Metrics

A metric is a way of measuring the distance between objects in a set.

Given a set S, a metric on S is a function d from $S \times S$ to $\mathbb{R}$ such that for all $x, y, z \in S$,

$$(i) \quad d(x,x) = 0;$$

$$(ii) \quad d(x,y) > 0, x \neq y;$$

$$(iii) \quad d(x,y) = d(y,x);$$

$$(iv) \quad d(x,z) \leq d(x,y) + d(y,z)$$

For most metrics, the first three properties are usually obviously satisfied.
The fourth property, known as the *Triangle Inequality*, commonly requires a bit more effort to verify.

Examples:

The following functions are metrics on the stated sets:

1. $S = \mathbb{R}$;

$$d(x,y) = |x - y| \ .$$

(i) $d(x,x) = |x - x| = |0| = 0$
(ii) $|x - y| \geq 0$, and $|x - y| = 0$ if and only if $x - y = 0$; that is $x = y$.
(iii) $d(y,x) = |y - x| = |x - y| = d(x,y)$.
(iv) For any real number, $x \leq |x|$.
Therefore

$$\begin{aligned} |x+y|^2 &= (x+y)^2 \\ &= x^2 + 2xy + y^2 \\ &\leq |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2 \end{aligned}$$

Since $|x| + |y| \geq 0$, we have

$$|x + y| \leq |x| + |y|$$

and hence

$$|x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| \ .$$

2. $S = \mathbb{R}^n$;

$$\begin{aligned} d(\underline{x}, \underline{y}) &= \left(\sum_{i=1}^{n}(x_1 - y_i)^2\right)^{1/2} \\ &= \big((\underline{x} - \underline{y}).(\underline{x} - \underline{y})\big)^{1/2} \\ &= ||\underline{x} - \underline{y}|| \end{aligned}$$

This is usually referred to as the Euclidean metric.

To prove (iv), consider

$$(\underline{x} + t\underline{y}).(\underline{x} + t\underline{y}) = ||\underline{x}||^2 + 2t(\underline{x}.\underline{y}) + t^2||\underline{y}||^2$$

where t is a real variable.

Since this quadratic expression in t is always ≥ 0, the discriminant

$$4||\underline{x}||^2\,||\underline{y}||^2 - 4(\underline{x}.\underline{y})^2$$

is non-negative.

Therefore

$$\begin{gathered} \underline{x}.\underline{y} \leq ||\underline{x}||\,||\underline{y}|| \\ ||\underline{x} + \underline{y}||^2 = ||\underline{x}||^2 + 2\underline{x}.\underline{y} + ||\underline{y}||^2 \\ \leq ||\underline{x}||^2 + 2||\underline{x}||\,||\underline{y}|| + ||\underline{y}||^2 = \big(||\underline{x}|| + ||\underline{y}||\big)^2 \\ ||\underline{x} + \underline{y}|| \leq ||\underline{x}|| + ||\underline{y}|| \end{gathered}$$

and

$$\begin{aligned} d(x, z) &= ||\underline{x} - \underline{z}|| \\ &= ||\underline{x} - \underline{y} + \underline{y} - \underline{z}|| \\ &\leq ||\underline{x} - \underline{y}|| + ||\underline{y} - \underline{z}|| \\ &\leq d(x, y) + d(y, z) \end{aligned}$$

3. $S = \mathbb{C}$;

$$d(z_1, z_2) = |z_1 - z_2|\ .$$

Since

$$|z_1 - z_2| = \big((x_1 - x_2)^2 + (y_1 - y_2)^2\big)^{1/2}$$

this is covered by the previous example.

4. $S = \mathbb{R}^n$;

$$d(\underline{x}, \underline{y}) = \sum_{i=1}^{n}|x_i - y_i|$$

This is sometimes known as the *taxi cab metric*.

Since

$$|x_i - z_i| = |x_i - y_i + y_i - z_i| \le |x_i - y_i| + |y_i - z_i|$$
$$d(\underline{x}, \underline{z}) \le d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z}) \ .$$

5. $S = \mathbb{R}^n$;

$$d(\underline{x}, \underline{y}) = \max_i |x_i - y_i|$$

This is known as the sup metric.

$$\begin{aligned} d(\underline{x}, \underline{z}) = & \ |x_j - z_j| \quad \text{for some } j \\ & \le |x_j - y_j| + |y_j - z_j| \\ & \le d(\underline{x}, \underline{y}) + d(\underline{y}, \underline{z}) \end{aligned}$$

6. S arbitrary;

$$d(x, y) = 1, x \neq \ y$$
$$d(x, x) = 0 \ .$$

(This is known as the discrete metric.)

