate is therefore

$$\|f\|^{\frac{3}{2}}_{L^{\frac{3}{2}}(R^3)} \le \sqrt{\frac{1}{2^3}\,\|D_1 f\|_{L^1(R^3)}\,\|D_2 f\|_{L^1(R^3)}\,\|D_3 f\|_{L^1(R^3)}},$$

or

$$\|f\|_{L^{\frac{3}{2}}(R^3)} \le \frac{1}{2}\left(\|D_1 f\|_{L^1(R^3)}\,\|D_2 f\|_{L^1(R^3)}\,\|D_3 f\|_{L^1(R^3)}\right)^{\frac{1}{3}}. \qquad (S_7)$$

In English: the $L^{\frac{3}{2}}$-norm of f is dominated by (half of) the geometric mean of the L^1-norms of the partial derivatives of f. If the derivatives belong to the (minimally smooth) space L^1, then the function itself belongs to the "smoother" space $L^{\frac{3}{2}}$ — "smoother" in the sense that $\frac{3}{2} > 1$. Estimate (S_7) is our first example of a Sobolev embedding.

(First) Estimates in R^n: Estimates (S_5) and (S_7) are easy to extend from R^3 to R^n. Suppose $f \in C_0^1(R^n)$. Estimates (S_1) and (S_2) generalize to

$$|f(x_1, x_2, \ldots, x_n)| \le \frac{1}{2}\int_{-\infty}^{\infty} |D_i f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)|\,dt \equiv g_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

The dominating function g_i is a function of $n-1$ variables; it is missing x_i, which has been integrated out.

What estimate corresponds to the pointwise Estimate (S_6)? We want to integrate over R^n, using Hölder's inequality on each of the iterated integrals. Each iterated integral will have a constant term, which we factor out of the integral. Hölder's inequality (H_7) with $p_i = n-1$ will then apply to the remaining $n-1$ factors, so it is wise to have a power of $\frac{1}{p_i} = \frac{1}{n-1}$ on every factor. The "right" estimate corresponding to Estimate (S_6) is therefore

$$|f(x_1, x_2, \ldots, x_n)|^{\frac{n}{n-1}} \le \prod_{i=1}^{n} g_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)^{\frac{1}{n-1}}.$$

The sequence of iterated integrations begins just as in the case of R^3 as

$$\begin{aligned}\int_{-\infty}^{\infty} |f(x_1, x_2, \ldots, x_n)|^{\frac{n}{n-1}}\,dx_1 &\le g_1(x_2, \ldots, x_n)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} g_2(x_1, x_3, \ldots)^{\frac{1}{n-1}} \cdots g_n(x_1, x_2, \ldots)^{\frac{1}{n-1}}\,dx_1 \\ &\le g_1(x_2, \ldots, x_n)^{\frac{1}{n-1}} \left(\int_{-\infty}^{\infty} g_2(x_1, x_3, \ldots)\,dx_1\right)^{\frac{1}{n-1}} \cdots \left(\int_{-\infty}^{\infty} g_n(x_1, x_2, \ldots)\,dx_1\right)^{\frac{1}{n-1}}.\end{aligned}$$

Iterating the integrations one additional variable at a time, we arrive at

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |f(x_1, x_2, \ldots, x_n)|^{\frac{n}{n-1}}\,dx_1 \cdots dx_n \le \left(\|g_1\|_{L^1(R^{n-1})}\,\|g_2\|_{L^1(R^{n-1})} \cdots \|g_n\|_{L^1(R^{n-1})}\right)^{\frac{1}{n-1}}.$$

The left side is $\|f\|^{\frac{n}{n-1}}_{L^{\frac{n}{n-1}}(R^n)}$, so

$$\|f\|_{L^{\frac{n}{n-1}}(R^n)} \le \Big(\|g_1\|_{L^1(R^{n-1})}\,\|g_2\|_{L^1(R^{n-1})}\cdots\|g_n\|_{L^1(R^{n-1})}\Big)^{\frac{1}{n}}.$$

