

Notes on Parameter determination

Jeremiah MF Kelly

10/11/2020

Dark adaptation models

There are at least two models used in the literature. The first is a twin exponential decay

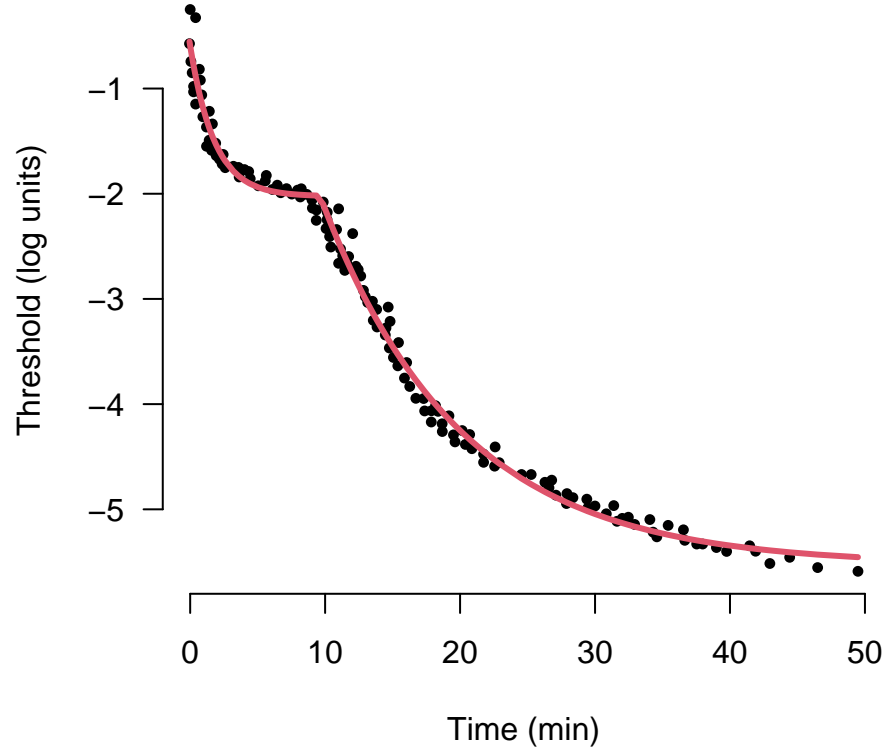


Figure 0.1: This is my plot

the second, an exponential, bilinear model, is more complex and was first described by Mahroo and Lamb (ref)

Put stuff about each model here, in particular that the MLP model has two distinct rod phases.

We are concerned with a reduced form of the Mahroo and Lamb Model with only five parameters. The threshold Thr ($\log(cd.m^{-2})$) with respect to *time* (*minutes*) of adaptation is defined as the sum of the cone response;

$$Thrs_{cone} = CT + CC * \exp\left(\frac{-time}{\tau}\right),$$

where CT is the absolute cone threshold ($\log(cd.m^{-2})$), CC ($\log(cd.m^{-2})$) the cone offset at time zero, and

