

5.2 Kleiber's law and linear regression

After collecting and plotting a considerable amount of data comparing the body mass versus metabolic rate (a measure of at rest energy expenditure) of a variety of animals, early twentieth-century biologist Max Kleiber noted an interesting relationship between the two values. Denoting by x_p and y_p the body mass (in kg) and metabolic rate (in kJ/day) of a given animal respectively, treating the body mass as the input feature Kleiber noted (by visual inspection) that the natural log of these two values were linearly related. That is

$$w_0 + \log(x_p)w_1 \approx \log(y_p). \quad (5.41)$$

In [Figure 1.9](#) we show a large collection of transformed data points

$$\{(\log(x_p), \log(y_p))\}_{p=1}^P \quad (5.42)$$

each representing an animal ranging from a small black-chinned hummingbird in the bottom-left corner to a large walrus in the top-right corner.

- Fit a linear model to the data shown in [Figure 1.9](#).
- Use the optimal parameters you found in part (a) along with the properties of the log function to write the nonlinear relationship between the body mass x and the metabolic rate y .
- Use your fitted line to determine how many calories an animal weighing 10 kg requires (note each calorie is equivalent to 4.18 J).

5.6 Compare the Least Squares and Least Absolute Deviation costs

Repeat the experiment outlined in [Example 5.4](#). You will need to implement the Least Absolute Deviations cost, which can be done similarly to the Least Squares implementation in [Section 5.2.4](#).

5.8 The Least Absolute Deviations cost is convex

Prove that the Least Absolute Deviations cost is convex using the zero-order definition of convexity given below.

An unconstrained function g is convex if and only if any line segment connecting two points on the graph of g lies *above* its graph. Figure 5.13 illustrates this definition of a convex function.

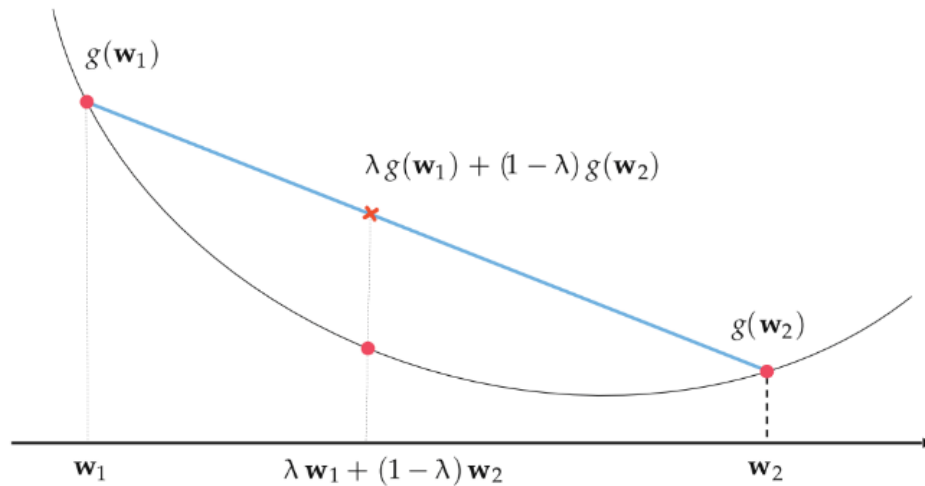


Figure 5.13 Figure associated with Exercise 5.8.

Stating this geometric fact algebraically, g is convex if and only if for all \mathbf{w}_1 and \mathbf{w}_2 in the domain of g and all $\lambda \in [0, 1]$, we have

$$g(\lambda \mathbf{w}_1 + (1 - \lambda) \mathbf{w}_2) \leq \lambda g(\mathbf{w}_1) + (1 - \lambda) g(\mathbf{w}_2). \quad (5.43)$$

5.11 Multi-output regression

Repeat the experiment outlined in Example 5.9. You can use the implementation described in Section 5.6.3 as a base for your implementation.