This estimate is a Hölder inequality for functions of a special form, and we record the result as a lemma:

Lemma 1: If, for each $i = 1, \dots, n$, the estimate

$$|f(x_1, \dots, x_n)| \le g_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \tag{S_8}$$

holds pointwise, then

$$\|f\|_{L^{\frac{n}{n-1}}(R^n)} \le \left(\prod_{i=1}^{n}\|g_i\|_{L^1(R^{n-1})}\right)^{\frac{1}{n}}. \tag{S_9}$$

In English: The geometric mean of the L^1-norms of the g_i dominates the $L^{\frac{n}{n-1}}$-norm of f. The peculiarity is that f's norm is taken over all of R^n, while each of the g_i's norms is taken over an $(n-1)$-dimensional coordinate subspace of R^n.

We now apply the Lemma to the case where each of the norms on the right side is of the form

$$\begin{aligned}\|g_1\|_{L^1(R^{n-1})} &= \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-1\times}\frac{1}{2}\int_{-\infty}^{\infty}|D_i f(x_1, \dots, t, \dots, x_n)|\,dt\,x_1\,\cdots\,dx_n\\ &= \frac{1}{2}\,\|D_i f\|_{L^1(R^n)}\end{aligned}$$

Together with Estimate (S_9), this means

$$\|f\|_{L^{\frac{n}{n-1}}(R^n)} \le \frac{1}{2}\left(\prod_{i=1}^{n}\|D_i f\|_{L^1(R^n)}\right)^{\frac{1}{n}}, \tag{S_{10}}$$

at least for all $f \in C_0^1(R^n)$. This is our first example of a Sobolev embedding for functions on R^n.

The general case: What if the derivatives $D_i f$ belong to L^p instead of L^1? Surely "smoother" derivatives imply "smoother" functions. To discover the result, we need the integral of $|D_i f|^p$ in place of the integral of $D_i f$ in The Fundamental Theorem of Calculus. Hölder's inequality is the natural choice.

Replace f with $|f|^\gamma$ for some $\gamma \ge 1$ in the computations above. The Fundamental Theorem of Calculus says

$$|f(x_1, \dots, x_n)|^\gamma \le \frac{\gamma}{2}\int_{-\infty}^{\infty}|f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^{\gamma-1}\,|D_i f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|\,dt.$$

The right side is our g_i, and by Hölder's inequality,

$$\begin{aligned}\|g_i\|_{L^1(R^{n-1})} &= \left\|\int_{-\infty}^{\infty}|f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|^{\gamma-1}\,|D_i f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)|\,dt\right\|_{L^1(R^{n-1})}\\ &\le \|f\|^{\gamma-1}_{L^{p'(\gamma-1)}(R^n)}\,\|D_i f\|_{L^p(R^n)}.\end{aligned}$$

Consequently, Lemma 1 implies that

$$\begin{aligned}\|f\|^{\gamma}_{L^{\frac{\gamma n}{n-1}}(R^n)} &\le \frac{\gamma}{2}\left(\prod_{i=1}^{n}\|f\|^{\gamma-1}_{L^{p'(\gamma-1)}(R^n)}\,\|D_i f\|_{L^p(R^n)}\right)^{\frac{1}{n}}\\ &= \frac{\gamma}{2}\,\|f\|^{\gamma-1}_{L^{p'(\gamma-1)}(R^n)}\left(\prod_{i=1}^{n}\|D_i f\|_{L^p(R^n)}\right)^{\frac{1}{n}}.\end{aligned} \tag{S_{11}}$$

for all $f \in C_0^1(R^n)$. Note that we should keep an eye on the $L^{p'(\gamma-1)}$ "norm" to ensure that that $p'(\gamma-1) \geq 1$.

