Levi-Civita Connection and Covariant Derivative

Let $\mathcal{N}$ be a euclidean manifold of dimension n. We will use $\bar{x}^{\bar{\alpha}}$ ($\bar{\alpha}: \ 1, \ldots, n$) to represent cartesian coordinates on $\mathcal{N}$. Because $\mathcal{N}$ is euclidean, we can treat the entire manifold as a vector space centered on the origin of our coordinates. Let $\mathcal{M}$ be an m dimensional submanifold embedded in $\mathcal{N}$. We will use x^μ ($\mu: \ 1, \ldots, m$) to represent general coordinates on a patch of $\mathcal{M}$. Because $\mathcal{N}$ can be treated as an n dimensional vector space, we can take the unbarred coordinates as parameters to define a parameterized hypersurface in $\mathbb{R}^n$ such that the values of this function are n dimensional vectors, *i.e.* $\bar{\mathbf{x}}(\mathbf{x}): \ \mathbb{R}^m \to \mathbb{R}^n$. We can then define m vectors in $\mathbb{R}^n$:

$$\mathbf{e}_\mu \coloneqq \partial_\mu \bar{\mathbf{x}}$$

where each vector has n components:

$$e^{\bar{\alpha}}_\mu = \frac{\partial \bar{x}^{\bar{\alpha}}}{\partial x^\mu}$$

These are the coordinate basis vectors for our unbarred coordinates, and they span the tangent space at each point in the portion of $\mathcal{M}$ covered by the coordinate chart. Since

$$\frac{\partial x^\nu}{\partial \bar{x}^{\bar{\alpha}}} \frac{\partial \bar{x}^{\bar{\alpha}}}{\partial x^\mu} = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu$$

and by definition, $\langle \boldsymbol{\omega}^\nu, \mathbf{e}_\mu \rangle \coloneqq \delta^\nu_\mu$ (where $\boldsymbol{\omega}^\nu$ is the set of m basis one-forms that are dual to our basis vectors $\mathbf{e}_\mu$), we have

$$\omega^\nu_{\bar{\alpha}} \equiv \frac{\partial x^\nu}{\partial \bar{x}^{\bar{\alpha}}}$$

Ultimately, we will want to be able to do calculus on $\mathcal{M}$ without needing to reference a higher dimensional embedding space. However, I think it's conceptually easier to start with the embedding space and see if we can find a useful way to dispose of it later. Let's start with differentiating our basis vectors. Upon contemplating the vector:

$$(\partial_\mu e_\nu)^{\bar{\alpha}} = \frac{\partial^2 \bar{x}^{\bar{\alpha}}}{\partial x^\mu \partial x^\nu}$$

we see that, in general, it will not lie in $\mathcal{M}$'s tangent space. So we can define this vector as some linear combination of our m basis vectors, $\mathbf{e}_\theta$, and $n-m$ linearly independent normal vectors, $\mathbf{n}_i$, like this:

$$\partial_\mu \mathbf{e}_\nu \coloneqq \Gamma^\theta_{\mu\nu} \mathbf{e}_\theta + K^i_{\mu\nu} \mathbf{n}_i \tag{1}$$

Those Gamma components are called "Christoffel symbols". Here is the same equation but in component form:

$$(\partial_\mu e_\nu)^{\bar{\alpha}} \coloneqq \Gamma^\theta_{\mu\nu} e^{\bar{\alpha}}_\theta + K^i_{\mu\nu} n^{\bar{\alpha}}_i \tag{2}$$

(note the distinctions between n, an integer denoting $\mathcal{N}$'s dimension, $\mathbf{n}_i$, a set of $n-m$ vectors, and $n^{\bar{\alpha}}_i$, the components of said vectors). We may simply define $\mathbf{n}_i$ with the following equation:

$$\delta_{\bar{\alpha}\bar{\beta}} e^{\bar{\alpha}}_\lambda n^{\bar{\beta}}_i \coloneqq 0 \ \ \forall \ \{\lambda, \ i\}$$

