THE SECOND DERIVATIVE TEST ON RIEMANNIAN MANIFOLDS

KEATON QUINN

With the added structure of a Riemannian metric g on a smooth manifold M we can extend the gradient, the Hessian, and the second derivative test to the curved setting.

Definition. *Let $f : M \to \mathbb{R}$ be a smooth function. The gradient of f is the smooth vector field ∇f dual to the 1-form df. That is, for all vector fields X on M,*

$$df(X) = g(\nabla f, X).$$

Just like in $\mathbb{R}^n$, the gradient is perpendicular to level sets.

Lemma. *Let $c \in \mathbb{R}$ be a regular value of f. Then ∇f is orthogonal to $f^{-1}(c)$.*

Proof. Since c is a regular value, $f^{-1}(c)$ is a smooth submanifold of M with tangent space at a point $T_p f^{-1}(c) = \ker df_p$. Let v be tangent to $f^{-1}(c)$ at p. Then

$$g(\nabla f, v) = df_p(v) = 0.$$

□

At noncritical points, the gradient also points in the direction of the greatest increase in f, as it does in the flat setting. Similarly, the negative of the gradient points in the direction of greatest decrease.

Proposition. *At a noncritical point p, the derivative df_p restricted to the unit tangent sphere U_pM at p has a maximum at $\frac{\nabla f}{\|\nabla f\|}$ and a minimum at $-\frac{\nabla f}{\|\nabla f\|}$.*

Proof. Consider $df_p : U_pM \to \mathbb{R}$. Since U_pM is compact, df_p has a maximum and a minimum. We can locate them using the derivative: Let $w \in T_v U_p M = v^\perp$, then

$$d(df_p)_v(w) = df_p(w)$$

since df_p is linear. If v is a critical point of df_p then $0 = d(df_p)_v = df_p(w)$ for all w perpendicular to v. Equivalently, v is a critical point of df_p if

$$g(\nabla f_p, w) = 0 \text{ for all } w \in v^\perp$$

Hence ∇f_p is orthogonal to $v^\perp$. Since $\nabla f_p \neq 0$ and since $v^\perp$ is $(n-1)$ dimensional, we must have ∇f in the span of v. So, the two critical points v are either

$$v = \frac{\nabla f}{\|\nabla f\|} \text{ or } v = -\frac{\nabla f}{\|\nabla f\|}.$$

But,

$$df_p\left(\pm\frac{\nabla f}{\|\nabla f\|}\right) = \pm g\left(\nabla f, \frac{\nabla f}{\|\nabla f\|}\right) = \pm\|\nabla f\|.$$

Hence, f is increasing most rapidly in the direction of the gradient, and decreasing most rapidly in the direction opposite the gradient. □

Date: Last revised July 14, 2020.

The Hessian of f is the symmetric 2-tensor given as the second covariant derivative of f, i.e., $\mathrm{Hess} f = \nabla(\nabla f)$ and is the natural notion of a second derivative of a function on a Riemannian manifold. Its action on vector fields X and Y is given by

$$\mathrm{Hess} f(X, Y) = X(Yf) - df(\nabla_X Y) = g(\nabla_X(\nabla f), Y).$$

Now, suppose p is a non-degenerate critical point of f. By this we mean $\mathrm{Hess} f_p \neq 0$. To determine whether $f(p)$ is a maximum or a minimum, we can check the signature of the Hessian. To prove this we first need a chain rule for the second derivative on a Riemannian manifold.

Lemma. *If $\gamma : U \to M$ is a smooth curve in M, then*

$$\frac{d^2}{dt^2}(f \circ \gamma)(t) = \mathrm{Hess} f_{\gamma(t)}(\gamma'(t), \gamma'(t)) + df_{\gamma(t)}\left(\nabla_{\gamma'(t)}\gamma'(t)\right)$$

Proof. The following takes place on the pullback bundle $\gamma^* TM$. Metric compatibility of the Levi-Civita connection ∇ says

$$dg(X, Y) = g(\nabla X, Y) + g(X, \nabla Y)$$

and so

$$\begin{aligned} \frac{d^2}{dt^2} f \circ \gamma = d(df_\gamma(\gamma')) = dg(\nabla f, \gamma') &= g(\nabla_{\gamma'}(\nabla f), \gamma') + g(\nabla f, \nabla_{\gamma'}\gamma') \\ &= \mathrm{Hess} f(\gamma', \gamma') + df(\nabla_{\gamma'}\gamma') \end{aligned}$$

□

Theorem (The Second Derivative Test)**.** *Suppose p is a non-degenerate critical point of f. Then f has a maximum at p if and only if* $\mathrm{Hess} f$ *is negative-definite. The function f has a minimum at p if and only if* $\mathrm{Hess} f$ *is positive-definite.*