It would be very convenient indeed if we chose γ so that the two norms of f on the two sides of Estimate (S_{11}) were the same. In other words, we should choose γ so that

$$\frac{\gamma n}{n-1} = p'(\gamma - 1).$$

This condition means, since $p' = \frac{p}{p-1}$,

$$\begin{aligned}\gamma &= \frac{p'(n-1)}{p'(n-1)-n}\\ &= \frac{n-1}{n-p}p.\end{aligned}$$

Let q be the common exponent $\frac{\gamma n}{n-1} = p'(\gamma-1)$:

$$q = \frac{\gamma n}{n-1} = \frac{n}{n-p}p. \tag{S_{12}}$$

Recall that we want $q \geq 1$ so the L^q-norm really is a norm, so we want $p < n$. (In this case, $q > p$, so L^q is "nicer" than L^p, and we have achieved our goal.) Then Estimate (S_{11}) reads

$$\|f\|^{\gamma}_{L^q(R^n)} \leq \frac{\gamma}{2}\|f\|^{\gamma-1}_{L^{q)}(R^n)}\left(\prod_{i=1}^n \|D_i f\|_{L^p(R^n)}\right)^{\frac{1}{n}},$$

and we may cancel the $\|f\|^{\gamma-1}_{L^{q)}(R^n)}$ from both sides, leaving

$$\|f\|_{L^q(R^n)} \leq \frac{n-1}{2(n-p)}p\left(\prod_{i=1}^n \|D_i f\|_{L^p(R^n)}\right)^{\frac{1}{n}} \tag{S_{13}}$$

for all $f \in C_0^1(R^n)$. For reasons that will become clear later, the relation between p and q in Equation (S_{12}) is best written in terms of their reciprocals (just as it is in Hölder's inequality):

$$\frac{n}{q} = \frac{n}{p} - 1. \tag{S_{14}}$$

We record this result as a theorem:

Theorem 2: If $\frac{n}{q} = \frac{n}{p} - 1 > 0$, then Estimate (S_{13}) holds.

Remark: Note that the embedding constant in Estimate (S_{13}),

$$\frac{n-1}{2(n-p)}p = \frac{n-1}{2\left(\frac{n}{p}-1\right)} = \frac{n-1}{2n}q$$

goes to ∞ as $p \uparrow n$. This behavior is typical, and we must exercise caution. We can't hope to embed into L^∞ when $p = n$: translation is continuous in L^p but not in L^∞.

Embeddings into C^λ: What if $\frac{n}{q} = \frac{n}{p} - 1 < 0$ in Theorem 2? The situation is even better: f is Hölder continuous!

The proofs of Sobolev's Embeddings into the (Hölder) continuous functions have a slightly different flavor than the proofs of embeddings into L^q. Our motivation is Estimate (H_4), which was easy to prove. Let us attempt the corresponding computation in R^n for $n > 1$.

If $f \in C^1(R^n)$, the Fundamental Theorem of Calculus and the Chain Rule say

$$f(y) - f(x) = \int_0^1 \frac{d}{dt} f\left(x + t(y-x)\right)\, dt = \int_0^1 \sum_{i=1}^n D_i f\left(x + t(y-x)\right)(y_i - x_i)\, dt. \qquad (S_{15})$$

Fix an x and let Ω be a convex subset of R^n which contains x. We integrate both sides of Equation (S_{15}) with respect to y over Ω:

$$\int_\Omega f(y) - f(x)\, dy = \int_\Omega \int_0^1 \sum_{i=1}^n D_i f\left(x + t(y-x)\right)(y_i - x_i)\, dt\, dy.$$

On the left side, $f(x)$ is a constant. We let $|\Omega|$ denote the measure (= volume) of Ω and conclude that

$$\frac{1}{|\Omega|} \int_\Omega f(y)\, dy - f(x) = \frac{1}{|\Omega|} \int_\Omega \int_0^1 \sum_{i=1}^n D_i f\left(x + t(y-x)\right)(y_i - x_i)\, dt\, dy. \qquad (S_{16})$$