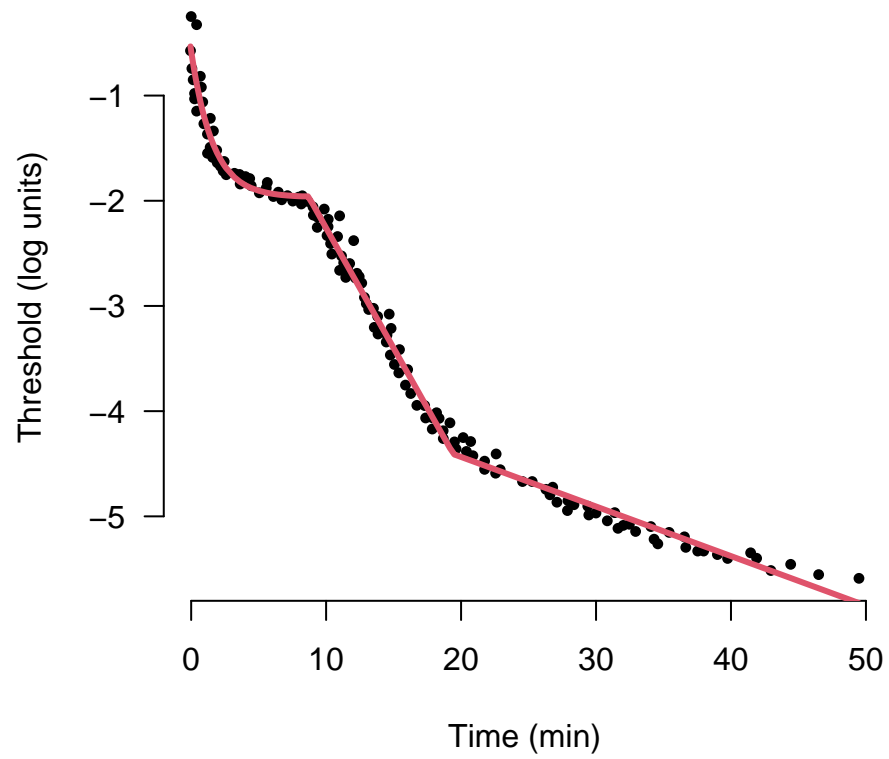


Figure 0.2: The exponential, bilinear model of dark adaptation. This model has seven parameters, cone threshold, cone coefficient, cone time constant, rod cone break time, the slope of the S2 phase, rod-rod break time, slope of the S3 phase

the time constant of recovery τ (*minutes*), and the rod threshold;

$$Thr_{s_{rod}} = \frac{S2 * (time - \alpha)}{1 + \exp(-k * (time - \alpha))},$$

where $S2$ ($\log(cd.m^{-2}).min^{-1}$) is the rate limited rod recovery rate, α (*min*) is the time when rod sensitivity is greater than cone sensitivity, and k is a dimensionless value. The denominator ensures that rod contribution to sensitivity is nil before α and additive after, the value of k affects the rate of switch transition and is discussed in the appendix.

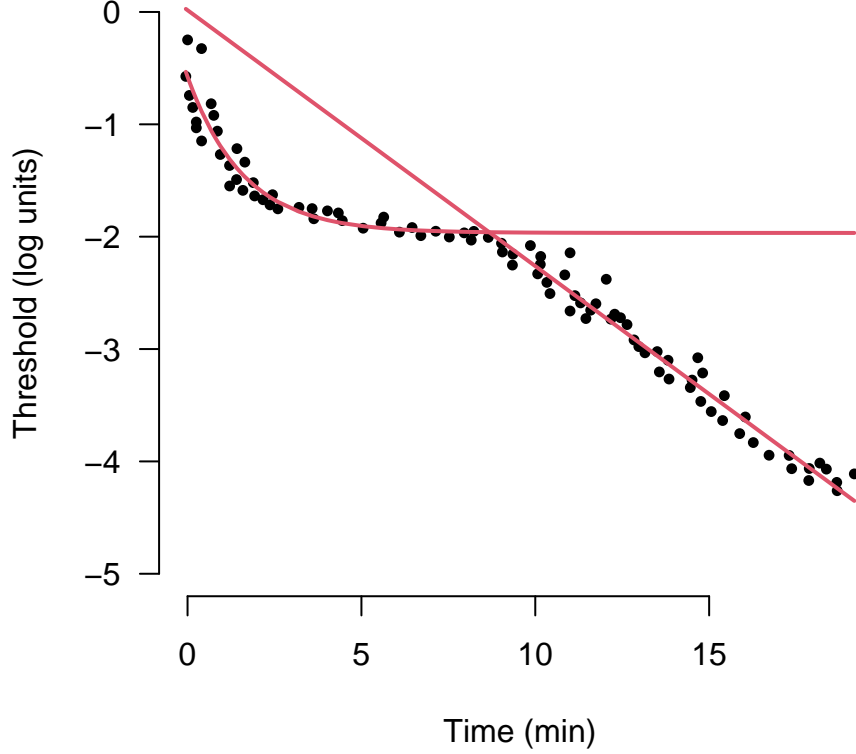


Figure 0.3: This is my plot

Hence the estimated threshold for a time t and model parameter estimate $\hat{\theta}$ is;

$$Thr_s(\hat{\theta}, t) = \theta_1 + \theta_2 \exp\left(\frac{-t}{\theta_3}\right) + \theta_4 \frac{(t - \theta_5)}{1 + \exp(-k * (t - \theta_5))} \quad (1)$$

and it follows that the residuals can be written as a vector \mathbf{R} where

$$\mathbf{R}(\hat{\theta}) = \begin{bmatrix} thr_{s1} - Thr_s(\hat{\theta}, time_1) \\ thr_{s2} - Thr_s(\hat{\theta}, time_2) \\ thr_{s3} - Thr_s(\hat{\theta}, time_3) \\ \vdots \\ thr_{sn} - Thr_s(\hat{\theta}, time_n) \end{bmatrix} \quad (1)$$

