CS 468, Lecture 11: Covariant Differentiation

Adrian Butscher
(scribe: Ben Mildenhall)

May 6, 2013

1 Introduction

We have talked about various extrinsic and intrinsic properties of surfaces. Extrinsic geometry is basically 'how a surface looks from the outside' (when embedded in $\mathbb{R}^3$, for example). Variations of the unit normal vector field and the second fundamental form and associated curvatures are extrinsic. They depend on the ambient space containing the surface.

Last time we talked about an *intrinsic* object: the induced metric $g_u = [D\phi_u]^T [D\phi_u]$. The induced metric allows us to calculate the Euclidean inner product of tangent vectors expressed in tangent space bases. Since it gives us the inner product, it can also be used to calculate intrinsic lengths. If we take some curve $c(t)$ defined for $t \in [0,1]$ and mapping into parameter space and then map it to a curve on the surface using our parametrization ϕ, we get some curve $\gamma = \phi \circ c$. Then using the induced metric, we have

$$\langle \dot\gamma(t), \dot\gamma(t) \rangle = [\dot c(t)]^\top g_{c(t)} [\dot c(t)]$$

and so the length of γ along the surface is

$$\int_0^1 \sqrt{\langle \dot\gamma(t), \dot\gamma(t) \rangle} dt = \int_0^1 \sqrt{[\dot c(t)]^\top g_{c(t)} [\dot c(t)]} dt$$

There's 'something missing' in our intrinsic geometry story so far. The geodesic equation:

$$\vec{k}_\gamma(t) \perp T_{\gamma(t)} S$$

tells us that if γ is a shortest path, then its $\vec{k}$ vector is perpendicular to the tangent plane (derived last week).

The loose end is that this equation looks entirely extrinsic. We need to show the geodesic equation is a second order ODE expressible in terms of g alone (so it is an *intrinsic* thing). (This is helpful because there's lots of theory for solving second order ODEs.)

2 Differentiation

2.1 Differentiation in Euclidean space

Let $V = [V^1, V^2, V^3]^\top$ be a vector in $T_p\mathbb{R}^3$ and let $c : I \to \mathbb{R}^3$ be a curve with $c(0) = p$ and $\dot{c}(0) = V$.

The derivative of a scalar function f in the direction of a vector V is given by

$$D_V f := \left.\frac{df(c(t))}{dt}\right|_{t=0} = \sum_{i=1}^{3} V^i \frac{\partial f}{\partial x^i} = \langle \nabla f(p), V \rangle$$

The derivative of a vector field $Y(x) := [Y^1(x), Y^2(x), Y^3(x)]^\top$ in the V direction is

$$D_V Y := \begin{bmatrix} D_V Y^1 \\ D_V Y^2 \\ D_V Y^3 \end{bmatrix} = \begin{bmatrix} \vdots \\ \sum_{i=1}^{3} V^i \frac{\partial Y^j}{\partial x^i} \\ \vdots \end{bmatrix}$$

2.2 Differentiation on a surface

We can differentiate scalar functions on a surface just as we do scalar functions on all of R^3. If we have some $f : S \to \mathbb{R}$ and want to find its derivative at $p \in S$ in the direction of some vector $V \in T_pS$, we can just take a curve $c : I \to S$ with $c(0) = p$, $\frac{dc(0)}{dt} = V$ and define $D_V f := \frac{d}{dt} f(c(t))|_{t=0}$.

However, you can't do the same thing for the derivative of a tangent vector field $Y : S \to TS$. The difference between two of the vectors in the vector field as distance between them goes to zero has no reason to lie tangent to the surface, so $\frac{d}{dt} Y(c(t))|_{t=0}$ will not necessarily be tangent to S.

So what are possible alternatives? We could try using a parameterization. Then we need to worry about independence from parameterization. A hint it will be difficult is that we will need to differentiate the coordinate vectors $E_i := D\phi(\frac{\partial}{\partial u^i})$. In the parameter domain, we have the same coordinate system everywhere, but if you push that basis forward, it will be different at each point, so the vector field Y would have to be expressed in that different basis at each point.