The left side is the difference of the average of f over Ω and $f(x)$, so the right side measures how much $f(x)$ differs from its mean over Ω. Each of the terms in the integrand is of the form

$$D_i f\left(x + t(y-x)\right)(y_i - x_i) \le \left|D_i f\left(x + t(y-x)\right)\right| \sup_{y \in \Omega} |y - x|.$$

We estimate the integral of each of these terms separately. The rightmost term, the supremum, is independent of y, and factors out of any integrals over Ω.

To get the estimate we want, namely something involving the p-norm of the derivatives, we switch the order of integration. This computation is interesting enough to be recorded as a lemma:

Lemma 3: If Ω is convex, $x \in \Omega$, and g is integrable over Ω, then

$$\int_\Omega \int_0^1 g\left(x + t(y-x)\right)\, dt\, dy = \int_0^1 t^{-n} \int_{\Omega_t} g(z)\, dz\, dt, \qquad (S_{17})$$

where Ω_t is Ω dilated by a factor of t, taking x as the center of dilation.

Proof: Fubini's Theorem permits the change of order in the integration

$$\int_\Omega \int_0^1 g\left(x + t(y-x)\right)\, dt\, dy = \int_0^1 \int_\Omega g\left(x + t(y-x)\right)\, dy\, dt.$$

The change is particularly easy because the limits of integration in y are independent of t.

Since $dy = dy_1\, dy_2 \cdots dy_n$ is the volume element in R^n, and since the dilation scales in each of the n directions, it is easy to see that the substitution $z = x + t(y-x)$ (where t and x are constant) implies

$$dz = d(t(y-x)) = t^n\, dy.$$

The only hard part about the computation is the limits of integration. As y varies over all of Ω, $y - x$ varies over a copy of Ω translated by x. Similarly, $t(y-x)$ varies over a dilated copy of the translated Ω. Adding x back in (to get $x + t(y-x)$) translates the dilated, translated copy to a dilated copy where the dilation is about the point x. That's exactly what Equation (S_{17}) says. ////

Note that the measure (= volume) of Ω_t is

$$|\Omega_t| = t^n\, |\Omega|.$$

Equation (S_{17}) implies that Equation (S_{16}) is dominated by

$$\begin{aligned}\left|\frac{1}{|\Omega|}\int_\Omega f(y)\,dy - f(x)\right| &\le \frac{\sup_{y\in\Omega}|y-x|}{|\Omega|}\sum_{i=1}^n\int_\Omega\int_0^1 |D_i f\,(x+t(y-x))|\;dt\,dy\\ &= \frac{\sup_{y\in\Omega}|y-x|}{|\Omega|}\sum_{i=1}^n\int_0^1 t^{-n}\int_{\Omega_t}|D_i f\,(z)|\;dz\,dt\end{aligned}$$

Hölder's inequality now implies that

$$\begin{aligned}\left|\frac{1}{|\Omega|}\int_\Omega f(y)\,dy - f(x)\right| &\le \frac{\sup_{y\in\Omega}|y-x|}{|\Omega|}\sum_{i=1}^n\int_0^1 t^{-n}\left(\int_{\Omega_t}|D_i f\,(z)|^p\;dz\right)^{\frac{1}{p}}|\Omega_t|^{\frac{1}{p'}}\,dt\\ &= \frac{\sup_{y\in\Omega}|y-x|}{|\Omega|^{1-\frac{1}{p'}}}\sum_{i=1}^n\int_0^1 t^{-n+\frac{n}{p'}}\left(\int_{\Omega_t}|D_i f\,(z)|^p\;dz\right)^{\frac{1}{p}}dt\\ &= \frac{\sup_{y\in\Omega}|y-x|}{|\Omega|^{\frac{1}{p}}}\sum_{i=1}^n\int_0^1 t^{\frac{-n}{p}}\left(\int_{\Omega_t}|D_i f\,(z)|^p\;dz\right)^{\frac{1}{p}}dt\end{aligned}$$

