

# Notes on finding the lowest eigenvector

Jacob Stevenson

September 30, 2013

This describes a method for computing the smallest eigenvector of a Hessian matrix  $H_{ij} = \partial^2 E(\vec{x}) / (\partial x_i \partial x_j)$  of an energy function  $E(\vec{x})$ . In this method the gradients  $\vec{g} = \partial E(\vec{x}) / \partial \vec{x}$  of the energy function are used, but the exact Hessian is not. The lowest eigenvector is found by estimating the curvature at a give point  $\vec{x}$  along a direction  $\vec{v}$  using finite differences and then using an optimization algorithm to minimizing the curvature with respect to  $\vec{v}$ .

## 1 Central differences

The curvature is estimated by the the central differences formula, which is accurate up to order  $\delta^2$  where  $\delta$  is a small parameter.

$$\mu(\vec{x}, \vec{v}) = \frac{1}{\delta} \left[ \vec{g}(\vec{x} + \frac{\delta}{2} \vec{v}) - \vec{g}(\vec{x} - \frac{\delta}{2} \vec{v}) \right] \cdot \vec{v} \quad (1)$$

If N is the dimension of the original space then this is an optimization problem in N-1 dimensions because of the constraint that  $|\vec{v}| = 1$ .

The gradients of the above function can be computed analytically to aid in the optimization.

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} \sum_i \frac{\partial v_i}{\partial v_k} \left[ g_i(\vec{x} + \frac{\delta}{2} \vec{v}) - g_i(\vec{x} - \frac{\delta}{2} \vec{v}) \right] + \frac{1}{\delta} \sum_i v_i \frac{\partial}{\partial v_k} \left[ g_i(\vec{x} + \frac{\delta}{2} \vec{v}) - g_i(\vec{x} - \frac{\delta}{2} \vec{v}) \right] \quad (2)$$

In the next we use

$$\frac{\partial g_i(\vec{x} + \frac{\delta}{2} \vec{v})}{\partial v_k} = \frac{\delta}{2} \sum_j \frac{\partial g_i(\vec{x} + \frac{\delta}{2} \vec{v})}{\partial x_j} \frac{\partial v_j}{\partial v_k} = \frac{\delta}{2} \frac{\partial g_i(\vec{x} + \frac{\delta}{2} \vec{v})}{\partial x_k} \quad (3)$$

where  $\partial v_i / \partial v_k$  is the kroniker delta  $\delta_{ik}$ .

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} \left[ g_k(\vec{x} + \frac{\delta}{2} \vec{v}) - g_k(\vec{x} - \frac{\delta}{2} \vec{v}) \right] + \frac{1}{2} \sum_i v_i \frac{\partial}{\partial x_k} \left[ g_i(\vec{x} + \frac{\delta}{2} \vec{v}) + g_i(\vec{x} - \frac{\delta}{2} \vec{v}) \right] \quad (4)$$

To second order in  $\delta$  we can replace the second term with

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} \left[ g_k(\vec{x} + \frac{\delta}{2} \vec{v}) - g_k(\vec{x} - \frac{\delta}{2} \vec{v}) \right] + \sum_i v_i \frac{\partial g_i(\vec{x})}{\partial x_k} \quad (5)$$

~~We can now use the eigenvalue equation  $\vec{v} \cdot H = \mu \vec{v}$  to further simplify the second term.~~ We use the fact that the Hessian is symmetric to slightly rewrite the second term

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} \left[ g_k(\vec{x} + \frac{\delta}{2} \vec{v}) - g_k(\vec{x} - \frac{\delta}{2} \vec{v}) \right] + \sum_i v_i \frac{\partial g_k(\vec{x})}{\partial x_i} \quad (6)$$

The second term is now the change in the derivative along the direction  $\vec{v}$ . We know how to write this using the central difference method, which gives exactly the first term and we end up with.

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{2}{\delta} \left[ g_k(\vec{x} + \frac{\delta}{2} \vec{v}) - g_k(\vec{x} - \frac{\delta}{2} \vec{v}) \right] \quad (7)$$

We want to maintain the constraint  $|\vec{v}| = 1$ , so we subtract out the parallel component of the derivative

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial \vec{v}} - \left( \frac{\partial \mu(\vec{x}, \vec{v})}{\partial \vec{v}} \cdot \vec{v} \right) \vec{v} = \frac{2}{\delta} \left[ \vec{g}(\vec{x} + \frac{\delta}{2} \vec{v}) - \vec{g}(\vec{x} - \frac{\delta}{2} \vec{v}) \right] - \left( \frac{2}{\delta} \left[ \vec{g}(\vec{x} + \frac{\delta}{2} \vec{v}) - \vec{g}(\vec{x} - \frac{\delta}{2} \vec{v}) \right] \cdot \vec{v} \right) \vec{v} \quad (8)$$

Which simply reduces to

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial \vec{v}} - \left( \frac{\partial \mu(\vec{x}, \vec{v})}{\partial \vec{v}} \cdot \vec{v} \right) \vec{v} = \frac{2}{\delta} \left[ \vec{g}(\vec{x} + \frac{\delta}{2} \vec{v}) - \vec{g}(\vec{x} - \frac{\delta}{2} \vec{v}) \right] - 2\mu \vec{v} \quad (9)$$

## 2 Forward finite differences

Here the curvature is estimated by the forward finite differences method.

$$\mu(\vec{x}, \vec{v}) = \frac{1}{\delta} [\vec{g}(\vec{x} + \delta \vec{v}) - \vec{g}(\vec{x})] \cdot \vec{v} \quad (10)$$

This is accurate only up to which is accurate up to order  $\delta$ , however each calculation of  $\mu$  for a different  $\vec{v}$  requires only one additional gradient evaluation.

The gradients of the above function can be computed analytically to aid in the optimization.

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} \sum_i \frac{\partial v_i}{\partial v_k} [g_i(\vec{x} + \delta \vec{v}) - g_i(\vec{x})] + \frac{1}{\delta} \sum_i v_i \frac{\partial g_i(\vec{x} + \delta \vec{v})}{\partial v_k} \quad (11)$$

Following the arguments from the central differences method this can be reduced to

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} [g_k(\vec{x} + \delta \vec{v}) - g_k(\vec{x})] + \sum_i v_i \frac{\partial g_i(\vec{x} + \delta \vec{v})}{\partial x_k} \quad (12)$$

To first order in  $\delta$  the second term simplifies to

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{1}{\delta} [g_k(\vec{x} + \delta \vec{v}) - g_k(\vec{x})] + \sum_i v_i \frac{\partial g_k(\vec{x})}{\partial x_i} \quad (13)$$

In the above we also used the fact that the Hessian is symmetric. Expanding the derivative in the second term using forward finite differences gives the same equation as the first term and we're left with

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial v_k} = \frac{2}{\delta} [g_k(\vec{x} + \delta \vec{v}) - g_k(\vec{x})] \quad (14)$$

Subtracting out the component parallel to  $\vec{v}$  gives finally

$$\frac{\partial \mu(\vec{x}, \vec{v})}{\partial \vec{v}} - \left( \frac{\partial \mu(\vec{x}, \vec{v})}{\partial \vec{v}} \cdot \vec{v} \right) \vec{v} = \frac{2}{\delta} [\vec{g}(\vec{x} + \delta \vec{v}) - \vec{g}(\vec{x})] - 2\mu \vec{v} \quad (15)$$