To prove the triangle inequality, we note that if $z = x$,

$$d(x, z) = 0 \le d(x, y) + d(y, z)$$

for any choice of y,
while if $z \neq x$ then either $z \neq y$ or $x \neq y$ (at least) so that

$$d(x, y) + d(y, z) \ge 1 = d(x, z)$$

7. S is the set of all real continuous functions on $[a, b]$.

$$d(f, g) = \left(\int_a^b (f(x) - g(x))^2 dx \right)^{\frac{1}{2}} .$$

This is the continuous equivalent of the Euclidean metric in $\mathbb{R}^n$.

The proof of the triangle inequality follows the same form as in that case.

8. S as in 7.

$$d(f, g) = \max_{a \le x \le b} \ |f(x) - g(x)|.$$

This is the continuous equivalent of the sup metric.

The proof of the triangle inequality is virtually identical.

9. S arbitrary;
$d(x,y) = D(x,y)/(1+D(x,y))$, where $D(x,y)$ is another metric on S.

$$
\begin{aligned}
& d(x,y) + d(y,z) - d(x,z) \\
&= \frac{D(x,y)(1+D(y,z)+D(x,z)+D(y,z)D(x,z))}{(1+D(x,y))(1+D(y,z))(1+D(x,z))} \\
&+ \frac{D(y,z)(1+D(x,y)+D(x,z)+D(x,y)D(x,z))}{(1+D(x,y))(1+D(y,z))(1+D(x,z))} \\
&- \frac{D(x,z)(1+D(y,z)+D(x,y)+D(y,z)D(x,y))}{(1+D(x,y))(1+D(y,z))(1+D(x,z))} \\
&= \frac{D(x,y)+D(y,z)-D(x,z)}{(1+D(x,y))(1+D(y,z))(1+D(x,z))} \\
&+ \frac{2D(x,y)D(y,z)+D(x,y)D(y,z)D(x,z)}{(1+D(x,y))(1+D(y,z))(1+D(x,z))} \\
&\geq 0
\end{aligned}
$$

since the denominator is positive, and the numerator is non-negative.

10. Let f be a twice differentiable real function defined for $x \geq 0$, and such that

(a) $$f(0) = 0$$

(b) $$f'(x) > 0 \ x \geq 0$$

(c) $$f''(x) \leq 0 \ x \geq 0$$

Then if $d(x,y)$ is a metric on the set S, so is $f(d(x,y))$.

Since $f'(x) > 0$, we have from the Mean Value Theorem

$$
\begin{aligned}
f(x) - f(y) &= f'(\theta)(x-y) \\
f(x) &> f(y)\ ,\ x > y\ .
\end{aligned}
$$

In particular, since $f(0) = 0$,

$$f(x) > 0\ ,\ x > 0\ .$$

If $a \geq 0$ and $s \geq 0$, then by the Mean Value Theorem we also have

$$
\begin{aligned}
f'(a+s) - f'(s) &= f''(\theta)s \leq 0 \\
f'(a+s) &\leq f'(s)
\end{aligned}
$$

and if $b \geq 0$ also

$$
\begin{aligned}
\int_0^b f'(a+s)\,ds &\leq \int_0^b f'(s)\,ds \\
f(a+b) - f(a) &\leq f(b) - f(0) = f(b) \\
f(a+b) &\leq f(a) + f(b)
\end{aligned}
$$

Since $d(x,z) \leq d(x,y) + d(y,z)$,

$$
\begin{aligned}
f(d(x,z)) &\leq f(d(x,y) + d(y,z)) \\
&\leq f(d(x,y)) + f(d(y,z)
\end{aligned}
$$

Three important inequalities

Definitions:
If p and q are both > 1, we say that they are *dual indices* if

$$\frac{1}{p} + \frac{1}{q} = 1 \ .$$

For $\underline{x} \in \mathbb{R}^n$, we define

$$||\underline{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Young's inequality.
For x and y positive real numbers, and dual indices p and q,

$$xy \leq \frac{1}{p} x^p + \frac{1}{q} y^q \ .$$

Proof
Fix $y > 0$, and consider the function

$$f(x) = xy - \frac{1}{p} x^p$$

for $x > 0$.

$$f'(x) = y - x^{p-1} \ ; \ f''(x) = -(p-1)x^{p-2} < 0$$

Therefore f has a global maximum when $y = x^{p-1}$, $x = y^{1/(p-1)}$, and

$$xy - \frac{1}{p} x^p \leq y^{1+1/(p-1)} - \frac{1}{p} y^{p/(p-1)} = \frac{1}{q} y^q$$