Here I introduce $\delta_{\bar{\alpha}\bar{\phi}}$, which of course functions as the metric on $\mathcal{N}$ (given the cartesian coordinates). The components of the induced metric on $\mathcal{M}$ can be found using this equation:

$$g_{\mu\nu} = \frac{\partial \bar{x}^{\bar{\alpha}}}{\partial x^{\mu}} \frac{\partial \bar{x}^{\bar{\beta}}}{\partial x^{\nu}} \delta_{\bar{\alpha}\bar{\beta}} = e^{\bar{\alpha}}_{\mu} e^{\bar{\beta}}_{\nu} \delta_{\bar{\alpha}\bar{\beta}} \tag{3}$$

Though this isn't actually a tensor equation, it can easily be derived from one. Let's imagine a new set of coordinates on $\mathcal{N}$, $\bar{y}^{\alpha}$ ($\alpha :\ 1, \ldots, n$), that are identical to x^{μ} but with an additional $n - m$ independent coordinates (let's call these specific coordinates $\bar{y}^{i}$). If we replace the x's with $\bar{y}$'s in eq. 3, then clearly that *is* a tensor equation. By dropping all of the i components of *that* metric, we get the induced metric on $\mathcal{M}$, which is essentially the same thing that eq. 3 does. Similarly, the inverse of the induced metric can be given by $g^{\mu\nu} = \omega^{\mu}_{\bar{\alpha}} \omega^{\nu}_{\bar{\beta}}\ \delta^{\bar{\alpha}\bar{\beta}}$.

Let me now claim that a useful concept of taking the derivative of vectors on $\mathcal{M}$ that does not rely on reference to $\mathcal{N}$ would be to project the derivative vector into $\mathcal{M}$'s tangent space. With how we broke up the vector $\partial_{\mu}\mathbf{e}_{\nu}$ in eq. 1 into components that lie in the tangent space and components that are orthogonal to it, if we simply throw away the orthogonal components, we get the vector's projection into the tangent space. We now need a new symbol to distinguish this type of differentiation. Instead of ∂_{μ} we will use ∇_{μ}, which we'll call the "covariant derivative". So now we have $\nabla_{\mu}\mathbf{e}_{\nu} \coloneqq \Gamma^{\theta}_{\mu\nu}\mathbf{e}_{\theta}$. Then if we contract with our basis one-forms, we get $\Gamma^{\lambda}_{\mu\nu} = \langle \boldsymbol{\omega}^{\lambda}, \nabla_{\mu}\mathbf{e}_{\nu} \rangle$. As far as finding the actual values of $\Gamma^{\lambda}_{\mu\nu}$, this equation will be of no use to us because it uses the covariant derivative, and we *defined* the covariant derivative using the hitherto inscrutable Christoffel symbols. We therefore want to find some alternative equation for it. It turns out that contracting the full eq. 1 with our basis one-forms also gives a very similar equation, but brings us back to thinking in terms of our embedding space. First off, it will be necessary to show that $\langle \boldsymbol{\omega}^{\mu}, \mathbf{n}_{i} \rangle = 0$. This might be intuitive, but let's quickly prove it anyway. Given that $\omega^{\mu}_{\bar{\beta}} = g^{\mu\nu}\delta_{\bar{\alpha}\bar{\beta}} e^{\bar{\alpha}}_{\nu}$, we have $\omega^{\mu}_{\bar{\beta}} n^{\bar{\beta}}_{i} = g^{\mu\nu}\delta_{\bar{\alpha}\bar{\beta}} e^{\bar{\alpha}}_{\nu} n^{\bar{\beta}}_{i}$, and if you look at the last three terms on the RHS, you'll recall that we already defined that to equal zero, so there you have it.