3 Covariant Differentiation

We start with a geometric definition on S.
Definition. *Let Y be a vector field on S and $V_p \in T_pS$ a vector.*

$$\nabla_V Y := [D_V Y]^{\parallel}$$

where $D_V Y$ is the Euclidean derivative $\frac{d}{dt}Y(c(t))|_{t=0}$ for c a curve in S with $c(0) = p, \dot{c}(0) = V_p$.

So we start with $D_V Y$ not parallel to S, then we project it down onto the tangent plane of the surface at p to solve the previously mentioned problem. (Sidenote: $\nabla_V Y$ is pronounced 'nabla' or 'del V of Y.')

This definition gives us a relationship with the second fundamental form

$$D_V Y = [D_V Y]^{\perp} + [D_V Y]^{\parallel} = A(V, Y)N + \nabla_V Y$$

because by the product rule,

$$[D_V Y]^{\perp} = \langle D_V Y, N\rangle = D_V \langle Y, N\rangle - \langle Y, D_V N\rangle$$

The first part $\langle Y, N\rangle$ is zero because Y is parallel to the surface. The second part is $-\langle Y, D_V N\rangle = A(Y, V)$, from our definition of the second fundamental form.

3.1 Five Properties of the Covariant Derivative

As defined, $\nabla_V Y$ depends only on V_p and Y to first order along c. It's a very local derivative. It also satisfies the following five properties:

1. C^∞-linearity in the V-slot. $\nabla_{V_1 + fV_2} Y = \nabla_{V_!} Y + f\nabla_{V_2} Y$ where $f : S \to R$.

   This property seems trivial, but something is going on that needs some thought here. What is the directional derivative? In $\mathbb{R}^3$,

   $$D_{V_1 + fV_2}(h) = [Dh]^\top [V_1 + fV_2] = [Dh]^\top V_1 + f[Dh]^\top V_2$$

   What could go wrong? We used a curve to define a derivative. Suppose we have a curve c_1 with tangent vector V_1 and a curve c_2 with tangent vector V_2. What's the curve that generates $V_1 + fV_2$? It is not clear what this curve should be. We don't have partial derivatives to get around this on a surface. The intuition is that to first order, none of this matters. All of this can be defined in terms of a parameter domain *and* can be proved to be independent of parameterization. It ends up not being an issue.

2. $\mathbb{R}$-linearity in the Y-slot. $\nabla_V(Y_1 + aY_2) = \nabla_V Y_1 + a\nabla_V Y_2$ where $a \in \mathbb{R}$. This means differentiation is a linear operation, as usual.

3. Product rule in the Y-slot. $\nabla_V(fY) = f \cdot \nabla_V Y + (\nabla_V f) \cdot Y$ where $f : S \to R$. Again, this comes from the properties of ordinary derivatives. The derivative of the function f is the usual scalar derivative.

4. The metric compatibility property. $\nabla_V \langle Y, Z \rangle = \langle \nabla_V Y, Z \rangle + \langle Y, \nabla_V Z \rangle$. This is a generalization of something that is true in Euclidean space:

$$\begin{aligned} \langle Y_1, Y_2 \rangle &= \sum_{i=1}^{3} Y_1^i Y_2^i \\ \Rightarrow D_V \langle Y_1, Y_2 \rangle &= \frac{d}{dt} \langle Y_1(c(t)), Y_2(c(t)) \rangle \Big|_{t=0} \\ &= \frac{d}{dt} \sum_{i=1}^{3} Y_1^i Y_2^i \Big|_{t=0} \\ &= \sum_{i=1}^{3} \frac{dY_1^i}{dt}\Big|_{t=0} Y_2^i + \sum_{i=1}^{3} Y_1^i \frac{dY_2^i}{dt}\Big|_{t=0} \\ &= \langle D_V Y_1, Y_2 \rangle + \langle Y_1, D_V Y_2 \rangle \end{aligned}$$

So this property follows from the product rule (as applied when going from line 3 to 4).
This property means the covariant derivative interacts in the 'nicest possible way' with the inner product on the surface, just as the usual derivative interacts nicely with the general Euclidean inner product.

