

Basic setup and conventions

We work with a Riemannian manifold (M, g) with ∇ denoting the Levi-Civita connection.

Vector fields will be denoted X, Y, Z, W, \dots . Most of the time we will abuse notation and write X, Y, Z, W, \dots for tangent **vectors**. We do this because many tensors, such as the curvature tensor are defined for vector fields, are tensorial so at a fixed point only depend on the value of the vector fields in question at that fixed point.

The Curvature Tensor

Let X, Y, Z be vector fields. Define a new vector field by

$$\text{Rm}(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z.$$

Notice that $\nabla_X \nabla_Y Z$ will include the variation of Y along X - namely $\nabla_X Y$. This is undesirable since we want to measure the curvature of the space itself at each point using **Rm**, and this should not depend on how any particular vector field varies. Likewise for $\nabla_Y \nabla_X Z$. The term $\nabla_{[X, Y]} Z$ compensates precisely for this undesirable effect.

Another way of expressing this compensation is to say that **Rm** is **tensorial** in X, Y so that for any smooth function $f \in C^\infty(M)$ we have

$$\text{Rm}(fX, Y)Z = f \text{Rm}(X, Y)Z = \text{Rm}(X, fY)Z.$$

Exercise: Using the Leibniz rule for the connection ∇ and the corresponding rule for the Lie bracket, prove the claimed tensoriality in X, Y .

As a consequence, although as written, **Rm** is defined for vector **fields**, tensoriality induces a well defined map defined on tangent vectors. As mentioned in setup and conventions, we will typically not differentiate by vector fields and tangent vectors when dealing with tensorial equations. But **just this time**, let us be very explicit: Let $X, Y, Z \in T_x M$ be tangent vectors, let $\bar{X}, \bar{Y}, \bar{Z}$ and $\tilde{X}, \tilde{Y}, \tilde{Z}$ be vector fields defined on a neighbourhood of x such that

$$\bar{X}(x) = X, \bar{Y}(x) = Y, \bar{Z}(x) = Z$$

$$\tilde{X}(x) = X, \tilde{Y}(x) = Y, \tilde{Z}(x) = Z.$$

Then tensoriality implies that

$$(\text{Rm}(\bar{X}, \bar{Y})\bar{Z})(x) = (\text{Rm}(\tilde{X}, \tilde{Y})\tilde{Z})(x).$$

Thus we may define unambiguously,

$$\text{Rm}(X, Y)Z = (\text{Rm}(\bar{X}, \bar{Y})\bar{Z})(x)$$

where $\bar{\cdot}$ denotes any arbitrary extension of X, Y, Z . Tensoriality then guarantees the result is independent of the extension.

What is rather more surprising, given that X is being differentiated twice, is that **Rm** is tensorial in Z also! This means that **Rm** may be evaluated on tangent vectors X, Y, Z at a point and thus may be interpreted as giving information (via ∇ which itself is determined by g) about (M, g) at a point. This information is in fact a measure of curvature.

One question stands out: **Why is $\nabla_{[X,Y]}Z$ the right correction term?** There are a few ways we might answer this question such as "because it works!" and "check in coordinates". The answer we will give here is obtained by interpreting Rm as the **commutator of second derivatives**.

Second Derivative

The second derivative of a vector field, in directions X, Y is defined to be

$$\nabla_{X,Y}^2 Z := \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z) = \nabla_X(\nabla_Y Z) - \nabla_{\nabla_X Y} Z.$$

Exercise: Check that $\nabla_{X,Y}^2 Z$ is tensorial in X, Y .

The reason for this definition is that once again, $\nabla_X(\nabla_Y Z)$ will include the variation, $\nabla_X Y$ of Y along X so we must subtract it off so that it doesn't contribute to $\nabla^2 Z$. Essentially the way to understand how to choose what to subtract off is by the product rule. First, for those more comfortable with coordinates, we have

$$\nabla_Y Z = Y^i \partial_i Z^j \partial_j + Y^i Z^j \Gamma_{ij}^k \partial_k.$$

This looks pretty good: we are differentiating Z in the direction Y and the result depends only on Y, Z and the first derivatives of Z . Now we apply ∇_X :

$$\nabla_X \nabla_Y Z = X^\ell \partial_\ell (Y^i \partial_i Z^j) \partial_j + X^\ell Y^i \partial_i Z^j \Gamma_{\ell j}^m \partial_m + \dots$$

where I got tired of computing this way so I just put \dots to indicate there are more terms! The point though is that there are derivatives of Y^i in there but we really only want to compute the variation of Z . In particular notice that applying the product rule will give a term

$$X^\ell \partial_\ell Y^i \partial_i Z^j \partial_j$$

which we recognise as the first term occurring in

$$\nabla_{\nabla_X Y} Z = X^\ell \partial_\ell Y^i \partial_i Z^j \partial_j + \dots$$

If one is so inclined, this computation may be fully carried out to verify that the result only depends on the components X^i, Y^j, Z^k and the first two derivatives of Z : $\partial_i Z^k, \partial_i \partial_j Z^k$.