Proof. Let γ be a curve in M through p. Since p is a critical point of f,

$$\left.\frac{d}{dt} f \circ \gamma\right|_{t=0} = df_p(\gamma'(0)) = 0$$

so 0 is a critical point of $f \circ \gamma$. We may then apply the second derivative test from single variable calculus using the following

$$\frac{d^2}{dt^2}(f \circ \gamma) = \mathrm{Hess} f(\gamma', \gamma') + df_{\gamma(t)}(\nabla_{\gamma'}\gamma').$$

When $t = 0$ this reads

$$\left.\frac{d^2}{dt^2} f \circ \gamma\right|_{t=0} = \mathrm{Hess} f(\gamma'(0), \gamma'(0)).$$

If the Hessian is negative definite, then

$$\left.\frac{d^2}{dt^2} f \circ \gamma\right|_{t=0} < 0$$

for all curves γ. This tells us $f \circ \gamma(0) = f(p)$ is a local maximum no matter the curve through p, and this is equivalent to $f(p)$ being a local maximum of f.

Conversely, if $f(p)$ is a local maximum, then $f \circ \gamma$ has negative second derivative at 0 for all curves γ though p. This is equivalent to $\mathrm{Hess} f(v, v) < 0$ for all $v \in T_pM$ as we may take a curve γ with $\gamma'(0) = v$. Hence, the Hessian is negative definite. The other case is similar. □

1. The second derivative on a Riemannian manifold

The double tangent bundle TTM to a Riemannian manifold splits as follows. Let $\pi : TM \to M$ be the projection. Define $V_{(p,v)}TM = \ker d\pi_{(p,v)}$. From the regular value theorem we see that

$$\ker d\pi_{(p,v)} = T_{(p,v)}\pi^{-1}(p) = T_{(p,v)}T_pM \simeq T_pM.$$

This space is called the vertical space at (p, v) and it the map which sends $y \in T_pM$ to $\tilde{y} \in V_{(p,v)}TM$ using this identification is called the vertical lift of x. For a smooth function $f : M \to \mathbb{R}$ we have the useful property $d(df)_{(p,v)}(\tilde{y}) = df_p(y)$.

The Riemannian metric on M singles out a complementary subspace $H_{(p,v)}TM$ called the horizontal space at (p, v). Let $x \in T_pM$ and let γ be a curve with $\gamma(0) = p$ and $\gamma'(0) = x$. Let $v(t)$ be the parallel transport of v along γ, so $\nabla_x v(0) = 0$. The call $\tilde{\gamma}(t) = (\gamma(t), v(t))$, this defines a curve in TM starting at (p, v) and $\tilde{\gamma}'(0)$ is a tangent vector in $T_{(p,v)}TM$. The horizontal space $H_{(p,v)}TM$ consists of all vector tangent to TM at (p, v) obtained in this way and we call the vector $\tilde{\gamma}'(0) =: \tilde{x}$ the horizontal lift of x. We have

$$T_{(p,v)}TM = H_{(p,v)}TM \oplus V_{(p,v)}TM$$

Now, let $f : M \to \mathbb{R}$ be a smooth function. Then $df : TM \to \mathbb{R}$ by $df(p, v) = df_p(v)$ is also a smooth function and so we can consider its derivative $d(df) : TTM \to \mathbb{R}$. We would like to understand this second derivative in terms of the splitting of the double tangent space into the horizontal and vertical space, and particularly in terms of horizontal and vertical lifts of vectors from T_pM.

To this end, let $w \in T_{(p,v}TM$ and decompose it as $w = \tilde{x} + \tilde{y}$ where $\tilde{x}$ is the horizontal lift of x and $\tilde{y}$ is the vertical lift of y. Then $d(df)_{(p,v)}(w) = d(df)_{(p,v)}(\tilde{x}) + d(df)_{(p,v)}(\tilde{y})$. We know that $d(df)_{(p,v)}(\tilde{y}) = df_p(y)$ by the property of a vertical lift of a vector. For the remaining factor, we know that $\tilde{x}$ is the tangent vector at zero to the curve $(\gamma(t), v(t))$. We can therefore compute

$$\begin{aligned} d(df)_{(p,v)}(\tilde{x}) = \frac{d}{dt}\Big|_{t=0} df(\gamma(t), v(t)) = \frac{d}{dt}\Big|_{t=0} df_{\gamma(t)}(v(t)) = \frac{d}{dt}\Big|_{t=0} g(\nabla f_{\gamma(t)}, v(t))_{\gamma(t)} \\ = g(\nabla_{\gamma'(t)}\nabla f_{\gamma(t)}, v(t))\big|_{t=0} + g(\nabla f_{\gamma(t)}, \nabla_{\gamma'(t)}v(t))\big|_{t=0} \\ = g(\nabla_x \nabla f_p, v)_p + g(\nabla f_p, \nabla_x v(0))_p \\ = \nabla^2 f(v, x) \end{aligned}$$

since $\nabla_x v(0) = 0$.

Hence, if we consider instead $d(df)_{(p,v)} : T_pM \times T_pM \to \mathbb{R}$ then we have

$$d(df)_{(p,v)}(x, y) = \nabla^2 f(v, x)_p + df_p(y)$$

(note that the order of x and y matters).