Finding the parameters

We are interested in the parameters (θ) of the model described above, a common technique takes a number (n) of threshold measurements over time and calculates the difference between the actual measurements

and estimated values for a given set of parameters. These differences are squared and summed, then the parameters of the model are modified until the sum of squares is at a minimum, the final parameter values are then considered optimal. The sum of the squared differences of the measured threshold and the estimated threshold is minimised, such that when $L(\theta)$ is at a minimum, the values of θ are the best estimate. This objective function can be written using \odot to express element wise multiplication

$$L(\hat{\theta}) = \sum_{i=1}^n \mathbf{R}(\hat{\theta}) \odot \mathbf{R}(\hat{\theta}).$$

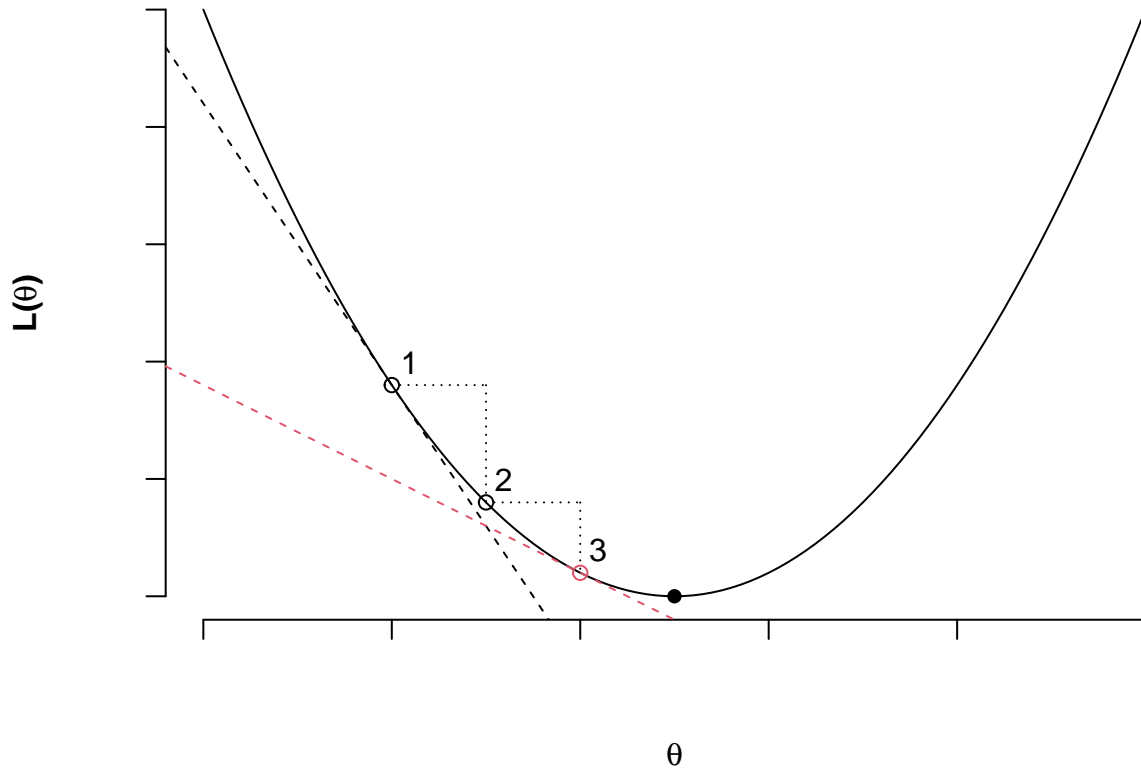
We use element wise multiplication notation to retain the \sum notation.

When $L(\hat{\theta})$ is at a minimum, $\frac{dL}{d\theta} = 0$. We can use the gradient of the slope of the objective function to locate the minimum of the function using a gradient descent method.

Gradient decent

We assume the objective function is convex over the parameter space, so that we can use the partial derivatives of the objective function to estimate the values of θ . A function composed of non decreasing convex functions will be convex (see <https://math.stackexchange.com/questions/108393/is-the-composition-of-n-convex-functions-itself-a-convex-function>, 11/11/2020). (Possible weakness here)

The method is illustrated in the figure below, and is briefly as follows. The gradient of the objective function for a set of parameters values (location 1), summarised as $\hat{\theta}_1$ is found, these values are used to make a further estimate $\hat{\theta}_2$ and a new gradient is found, which in turn is used to estimate new parameter values. The process is repeated until further changes yield little or no improvement in the value of $L(\hat{\theta}_n)$.