The last integral in t is finite provided $\frac{n}{p}-1<0$. In this case,

$$\left|\frac{1}{|\Omega|}\int_\Omega f(y)\,dy - f(x)\right| \le \frac{\sup_{y\in\Omega}|y-x|}{\left(1-\frac{n}{p}\right)|\Omega|^{\frac{1}{p}}}\sum_{i=1}^n\|D_i f\|_{L^p(\Omega)}\qquad\text{provided } \frac{n}{p}-1<0. \tag{S_{18}}$$

Estimate (S_{18}) says that $f(x)$ differs from its mean over Ω by no more than the $L^p(\Omega)$ norm of the derivatives of f times a term that depends only on the measure and diameter of Ω. This estimate is enough to prove that f is bounded if f and its first derivatives are in $L^p(R^n)$:

Lemma 4: If $\frac{n}{p}-1<0$ and f and its first derivatives are in $L^p(R^n)$, then $f\in L^\infty(R^n)$.

Proof: Take Ω be the ball of radius 1 with center x. Then Estimate (S_{18}) and the triangle and Hölder's inequalities imply

$$\begin{aligned}|f(x)| &\le \frac{1}{|\Omega|}\left|\int_\Omega f(y)\,dy\right| + \frac{\sup_{y\in\Omega}|y-x|}{\left(1-\frac{n}{p}\right)|\Omega|^{\frac{1}{p}}}\sum_{i=1}^n\|D_i f\|_{L^p(\Omega)}\\ &\le \frac{1}{|\Omega|}\left(\int_\Omega|f(y)|^p\;dy\right)^{\frac{1}{p}}|\Omega|^{\frac{1}{p'}} + \frac{1}{\left(1-\frac{n}{p}\right)|\Omega|^{\frac{1}{p}}}\sum_{i=1}^n\|D_i f\|_{L^p(\Omega)} \qquad (S_{19})\\ &\le \frac{1}{|\Omega|^{\frac{1}{p}}}\|f\|_{L^p(R^n)} + \frac{1}{\left(1-\frac{n}{p}\right)|\Omega|^{\frac{1}{p}}}\sum_{i=1}^n\|D_i f\|_{L^p(R^n)}\,.\end{aligned}$$

The last line is independent of x and therefore provides a uniform bound on f. ////

We turn now to the question of Hölder continuity. The triangle inequality and Estimate (S_{18}) say that

$$\begin{aligned}|f(z)-f(x)| &\le \left|f(z)-\frac{1}{|\Omega|}\int_\Omega f(y)\,dy\right| + \left|\frac{1}{|\Omega|}\int_\Omega f(y)\,dy - f(x)\right|\\ &\le \left(\frac{\sup_{y\in\Omega}|y-z| + \sup_{y\in\Omega}|y-x|}{\left(1-\frac{n}{p}\right)|\Omega|^{\frac{1}{p}}}\right)\sum_{i=1}^n\|D_i f\|_{L^p(\Omega)}\end{aligned}$$

provided $\frac{n}{p} - 1 < 0$ and both x and z are in Ω. If Ω is a "regular" shape — a ball of radius h or a cube with side length h, for example, — then the diameter and the measure (= volume) of Ω are related: for some constant K,

$$\frac{\sup_{y\in\Omega}|y-x|}{\left(1-\frac{n}{p}\right)|\Omega|^{\frac{1}{p}}} \le K\frac{h}{\left(1-\frac{n}{p}\right)h^{\frac{n}{p}}} = K'h^{1-\frac{n}{p}}. \qquad (S_{20})$$