Exercise: Carry out the computation if you are so inclined.

The Hessian of a function

For comparison, consider the hessian matrix of a real valued function defined on \mathbb{R}^n :

$$d^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1}(x) & \cdots & \frac{\partial^2 f}{\partial x^1 \partial x^n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x^n \partial x^1}(x) & \cdots & \frac{\partial^2 f}{\partial x^n \partial x^n}(x) \end{pmatrix}$$

This matrix records how f varies to second order at x . Once this matrix has been computed, second derivatives of f in directions $X = (X^1, \dots, X^n)$ and $Y = (Y^1, \dots, Y^n)$ may be computed as

$$d^2 f(X, Y) = Y^T d^2 f X.$$

However, if X, Y are vector fields, then in general,

$$d^2 f \neq \partial_X(\partial_Y f)$$

where

$$\partial_X f = df(X)$$

or equivalently $\partial_X f = X(f)$ with X acting as a derivation. The problem is of course again the fact that Y will also be differentiated:

$$\partial_X(\partial_Y f) = X^i \partial_i (Y^j \partial_j f) = X^i Y^j \partial_i \partial_j f + X^i \partial_i Y^j \partial_j f = d^2 f(X, Y) + df(D_X Y)$$

so that

$$d^2 f(X, Y) = \partial_X(\partial_Y f) - df(D_X Y) = \partial_X(\partial_Y f) - \partial_{D_X Y} f.$$

Now the point of tensorality is that just from the matrices for $d^2 f$ and df at a point x , the second derivative $\partial_X(\partial_Y f)$ at x may be computed by linear algebra alone (i.e. matrix multiplication) with no further differentiation required. **This is because of tensorality: $d^2 f(X, Y)$ only depends on the value of X, Y at the point x and not in a neighbourhood.** In other words, we may pre-compute the matrices df and $d^2 f$ once and for all, then apply them to any vectors to compute first and second derivatives. We may also approximate f to second order at any point without needing to compute any more derivatives.

As a simple comparison, this idea is essentially used by a calculator (or computer) to compute \sin, \cos, \exp etc. The Taylor series is calculated once and for all (giving an expression for the coefficients that can be calculate easily or by storing in a table sufficiently many of the coefficients) and then hard wired into the calculator. Further calculation is by elementary arithmetic operators.

Thus the moral is to compute the maps $x \mapsto df(x)$ and $x \mapsto d^2 f(x)$ from which any second derivatives may be later computed using linear algebra. This only works by using the tensorial first and second derivatives so we may later work pointwise!

Tensoriality of second derivatives

Now the definition of $d^2 f$ should be compared immediately with the definition of $\nabla^2 Z$. Formally, it is the same thing just with f replaced by Z and D replaced by ∇ . This is suggestive that we have the correct expression for $\nabla^2 Z$.

Let us now rephrase the expression for $\nabla^2 Z$ and see how the tensorality arises.

The first observation is that ∇Z is an endomorphism of TM . That is an element of

$$\text{Hom}(TM, TM) \simeq T^*M \otimes TM.$$

Then we may interpret $\nabla Z(X) = \nabla_X Z$ in terms of contractions (traces) and tensor products:

$$\nabla Z(X) = \text{Tr} \nabla Z \otimes X$$

where the trace is taken by contracting the T^*M part of ∇Z with X . Notice in particular for so-called **indecomposable** elements of $T^*M \otimes TM$, namely those of the form $\alpha \otimes X$ with α a one-form we have

$$\text{Tr} \alpha \otimes X = \alpha(X).$$

Now we'd like to be able to differentiate α . As before, if we differentiate the function $\alpha(X)$ we will pick up derivatives of both α and X . So to isolate the derivative of α we could subtract off the derivative of X . Then we make the definition

$$\nabla \alpha(X, Y) = \partial_X(\alpha(Y)) - \alpha(\nabla_X Y).$$

Exercise: Check this is tensorial in X and Y .