If we now multiply eq. 2 by $\omega^{\lambda}_{\bar{\alpha}}$ we get $\Gamma^{\lambda}_{\mu\nu} = \omega^{\lambda}_{\bar{\alpha}} (\partial_{\mu} e_{\nu})^{\bar{\alpha}}$. So now our goal is to see if we can manipulate this equation to get rid of any terms that require knowledge of how $\mathcal{M}$ is embedded in $\mathcal{N}$. In our scenario, we do use this knowledge to calculate the induced metric on $\mathcal{M}$ (see eq. 3), but beings that live completely within the confines of $\mathcal{M}$ are also perfectly capable of figuring out what that metric is *without* knowledge of any higher dimensions. So getting an equation for the Christoffel symbols in terms of metric components sounds like a good goal. Let's start with this:

$$\begin{aligned}
\Gamma^{\lambda}_{\mu\nu} &= \omega^{\lambda}_{\bar{\alpha}} (\partial_{\mu} e_{\nu})^{\bar{\alpha}} \\
&= {}^{1}\!/_{2}\ \omega^{\lambda}_{\bar{\alpha}} (\partial_{\mu} e_{\nu})^{\bar{\alpha}} + {}^{1}\!/_{2}\ \omega^{\lambda}_{\bar{\alpha}} (\partial_{\nu} e_{\mu})^{\bar{\alpha}}
\end{aligned}$$

Notice that in the last term I switched the μ and the ν around. Recall that $\partial_{\mu}\mathbf{e}_{\nu}$ can be defined as $\frac{\partial^2}{\partial x^{\mu} \partial x^{\nu}}(\bar{\mathbf{x}})$. Since the order of differentiation is commutative, $\partial_{\mu}\mathbf{e}_{\nu}$ is symmetric on μ and ν. (Incidentally, this means that $\Gamma^{\lambda}_{\mu\nu}$ is also symmetric on μ and ν.) Moving on, here are the next few steps:

$$\begin{aligned}
\Gamma^{\lambda}_{\mu\nu} &= {}^{1}\!/_{2}\ \omega^{\lambda}_{\bar{\alpha}} \left(\delta^{\bar{\alpha}}_{\bar{\beta}} (\partial_{\mu} e_{\nu})^{\bar{\beta}} + \delta^{\bar{\alpha}}_{\bar{\gamma}} (\partial_{\nu} e_{\mu})^{\bar{\gamma}} \right) \\
&= {}^{1}\!/_{2}\ \omega^{\lambda}_{\bar{\alpha}} \left(\delta^{\bar{\gamma}}_{\bar{\phi}}\ \delta^{\bar{\alpha}\bar{\phi}} \delta_{\bar{\beta}\bar{\gamma}} (\partial_{\mu} e_{\nu})^{\bar{\beta}} + \delta^{\bar{\beta}}_{\bar{\phi}}\ \delta^{\bar{\alpha}\bar{\phi}} \delta_{\bar{\beta}\bar{\gamma}} (\partial_{\nu} e_{\mu})^{\bar{\gamma}} \right) \\
&= {}^{1}\!/_{2}\ \omega^{\lambda}_{\bar{\alpha}} \left(e^{\bar{\gamma}}_{\eta} \omega^{\eta}_{\bar{\phi}}\ \delta^{\bar{\alpha}\bar{\phi}} \delta_{\bar{\beta}\bar{\gamma}} (\partial_{\mu} e_{\nu})^{\bar{\beta}} + e^{\bar{\beta}}_{\eta} \omega^{\eta}_{\bar{\phi}}\ \delta^{\bar{\alpha}\bar{\phi}} \delta_{\bar{\beta}\bar{\gamma}} (\partial_{\nu} e_{\mu})^{\bar{\gamma}} \right) \\
&= {}^{1}\!/_{2}\ \omega^{\lambda}_{\bar{\alpha}} \omega^{\eta}_{\bar{\phi}}\ \delta^{\bar{\alpha}\bar{\phi}} \delta_{\bar{\beta}\bar{\gamma}} \left((\partial_{\mu} e_{\nu})^{\bar{\beta}} e^{\bar{\gamma}}_{\eta} + (\partial_{\nu} e_{\mu})^{\bar{\gamma}} e^{\bar{\beta}}_{\eta} \right)
\end{aligned}$$