5. The 'torsion-free' property. $\nabla_{V_1} V_2 - \nabla_{V_2} V_1 = [V_1, V_2]$.

The *Lie bracket* $[V_1, V_2](f) := D_{V_1} D_{V_2}(f) - D_{V_2} D_{V_1}(f)$ is *tangent* to S if V_1, V_2 are.
'Intuition:' we know
$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$$
for all scalar functions $f \in C^2(\mathbb{R}^n)$. ('Second partial derivatives commute.') Or, we can say the operator
$$\frac{\partial}{\partial x^i}\frac{\partial}{\partial x^j} - \frac{\partial}{\partial x^j}\frac{\partial}{\partial x^i}$$
is zero. So what this property is to scalar functions, property five is to vector fields. It is basically 'taking care of the commutativity of second derivatives.'
(Note: A lecture supplement with more details on Lie differentiation and the Lie bracket has been posted on the website.)

3.2 Parameter Domain

We have given a geometric definition, so now we need to see what this looks like in a parameter domain. Let $\phi : \mathcal{U} \to S$ be a parametrization with $\phi(0) = p$. A basis of the tangent planes $T_{\phi(u)}S$ near p is given by $E_i(u) := \frac{\partial \phi}{\partial u^i}$.

Let $V_p = \sum_i a^i E_i(0)$ and $Y_{\phi(u)} := \sum_i b^i(u) E_i(u)$ (so we are expressing V_p and Y in parameter domain basis). E_1 and E_2 are just the usual basis vectors $e_1 = (1, 0)$ and $e_2 = (0, 1)$ in the parameter domain pushed forward at every point.
Use the five properties to compute $\nabla_V Y$:

$$\begin{aligned}
\nabla_V Y &= \nabla_{\sum_i a^i E_i}\left(\sum_j b^j E_j\right) \\
&= \sum_i a^i \nabla_{E_i}\left(\sum_j b^j E_j\right), \text{ by property 1} \\
&= \sum_{ij} a^i \nabla_{E_i}(b^j E_j), \text{ by property 2} \\
&= \sum_{ij} a^i \left(\nabla_{E_i}(b^j) E_j + b^j \nabla_{E_i} E_j\right), \text{ by property 3}
\end{aligned}$$

We know $\nabla_{E_i} E_j$ is expressible in the E-basis since it's tangent to S (by the definition of the covariant derivative). Expressing these terms in the E_k basis yields the 'Christoffel symbols' Γ^k_{ij}, where

$$\nabla_{E_i} E_j := \sum_k \Gamma^k_{ij} E_k$$

Substitute and switch around dummy variables to get

$$\begin{aligned}
\sum_{ij} a^i \left(\nabla_{E_i}(b^j) E_j + b^j \nabla_{E_i} E_j\right) &= \sum_{ik} a^i \nabla_{e_i}(b^k) E_k + \sum_{ij} a^i b^j \nabla_{E_i} E_j \\
&= \sum_{ik} a^i \frac{\partial b^k}{\partial u^i} E_k + \sum_{ij} a^i b^j \left(\sum_k \Gamma^k_{ij} E_k\right) \\
&= \sum_k \left(\sum_i a^i \frac{\partial b^k}{\partial u^i} + \sum_{ij} a^i b^j \Gamma^k_{ij}\right) E_k
\end{aligned}$$

Note: $\frac{\partial b^k}{\partial u^i}$ is the derivative $\nabla_{E_i} b^k(u)$ because it is the derivative of b_k in the E_i direction.

4 More Applications of ∇

4.1 Fundamental Lemma of Riemannian Geometry

The induced metric g and the Five Properties together determine a unique covariant derivative called the Levi-Civita connection.