Hölder's inequality
For $\underline{x}$ and $\underline{y}$ in $\mathbb{R}^n$,

$$\sum_{i=1}^{n} |x_i y_i| \leq ||\underline{x}||_p \, ||\underline{y}||_q$$

where p and q are dual indices.
Proof
If either $\underline{x} = \underline{0}$ or $\underline{y} = \underline{0}$ the result is trivially true.
Otherwise, for each i

$$\frac{|x_i|}{||\underline{x}||_p} \geq 0, \frac{|y_i|}{||\underline{y}||_q} \geq 0$$

and therefore by Young's inequality

$$\frac{|x_i||y_i|}{||\underline{x}||_p ||\underline{y}||_q} \leq \frac{1}{p} \frac{|x_i|^p}{||\underline{x}||_p^p} + \frac{1}{q} \frac{|y_i|^q}{||\underline{y}||_q^q}$$

Summing from $i = 1$ to n, we have

$$\frac{1}{||\underline{x}||_p ||\underline{y}||_q} \sum_{i=1}^{n} |x_i y_i| \leq \frac{1}{p} + \frac{1}{q} = 1 \ .$$

Minkowski's inequality
For $\underline{x}$ and $\underline{y}$ in $\mathbb{R}^n$, and $p > 1$,

$$||\underline{x} + \underline{y}||_p \le ||\underline{x}||_p + ||\underline{y}||_p \ .$$

$$\begin{aligned}
||\underline{x} + \underline{y}||_p^p &= \sum_{i=1}^{n} |x_i + y_i| \, |x_i + y_i|^{p-1} \\
&\le \sum_{i=1}^{n} |x_i||x_i + y_i|^{p-1} + \sum_{|i=1}^{n} |y_i||x_i + y_i|^{p-1} \\
&\le ||\underline{x}||_p \left(\sum_{i-1}^{n} |x_i + y_i|^{(p-1)q} \right)^{1/q} \\
&\quad + ||\underline{y}||_p \left(\sum_{i-1}^{n} |x_i + y_i|^{(p-1)q} \right)^{1/q} \\
&= \big(||\underline{x}||_p + ||\underline{y}||_p\big) \ \left[\left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{1/p} \right]^{p/q} \\
&= \big(||\underline{x}||_p + ||\underline{y}||_p\big) \, ||\underline{x} + \underline{y}||_p^{p-1}
\end{aligned}$$

Minkowski's inequality shows that

$$d(\underline{x}, \underline{y}) = ||\underline{x} - \underline{y}||_p$$

is a metric on $\mathbb{R}^n$.

If we consider the limits $p \to 1$ and $p \to \infty$, we obtain the taxicab and sup metrics which we have already seen.
The Euclidean metric corresponds to $p = 2$.

These metrics can be extended to infinite real sequences

$$\{x_i\}_{i=1}^{\infty} \ .$$

If l^p is the set of all sequences $\{x_i\}$ for which

$$\sum_{i=1}^{\infty} |x_i|^p \ \ ; \ p \ge 1$$

converges, then the norm

$$||\underline{x}||_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}$$

induces a metric on l^p in the same way as the finite form does on $\mathbb{R}^n$.

When $p = 1$, l^1 is the set of all sequences with absolutely convergent sums.
For $p \to \infty$, we take

$$||\underline{x}||_\infty = \sup_i |x_i|$$

and l^∞ is the set of all bounded sequences.

Note that the harmonic sequence $\{x_n = \frac{1}{n}\}$ is not in l^1, but is in l^p for all $p > 1$.

We can also extend these metrics to the continuous case.

For the set of functions continuous on $[a, b]$, we have the metrics

$$d_p(f, g) = ||f - g||_p = \left(\int_a^b |f(x) - g(x)|^p \, dx \right)^{1/p}$$

for $p \geq 1$.

Sequences and Convergence

In what follows, it is assumed that we have some set S with an associated metric d. The combination (S, d) is called a metric space.

Definition: A sequence $\{a_n\}$ converges in (S, d) if and only if there is $l \in S$ such that, for each $\epsilon > 0$ there is a positive integer N such that for **all** $n \geq N$ we have $d(a_n, l) < \epsilon$.

Note that convergence depends on both the set S, since the limit must belong to S, and on the metric d.

For example, if d is the discrete metric, a sequence converges if and only if all the terms from some point on are the same.

Definition: An *epsilon neighbourhood* of $a \in S$ is the set of all $x \in S$ such that $d(x, a) < \epsilon$.

(Note that this implies that $\epsilon > 0$.)

Definition: A set $Q \subset S$ is called a *neighbourhood* of a if Q contains an epsilon neighbourhood of a.

Theorem: A sequence $\{a_n\}$ converges to l iff every neighbourhood of l contains all but a finite number of terms of the sequence.

Suppose Q is a neighbourhood of l. Then for some $\epsilon > 0$, the set $\{d(x, l) < \epsilon\} \subset Q$.

Since $\{a_n\}$ converges to l, there is an integer N such that $d(a_n, l) < \epsilon$ for all $n > N$.

Therefore $a_n \in Q$ for all $n > N$.