Already, we can see the inverse metric out front, so we're moving in the right direction. Onto the next step:

$$\Gamma^{\lambda}_{\mu\nu} = {}^{1}\!/_{2}\ g^{\lambda\eta} \delta_{\bar{\beta}\bar{\gamma}} \left((\partial_{\mu} e_{\nu})^{\bar{\beta}} e^{\bar{\gamma}}_{\eta} + (\partial_{\nu} e_{\mu})^{\bar{\gamma}} e^{\bar{\beta}}_{\eta} + (\partial_{\mu} e_{\eta})^{\bar{\gamma}} e^{\bar{\beta}}_{\nu} - (\partial_{\eta} e_{\mu})^{\bar{\beta}} e^{\bar{\gamma}}_{\nu} + (\partial_{\nu} e_{\eta})^{\bar{\beta}} e^{\bar{\gamma}}_{\mu} - (\partial_{\eta} e_{\nu})^{\bar{\gamma}} e^{\bar{\beta}}_{\mu} \right)$$

Here we are adding zero inside the parentheses by adding two terms and then subtracting them. You'll notice that between the added and subtracted terms I've switched not only the subscripts around (which is fine due to the symmetry noted earlier), but also the gamma and beta indices. I can do this because of the symmetry of the $\delta_{\bar{\beta}\bar{\gamma}}$ out front. You'll see where I'm going with this once we note that $\partial_\mu(e_\nu^{\bar{\beta}} e_\eta^{\bar{\gamma}} \delta_{\bar{\beta}\bar{\gamma}}) = \delta_{\bar{\beta}\bar{\gamma}} \left((\partial_\mu e_\nu)^{\bar{\beta}} e_\eta^{\bar{\gamma}} + (\partial_\mu e_\eta)^{\bar{\gamma}} e_\nu^{\bar{\beta}} \right)$. This means we can simplify to

$$\Gamma^\lambda_{\mu\nu} = {}^1\!/\!{}_2 \; g^{\lambda\eta} \left(\partial_\mu(e_\nu^{\bar{\beta}} e_\eta^{\bar{\gamma}} \delta_{\bar{\beta}\bar{\gamma}}) + \partial_\nu(e_\eta^{\bar{\beta}} e_\mu^{\bar{\gamma}} \delta_{\bar{\beta}\bar{\gamma}}) - \partial_\eta(e_\mu^{\bar{\beta}} e_\nu^{\bar{\gamma}} \delta_{\bar{\beta}\bar{\gamma}}) \right)$$

And finally

$$\Gamma^\lambda_{\mu\nu} = {}^1\!/\!{}_2 \; g^{\lambda\eta} \left(\partial_\mu g_{\nu\eta} + \partial_\nu g_{\eta\mu} - \partial_\eta g_{\mu\nu} \right)$$

Great! We've successfully derived an alternate equation for the Christoffel symbols that still only uses structure intrinsic to $\mathcal{M}$. Let's now look at how we can use the Christoffel symbols.

Let $\mathbf{v}(\mathbf{x}) : \mathbb{R}^m \to \mathbb{R}^m$ be a vector field on $\mathcal{M}$ such that $\mathbf{v}(\mathbf{x}) = v^\mu(\mathbf{x}) \cdot \mathbf{e}_\mu$. We would like to differentiate this vector field. If the covariant derivative we defined earlier is any kind of self-respecting derivative, it will obey the Leibniz rule. This means that we ought to have $\nabla_\mu \mathbf{v} = (\nabla_\mu v^\lambda)\mathbf{e}_\lambda + v^\nu \cdot \nabla_\mu \mathbf{e}_\nu$. Right off the bat, we can change the nabla in the first term on the RHS to a regular old partial derivative. This is because taking the derivative of a scalar on $\mathcal{M}$ in some direction in $\mathcal{M}$ will obviously give an object that already lies in $\mathcal{M}$'s tangent space (technically its cotangent space), so there is no projecting that needs to be done. So what we really have is:

$$\nabla_\mu \mathbf{v} = (\partial_\mu v^\lambda)\mathbf{e}_\lambda + v^\nu \cdot \nabla_\mu \mathbf{e}_\nu \tag{4}$$

Let's now take a step back into $\mathcal{N}$ and evaluate the normal derivative in the higher dimensional embedding space. We have

$$\partial_\mu \mathbf{v} = (\partial_\mu v^\lambda)\mathbf{e}_\lambda + v^\nu \cdot \partial_\mu \mathbf{e}_\nu \tag{5}$$

It's that second term that will give a vector that lies outside of $\mathcal{M}$'s tangent space that will need to be projected into it. In eq. 4, the second term takes the projections of the derivative of the basis vectors and then scales them by components of $\mathbf{v}$. If we want to get the covariant derivative of $\mathbf{v}$ from eq. 5, we would first scale the derivative of the basis vectors by components of $\mathbf{v}$, *then* project into the tangent space. Since the operations of scaling and projecting commute, we see that eq. 4 is a valid way to express the covariant derivative of an arbitrary vector $\mathbf{v}$, and therefore the covariant derivative does indeed obey the Leibniz rule.