The constant K' depends on the geometry of Ω (ball or cube, for example), but not on the size of Ω. In particular, if we choose $h = |z-x|$, so the two points x and z lie on a diameter of a ball, Estimate (S_{18}) becomes

$$|f(z)-f(x)| \le 2K'|z-x|^{1-\frac{n}{p}}\sum_{i=1}^{n}\|D_i f\|_{L^p(\Omega)} \qquad \text{provided } \tfrac{n}{p}-1<0. \qquad (S_{21})$$

This is our last Theorem:

Theorem 5: If $-\lambda = \frac{n}{p} - 1 < 0$, and if all the partial derivatives of f belong to $L^p(R^n)$, then $f \in C^\lambda$.

Remarks: We conclude with some remarks about the embedding theorem.

Sobolev Spaces: We only proved Estimates (S_{13}) and (S_{21}) for $f \in C_0^1(R^n)$. Both estimates are of the form

$$\|f\|_Y \le KH\left(\|D_1 f\|_{L^p(R^n)}, \|D_2 f\|_{L^p(R^n)}, \dots, \|D_n f\|_{L^p(R^n)}\right), \qquad (S_{22})$$

where H is either a product or a sum and Y is L^q or C^λ. Suppose $\{f_j\}$ is a sequence of functions in C_0^1 whose derivatives $\{D_i f_j\}$ are all Cauchy in $L^p(R^n)$. Estimate (S_{22}) implies that $\{f\}$ is Cauchy in Y, where Y is either L^q or C^λ, depending on the sign of $\frac{n}{p} - 1$. But Y is complete, so Estimates (S_{13}) and (S_{21}) actually hold in the case where the (distributional) derivatives of f belong to L^p. The space of functions on R^n whose first derivatives belong to L^p is the Sobolev Space $W^{1,p}(R^n)$. In symbols, then, the Sobolev Embedding Theorems say

$$W^{1,p}(R^n) \hookrightarrow \begin{cases} L^q(R^n) & \text{if } \frac{n}{p}-1=\frac{n}{q}>0, \text{ and} \\ C^\lambda(R^n) & \text{if } \frac{n}{p}-1=-\lambda<0. \end{cases}$$

This result is sometimes called **Sobolev's Lemma**.

Bounded Domains If $\Omega \subset R^n$ is a measurable set, then the closure of $C_0^1(\Omega)$ with respect to the "norm" $H\left(\|D_1 f\|_{L^p(R^n)}, \|D_2 f\|_{L^p(R^n)}, \dots, \|D_n f\|_{L^p(R^n)}\right)$ is denoted $W_0^{1,p}(\Omega)$. Sobolev's Lemma holds for Ω, too:

$$W_0^{1,p}(\Omega) \hookrightarrow \begin{cases} L^q(\Omega) & \text{or} \\ C_0^\lambda(\Omega). \end{cases}$$

The proof consists of extending any function in $C_0^1(\Omega)$ to all of R^3 by setting $f = 0$ outside Ω. Even if the result isn't quite C^1, the fundamental estimates from Lemma 1 and Theorems 2 and 5 still hold, and the proof goes through unchanged.

More Sobolev Spaces: Functions in $C_0^1(\Omega)$ have zero boundary values. In many boundary value problems, the boundary values are non-zero. If Ω is "nice enough", then the closure of $C^1(\Omega)$ with respect to the norm

$$\|f\|_{W^{1,p}(\Omega)} = \|f\|_{L^p(\Omega)} + \sum_i \|D_i f\|_{L^p(\Omega)}$$

is the Sobolev space $W^{1,p}(\Omega)$. It is a suitable space of functions with non-zero values on $\partial\Omega$. Sobolev embedding theorems for $W^{1,p}(\Omega)$ are complicated by the boundary $\partial\Omega$. You can see why, even in 3 dimensions: the Fundamental Theorem of Calculus has a boundary term,

$$f(x,y,z) = f(x_0,y,z) + \int_{x_0}^{x} D_1 f(t,y,z)\,dt,$$

where $x_0 = x_0(y,z)$ is the x-coordinate of the "left-most" point (x,y,z) on the boundary of Ω. The additional term on the right, $f(x_0,y,z)$, complicates the proof (slightly). We avoided the complication in the proof of Theorem 2 by eliminating the boundary — the boundary of R^3 is empty.