Conversely, if every neighbourhood of l contains all but a finite number of terms of the sequence, then for any $\epsilon > 0$ this is true for the epsilon neighbourhood $\{d(x, l) < \epsilon\}$.

That is, there is some integer N such that if $n > N$, $d(a_n, l) < \epsilon$.

Theorem: If $\{a_n\}$ converges, then the limit is unique.

Suppose that the sequence converges to l_1. Choose any other $l_2 \neq l_1$, and let $\epsilon = \frac{1}{2} d(l_1, l_2) > 0$.

Since l_1 is a limit for the sequence, all but a finite number of the terms of the sequence are in the neighbourhood $\{d(x, l_1) < \epsilon\}$.

Now, for each a_n in the sequence,

$$d(l_1, l_2) \leq d(l_1, a_n) + d(a_n, l_2)$$
$$d(l_2, a_n) \geq d(l_1, l_2) - d(a_n, l_1)$$

so that

$$d(l_2, a_n) > \epsilon$$

for all but a finite number of terms in the sequence.

Since the epsilon neighbourhood of l_2 contains at most a finite number of terms of the sequence, l_2 cannot be a limit for the sequence.

Definition: A set $Q \subset S$ is said to be bounded if for any $a \in Q$, there is a real number k such that $d(x, a) < k$ for all $x \in Q$.

Boundedness depends on both the set Q and the metric d.

For example, in $\mathbb{R}$ with the standard metric $d(x, y) = |x - y|$, a set Q is bounded if there is some constant k such that $|x| < k$ for all $x \in Q$.

On the other hand, if we choose the metric

$$d(x, y) = \frac{|x - y|}{1 + |x - y|} \ ,$$

all sets are bounded, since $d(x, y) < 1$ for all x, y.

Theorem: If $\{a_n\}$ converges to l, then $\{a_n\}$ is bounded.

Given $\epsilon = 1$, there is an integer N such that

$$d(a_n, l) < 1 \ , \text{ for all } n > N \ .$$

For any $a \in S$, consider the finite set of real numbers

$$d(a_1, a), d(a_2, a), \ldots, d(a_N, a), (d(l, a) + 1) \ .$$

Since the set is finite, it contains a maximum element m.

Choose $k > m$.

By construction $d(a_n, a) < k$ for $n \leq N$, while if $n > N$,

$$d(a_n, a) \leq d(a_n, l) + d(l, a) < 1 + d(l, a) < k \ .$$

Definition: A sequence $\{a_n\}$ is *Cauchy* iff for each $\epsilon > 0$ there is a positive integer N such that if $m, n \geq N$, then

$$d(a_n, a_m) < \epsilon.$$

As before, a convergent sequence is Cauchy.

Our concern is in determining under what conditions a Cauchy sequence is convergent.

If every Cauchy sequence in (S, d) converges, then we say that the metric space is *complete*.

Consider the set S of real functions, continuous on $[0,1]$.

We will make this a metric space by considering the metrics 7 and 8. and consider the sequence

$$\{x^n\}, 0 \leq \ x \leq \ 1$$

This is a Cauchy sequence with respect to the metric defined in 7.

Suppose, without loss of generality, that $m > n$.

$$\begin{aligned}\int_0^1 (x^m - x^n)^2 \, dx &= \int_0^1 (x^{2m} - 2x^{m+n} + x^{2n}) \, dx \\ &= \frac{1}{2m+1} - \frac{2}{m+n+1} + \frac{1}{2n+1} \\ &< \frac{1}{2n+1} + \frac{1}{2n+1} = \frac{2}{2n+1} < \frac{1}{n}\end{aligned}$$

Therefore

$$d(x^m, x^n) < \frac{1}{\sqrt{n}} < \epsilon \text{ for all } m > n > \left[\frac{1}{\epsilon^2}\right] \ .$$

However, the limit of this sequence is the function

$$f(x) = \begin{cases} 0 \ , \ 0 \leq x < 1 \\ 1 \ , \ x = 1 \end{cases}$$

which is not continuous, and therefore not in S.

Therefore this Cauchy sequence does not converge in (S, d).

On the other hand, $\max_{0\leq x\leq 1} |x^n - x^{2n}| = 1/4$, and hence $\{x^n\}, 0 \leq x \leq 1$, is not a Cauchy sequence with respect to the sup metric defined in 8.

$$\frac{d}{dx}(x^n - x^{2n}) = nx^{n-1} - 2nx^{2n-1} = 0$$

when $x^n = \frac{1}{2}$, which is a maximum.

Since $x^n - x^{2n} = \frac{1}{4}$, $d(x^n, x^{2n}) = \frac{1}{4}$.

Hence there is no N such that

$$d(x^m, x^n) < \frac{1}{8} \text{ for all } m > n > N \ .$$