Moving on, since we defined $\nabla_\mu \mathbf{e}_\nu$ to be $\Gamma^\lambda_{\mu\nu}\mathbf{e}_\lambda$, eq. 4 can now be written as $\nabla_\mu \mathbf{v} = (\partial_\mu v^\lambda + \Gamma^\lambda_{\mu\nu} v^\nu)\mathbf{e}_\lambda$, then in component form:

$$\nabla_\mu v^\lambda = \partial_\mu v^\lambda + \Gamma^\lambda_{\mu\nu} v^\nu$$

Let's see if we can now figure out how to take the covariant derivative of a one-form. We know that $\langle \boldsymbol{\omega}^\lambda, \mathbf{e}_\nu \rangle$ equals either one or zero depending on whether or not $\lambda = \nu$, and therefore taking any kind of derivative of this expression, be it covariant or otherwise, would give zero, *i.e.* $\nabla_\mu \langle \boldsymbol{\omega}^\lambda, \mathbf{e}_\nu \rangle = 0$. We know the covariant derivative obeys the Leibniz rule, and since contraction is just the sum of products of pairs of scalars, the Leibniz rule must extend to contraction. So we have $\nabla_\mu \langle \boldsymbol{\omega}^\lambda, \mathbf{e}_\nu \rangle = \langle \nabla_\mu \boldsymbol{\omega}^\lambda, \mathbf{e}_\nu \rangle + \langle \boldsymbol{\omega}^\lambda, \nabla_\mu \mathbf{e}_\nu \rangle$ and therefore $\langle \nabla_\mu \boldsymbol{\omega}^\lambda, \mathbf{e}_\nu \rangle = -\Gamma^\lambda_{\mu\nu}$. Working in the reverse direction we had worked earlier, we can see that $\nabla_\mu \boldsymbol{\omega}^\lambda = -\Gamma^\lambda_{\mu\nu} \boldsymbol{\omega}^\nu$, and if we define a general one-form $\boldsymbol{\sigma} \coloneqq \sigma_\nu \; \boldsymbol{\omega}^\nu$, we have $\nabla_\mu \boldsymbol{\sigma} = (\partial_\mu \sigma_\nu)\boldsymbol{\omega}^\nu + (\nabla_\mu \boldsymbol{\omega}^\lambda)\sigma_\lambda = (\partial_\mu \sigma_\nu - \Gamma^\lambda_{\mu\nu}\sigma_\lambda)\boldsymbol{\omega}^\nu$, and again in component form:

$$\nabla_\mu \sigma_\nu = \partial_\mu \sigma_\nu - \Gamma^\lambda_{\mu\nu}\sigma_\lambda$$

From here it's easy to derive the equation for the covariant derivative of any arbitrary tensor:

$$\nabla_\mu T^{\lambda_1 \ldots \lambda_p}_{\nu_1 \ldots \nu_q} = \partial_\mu T^{\lambda_1 \ldots \lambda_p}_{\nu_1 \ldots \nu_q} + \Gamma^{\lambda_1}_{\mu\alpha} T^{\alpha\lambda_2 \ldots \lambda_p}_{\nu_1 \ldots \nu_q} + \cdots + \Gamma^{\lambda_p}_{\mu\alpha} T^{\lambda_1 \ldots \alpha}_{\nu_1 \ldots \nu_q} - \Gamma^{\alpha}_{\mu\nu_1} T^{\lambda_1 \ldots \lambda_p}_{\alpha\nu_2 \ldots \nu_q} - \cdots - \Gamma^{\alpha}_{\mu\nu_q} T^{\lambda_1 \ldots \lambda_p}_{\nu_1 \ldots \alpha}$$