We've already seen that:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \quad \text{eq (2)}$$

We postulate that we can come up with a similar equation for covectors:

$$\nabla_\mu V_\mu = \partial_\mu V_\nu + \gamma^\lambda_{\mu\nu} V_\lambda \quad \text{eq (2)}$$

Although we won't prove it here, just from the structure of these equations, one can imagine that, like eq (1), eq (2) transforms like a tensor.

To relate the two equations, we make two assumptions:

$\nabla_\mu \delta^\nu_\lambda = 0$ (which is reasonable because the delta function contains 1's and 0's and the derivative of a constant is 0.)

For a scalar $\phi, \nabla_\mu \phi = \partial_\mu \phi$ (which makes sense since a scalar is the same in all coordinate systems. The change in basis vectors with change in coordinate system, therefore, would not be expected to have any effect on a scalar. Thus, no correction to the normal partial derivative should be needed.)

Given this, let's calculate the derivative of a scalar, $\nabla_\mu (V_\sigma W^\sigma)$. ($V_\sigma W^\sigma$ is a scalar because it's the inner product i.e., dot product, of two vectors). We have:

$$\begin{aligned}
\nabla_\mu (V_\sigma W^\sigma) &= \nabla_\mu (\delta^\sigma_\lambda V_\sigma W^\lambda) \\
&= (V_\sigma W^\lambda) \nabla_\mu \delta^\sigma_\lambda + \nabla_\mu (V_\sigma W^\sigma) \\
&= 0 + V_\sigma (\nabla_\mu W^\sigma) + W^\sigma (\nabla_\mu V_\sigma) \\
&= V_\sigma (\partial W^\sigma + \Gamma^\sigma_{\mu\lambda} W^\lambda) + W^\sigma (\partial V_\sigma + \gamma^\lambda_{\mu\sigma} V_\lambda) \\
&= V_\sigma \partial W^\sigma + V_\sigma \Gamma^\sigma_{\mu\lambda} W^\lambda + W^\sigma \partial V_\sigma + W^\sigma \gamma^\lambda_{\mu\sigma} V_\lambda \\
&= V_\sigma \partial W^\sigma + W^\sigma \partial V_\sigma + V_\sigma \Gamma^\sigma_{\mu\lambda} W^\lambda + W^\sigma \gamma^\lambda_{\mu\sigma} V_\lambda \quad \text{eq (3)}
\end{aligned}$$

We use our second assumption to say:

$$\nabla_\mu(V_\sigma W^\sigma) = \partial_\mu(V_\sigma W^\sigma)$$
$$= V_\sigma(\partial_\mu W^\sigma) + W^\sigma(\partial_\mu V_\sigma) \quad \text{eq (4)}$$

Substituting, eq (4) into eq (3), we get:

$$\nabla_\mu(V_\sigma W^\sigma) = \underbrace{V_\sigma \partial W^\sigma + W^\sigma \partial V_\sigma}_{\nabla_\mu(V_\sigma W^\sigma)} + V_\sigma \Gamma^\sigma_{\mu\lambda} W^\lambda + W^\sigma \gamma^\lambda_{\mu\sigma} V_\lambda \quad \text{eq (5)}$$

Therefore:

$$0 = \Gamma^\sigma_{\mu\lambda} V_\sigma W^\lambda + \gamma^\lambda_{\mu\sigma} W^\sigma V_\lambda \quad \text{eq (6)}$$

Relabeling dummy indices and rearranging terms gives us:

$$\gamma^\lambda_{\mu\sigma} W^\sigma V_\lambda = -\Gamma^\lambda_{\mu\sigma} W^\sigma V_\lambda \quad \text{eq (7)}$$

So, in general:

$$\gamma^\lambda_{\mu\sigma} = -\Gamma^\lambda_{\mu\sigma} \quad \text{eq (8)}$$

Substituting eq (8) into eq (2), we are left with:

$$\nabla_\mu V_\mu = \partial_\mu V_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda \quad \text{eq (9)}$$

Which is what we sought to prove.