Finding the gradient vector

This all builds into a gradient vector that is used to move from the initial estimate to the next.

then the differentials can be expressed as;

$$\frac{\partial L}{\partial \theta_1} = -2 \sum_{i=1}^n \left(\mathbf{R}(\hat{\theta}) \odot \mathbf{1} \right), \quad (2)$$

$$\text{[cone threshold]} \quad (3)$$

$$\frac{\partial L}{\partial \theta_2} = -2 \sum_{i=1}^n n \left(\mathbf{R}(\hat{\theta}) \odot \exp \left(- \frac{time_i}{\theta_3} \right) \right), \quad (4)$$

$$\text{[cone offset]} \quad (5)$$

$$\frac{\partial L}{\partial \theta_3} = -2 \sum_{i=1}^n n \left(\mathbf{R}(\hat{\theta}) \odot \frac{\exp \left(- \theta_2 \cdot \frac{time_i}{\theta_3} \cdot time_i \right)}{\theta_3^2} \right), \quad (6)$$

$$\text{[cone time constant]} \quad (7)$$

$$\frac{\partial L}{\partial \theta_4} = -2 \sum_{i=1}^n \left(\mathbf{R}(\hat{\theta}) \odot \frac{(time_i - \theta_5)}{(1 + \exp(time_i - \theta_5))} \right), \quad (8)$$

$$\text{[S2, rod rate of recovery]} \quad (9)$$

$$\frac{\partial L}{\partial \theta_5} = -2 \sum_{i=1}^n \left(\mathbf{R}(\hat{\theta}) \odot \left(\frac{\theta_4}{1 + \exp(time_i - \theta_5)} - \theta_4 \cdot \exp(time_i - \theta_5) \frac{(time_i - \theta_5)}{(1 + \exp(time_i - \theta_5))^2} \right) \right). \quad (10)$$

$$\text{[cone-rod break point]} \quad (11)$$

$$(12)$$

hence

$$\nabla_{\theta} \mathbf{L}(\theta) = \begin{bmatrix} \frac{\partial L}{\partial \theta_1} \\ \frac{\partial L}{\partial \theta_2} \\ \frac{\partial L}{\partial \theta_3} \\ \frac{\partial L}{\partial \theta_4} \\ \frac{\partial L}{\partial \theta_5} \end{bmatrix} \quad (13)$$

Having found the gradient vector then we do the descent thing

Plain gradient descent

$$\hat{\theta}_{n+1} = \hat{\theta}_n - \eta \nabla_{\theta} L(\hat{\theta}_n) \quad (14)$$

Momentum

$$\begin{aligned} v_0 &= \mathbf{0} \\ v_n &= \gamma v_{n-1} + \eta \nabla_{\theta} L(\hat{\theta}_n) \\ \hat{\theta}_{n+1} &= \hat{\theta}_n - v_n \end{aligned} \quad (15)$$

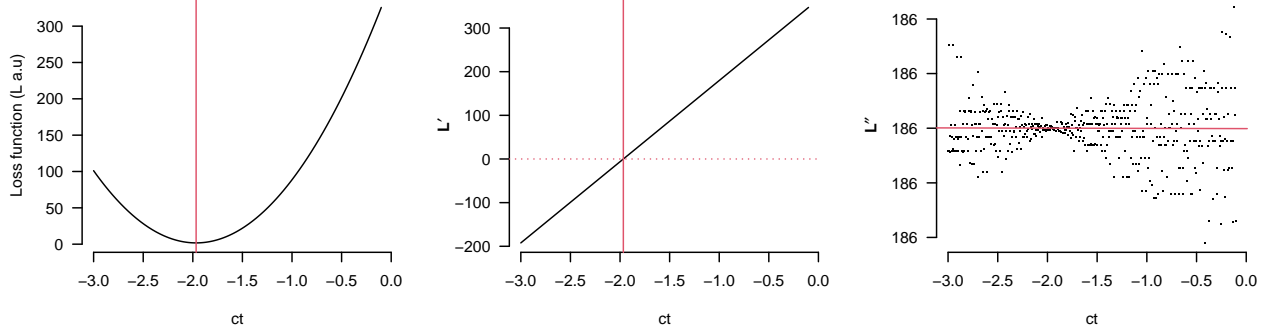
Nesterov

$$\begin{aligned} v_n &= \gamma v_{n-1} + \eta \nabla_{\theta} L(\hat{\theta}_n - \gamma v_{n-1}) \\ \hat{\theta}_{n+1} &= \hat{\theta}_n - v_n \end{aligned} \quad (16)$$