Counterexamples: When the domain Ω is bounded, we can frequently construct examples of non-embeddings using radial functions on regions with "cusps". As an example, consider a region in R^2 with a cusp at the origin defined in polar coordinates by $C = \{(r,\theta) \mid r < \theta^4 < 1\}$. The function

$$f(x,y) = \frac{1}{(x^2+y^2)^{\frac{1}{4}}} = \frac{1}{\sqrt{r}}$$

and its partial derivatives

$$D_1 f = \frac{-x}{2\,(x^2+y^2)^{\frac{5}{4}}} = \frac{-\cos(\theta)}{2r^{\frac{3}{2}}}$$

$$D_2 f = \frac{-y}{2\,(x^2+y^2)^{\frac{5}{4}}} = \frac{-\sin(\theta)}{2r^{\frac{3}{2}}}$$

all belong to $L^3(C)$, but f is unbounded even though $\frac{n}{p} - 1 = \frac{2}{3} - 1 < 0$.

The Cone Condition: Estimate (S_{21}) requires the ball that we called Ω to lie in the domain of f. If the domain of f (which is no longer Ω, I'm sorry to say, since Ω is the ball) has a boundary, a ball might not "fit" in the domain. In this case, a hemi-ball might fit, and the embedding would still hold, with a different K'. If a hemi-ball doesn't fit, maybe a cone would fit. Then the embedding would still hold. A domain in which every point x is the vertex of a (fixed) cone which (after rotation and translation) lies in the domain is said to satisfy the **cone property**. Lots of Sobolev embeddings require the domain satisfy the cone property.

Boundary Spaces:: If Ω is a "nice" subset of R^3, then its boundary $\partial\Omega$ is a "nice" 2-dimensional subset of R^3. Consider a 2-dimensional plane in R^3 — the x-y plane, for example. If $f \in W^{1,p}(R^3)$, there is no reason, *a priori*, to believe that f is even well-defined on the plane. After all, the plane has zero measure (= 3-dimensional volume), and "functions" in $L^p(R^3)$ are really equivalence classes of functions equal to each other almost everywhere. Fortunately, if $f \in W^{1,p}(R^3)$, then f restricted to the plane belongs to $L^q(R^2)$ where

$$\frac{2}{q} = \frac{3}{p} - 1.$$

(This is assuming $\frac{3}{p} - 1 > 0$, of course. If $\frac{3}{p} - 1 < 0$, then f is (Hölder) continuous, so the restriction of f to the plane is (Hölder) continuous as well.)

Sobolev's Condition: Now we can say why Equation (S_{14}) is the "right" way to write Sobolev's condition. The space of functions f which, together with all partial derivatives up to order m belong to $L^p(\Omega)$ is denoted $W^{m,p}(\Omega)$. Suppose Ω^k is a k-dimensional subset of Ω — for example, $\partial\Omega$ is likely an $(n-1)$-dimensional subset of R^n, and we set $k = n-1$. If the domains Ω and Ω^k are "nice enough", then

$$W^{m,p}(\Omega) \hookrightarrow L^q(\Omega^k) \qquad \text{if} \qquad \frac{k}{q} = \frac{n}{p} - m > 0.$$

This is the most general of Sobolev's Embeddings into L^q, and it is easy to remember in this form. (Far easier, say, than the usual form $q = \frac{kp}{n-mp}$.) Just remember "dimension (k or n) divided by power (q or p)" counts as a derivative (m). In this form, Sobolev's condition resembles Hölder's condition (H_1).