This relationship between g and ∇ is determined by the formula

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial u^j} + \frac{\partial g_{kj}}{\partial u^i} - \frac{\partial g_{ij}}{\partial u^k}\right)$$

where $\Gamma_{ijk} := g(\nabla_{E_i} E_j, E_k)$.

Note: $\Gamma_{ij}^k = \sum_\ell g^{k\ell}\Gamma_{ij\ell}$, where $g^{k\ell}$ are the components of g^{-1}.

4.2 Geodesic Equation

Recall: so far the geodesic equation $\vec{k}_\gamma(t) \perp T_{\gamma(t)}S$ is extrinsic. But we can re-express this as a purely *intrinsic* equation:

$$\begin{aligned}
\vec{k}_\gamma(t) \perp T_{\gamma(t)}S &\Leftrightarrow \ddot{\gamma}(t) \perp T_{\gamma(t)}S \\
&\Leftrightarrow [\ddot{\gamma}(t)]^{\parallel} = 0 \\
&\Leftrightarrow [D_{\dot{\gamma}}\dot{\gamma}(t)]^{\parallel} = 0 \\
&\Leftrightarrow \nabla_{\dot{\gamma}}\dot{\gamma}(t) = 0
\end{aligned}$$

In the parameter domain, once you substitute everything in this ends up being a system of second order ODEs with coefficients determined from g:

$$\gamma \text{ is a geodesic} \quad \Leftrightarrow \quad \frac{d^2\gamma^k}{dt^2} + \sum_{ij} \frac{d\gamma^i}{dt}\frac{d\gamma^j}{dt}\Gamma_{ij}^k = 0$$

4.3 Gradient of a function

We can use directional derivatives to give a geometric definition of gradient. The gradient of f is supposed to tell what direction to walk in to get to the 'top of the mountain' as quickly as possible, i.e. it is the direction of steepest ascent. So gradient is defined geometrically as 'something whose inner product gives you the directional derivative.' In Euclidean space the gradient is the transpose of the derivative matrix $[Df_p]^\top$, so

$$\left.\frac{df(c(t))}{dt}\right|_{t=0} := \langle \nabla f(p), V\rangle$$

But in the parameter domain, the metric changes at every point, so orthogonality does not mean the same thing over the surface. So how can we make the same thing happen? We can replace inner product with g. So now, $\nabla f = g^{-1} \cdot Df$. Then to find the gradient, you need the derivative of the function f as well as the metric g.

4.4 Vector analysis operators

We can compute parameter domain formulas for all the important covariant differential operators on a surface.

For the gradient of $f : S \to \mathbb{R}$, we have a geometric definition

$$\nabla f \text{ s.t. } D_V(f) := \langle \nabla f, V \rangle$$

and a corresponding parameter domain definition

$$[\nabla f]^i := \sum_j g^{ij} \frac{\partial f}{\partial u^j}$$

For the divergence of a vector field X, we have a geometric definition

$$\nabla \cdot X := \sum_j \langle \nabla_{E_i} X, E_i \rangle$$

where E_i is an orthonormal basis, and a corresponding parameter domain definition

$$\nabla \cdot X := \sum_i \left[\frac{\partial X^i}{\partial u^i} + \sum_j \Gamma^i_{ij} X^j \right]$$

We define the Laplacian of $f : S \to \mathbb{R}$ in terms of the divergence and gradient, so

$$\Delta f := \nabla \cdot (\nabla f)$$

and in the parameter domain,

$$\Delta f := \sum_{ij} g^{ij} \left[\frac{\partial^2 f}{\partial u^i \partial u^j} + \Gamma^k_{ij} \frac{\partial f}{\partial u^k} \right]$$

Note: We have an integration by parts formula

$$\int_S f \, \nabla \cdot X dA = - \int_S \langle \nabla f, X \rangle dA + \int_{\partial S} f \langle X, \vec{n}_{\partial S} \rangle d\ell$$