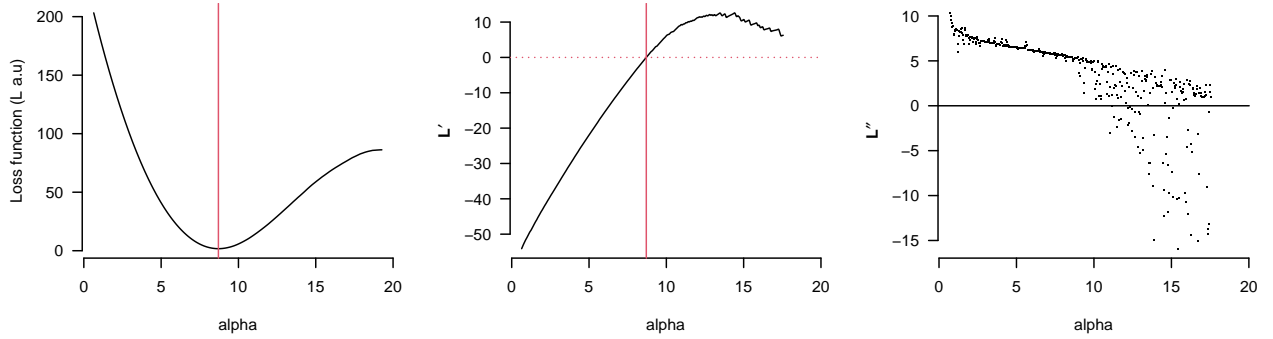
How do the gradient values vary as the estimate changes?

The first order partial differential equations are plotted below.

The parameters ct , cc and $S2$ are linear in the model



However the cone rod break time (α) behaves quite differently



Appendix

$$\nabla(\theta) = \quad (17)$$

$$\frac{\partial L}{\partial \theta_1} = \frac{\partial L}{\partial CT} = -2 \sum_{i=1}^n \left(thrs_i - \left(\theta_1 + \theta_2 \exp \left(-\frac{time_i}{\theta_3} \right) + \theta_4 \frac{(time_i - \theta_5)}{1 + \exp(time_i - \theta_5)} \right) \right) \quad (18)$$

$$\frac{\partial L}{\partial \theta_2} = \frac{\partial L}{\partial CC} = -2 \sum_{i=1}^n \exp \left(-\frac{time_i}{\theta_3} \right) \left(thrs_i - \left(\theta_1 + \theta_2 \exp \left(-\frac{time_i}{\theta_3} \right) + \theta_4 \frac{(time_i - \theta_5)}{1 + \exp(time_i - \theta_5)} \right) \right) \quad (19)$$

$$\frac{\partial L}{\partial \theta_3} = \frac{\partial L}{\partial \tau} = -2 \sum_{i=1}^n \frac{\exp \left(-\frac{time_i}{\theta_3} \right) \theta_2 \cdot time_i \left(thrs_i - \left(\theta_1 + \theta_2 \exp \left(-\frac{time_i}{\theta_3} \right) + \theta_4 \frac{(time_i - \theta_5)}{1 + \exp(time_i - \theta_5)} \right) \right)}{\theta_3^2} \quad (20)$$

$$\frac{\partial L}{\partial \theta_4} = \frac{\partial L}{\partial S2} = -2 \sum_{i=1}^n \frac{(time_i - \theta_5) \left(thrs_i - \left(\theta_1 + \theta_2 \exp \left(-\frac{time_i}{\theta_3} \right) + \theta_4 \frac{(time_i - \theta_5)}{1 + \exp(time_i - \theta_5)} \right) \right)}{(1 + \exp(time_i - \theta_5))} \quad (21)$$

$$\frac{\partial L}{\partial \theta_5} = \frac{\partial L}{\partial \alpha} = 2 \sum_{i=1}^n \left(\frac{\theta_4}{1 + \exp(\text{time}_i - \theta_5)} - \theta_4 \cdot \exp(\text{time}_i - \theta_5) \frac{(\text{time}_i - \theta_5)}{(1 + \exp(\text{time}_i - \theta_5))^2} \right) \cdot$$

$$\left(\text{thrs}_i - \left(\theta_1 + \theta_2 \exp \left(- \frac{\text{time}_i}{\theta_3} \right) + \theta_4 \frac{(\text{time}_i - \theta_5)}{1 + \exp(\text{time}_i - \theta_5)} \right) \right)$$