In terms of tensor products and traces we may express the definition as

$$\partial_X(\alpha(Y)) = \partial_X \text{Tr}(\alpha \otimes Y) = \text{Tr}(\nabla_X \alpha) \otimes Y + \text{Tr} \alpha \otimes \nabla_X Y = \nabla_X \alpha(Y) + \alpha(\nabla_X Y).$$

Given a connection ∇ on TM and the (uniquely determined by identifying vector fields with derivations) connection on $M \times \mathbb{R}$, we may define a unique connection on T^*M by requiring that the resulting three connections commute with traces and satisfy the Leibniz rule for the tensor product.

Now how do we differentiate ∇Z ? It is an endomorphism and we may do something similar for endomorphisms. So let T be an endomorphism so that $T(X)$ is a vector field. Note that for one-forms α we had $\alpha(X)$ is a function and we know how to differentiate functions. Well, given ∇ we also know how to differentiate vector fields suggesting that we define

$$(\nabla_X T)(Y) = \nabla_X(T(Y)) - T(\nabla_X Y).$$

In terms of traces

$$\nabla_X(T(Y)) = \nabla_X(\text{Tr} T \otimes Y) = \text{Tr} \nabla_X T \otimes Y + \text{Tr} T \otimes \nabla_X Y = \nabla_X T(Y) + T(\nabla_X Y).$$

Rearranging gives

$$(\nabla_X T)(Y) = \nabla_X(T(Y)) - T(\nabla_X Y).$$

Exercise: Check directly that this is tensorial in both X and Y . Do it both with the final expression and with the identities using traces and tensor products. Think about how requiring that the connection commutes with traces and satisfies the Leibniz product rule for tensor products leads to tensoriality.

Then for $T = \nabla Z$ we finally obtain

$$\nabla_{X,Y}^2 Z = \nabla^2 Z(X, Y) = (\nabla_X \nabla Z)(Y) = \nabla_X(\nabla Z(Y)) - \nabla Z(\nabla_X Y) = \nabla_X \nabla_Y Z - \nabla_{\nabla_X Y} Z$$

which is tensorial in both X and Y .

Ricci Identities and tensorality of second derivatives

Now that we understand second derivatives, we can express the curvature tensor Rm as the commutator of second derivatives:

$$\text{Rm}(X, Y)Z = \nabla_{X,Y}^2 Z - \nabla_{Y,X}^2 Z.$$

This equation is known as the **Ricci Identity**.

Exercise: Prove the Ricci Identity. **Hint:** Use the fact that ∇ is torsion-free $\nabla_X Y - \nabla_Y X = [X, Y]$.

Sometimes this expression is written

$$[\nabla_X, \nabla_Y]Z = \nabla_{X,Y}^2 Z - \nabla_{Y,X}^2 Z.$$

Be **careful** with this phrasing: $[\nabla_X, \nabla_Y]Z \neq \nabla_X(\nabla_Y Z) - \nabla_Y(\nabla_X Z)$! The right hand side is not tensorial.

Exercise: Define $\text{Rm}(X, Y)f = \nabla_{X,Y}^2 f - \nabla_{Y,X}^2 f$. Show that $\text{Rm}(X, Y)f = 0$. Equivalently, $\nabla^2 f(X, Y) = \nabla^2 f(Y, X)$. We might then say that $M \times \mathbb{R} \rightarrow M$ is a flat (i.e. not curved!) vector bundle.

Thus the curvature tensor measures the lack of commutativity of second derivatives of vector fields. Put another way, unlike for functions, $\nabla_{X,Y}^2 Z$ need not be symmetric. Instead we have

$$\nabla_{X,Y}^2 Z = \nabla_{Y,X}^2 Z + \text{Rm}(X, Y)Z.$$

Exercise: Show that in Euclidean space, $\nabla_{X,Y}^2 Z$ is symmetric in X, Y .

Now we observe that since we defined $\nabla^2 Z$ in a tensorial way, immediately we have $\text{Rm}(X, Y)Z$ is tensorial in X, Y . By defining Rm as the second order commutator, we also immediately obtained the correction term.

But still, we have the question **why is Rm tensorial in Z ?**

Exercise Show that

$$\nabla_X \nabla_Y f Z - \nabla_Y \nabla_X f Z - \nabla_{[X,Y]} f Z = f \text{Rm}(X, Y)Z + (\text{Rm}(X, Y)f)Z = f \text{Rm}(X, Y)Z.$$

Thus we conclude the tensorality in Z follows since $M \times \mathbb{R} \rightarrow M$ is a flat vector bundle.