Simple parameter estimation in the context of a Michaelis-Menten kinetic rate law

John Mason

October 2018

1 Background

This is a highly reduced basis to serve as an introduction to my broader parameter estimation work (forthcoming publication).

2 Model

For this demonstration, I will use a basic Michaelis-Menten kinetic rate law:

$$v = kE \frac{C}{C + K} \tag{1}$$

Note the simplified notation – for this level of complexity, I want to avoid superscripts and subscripts as much as possible. Here k is the catalytic rate constant, E is the enzyme concentration, C is the metabolite concentration, and K is the so-called Michaelis-Menten constant (an effective saturation constant).

3 Objective

Our goal is to match some target rate of reaction, $v^\star$. We are not considering any prior information about parameter values, so our system is not necessarily an optimization problem – we simply wish to solve $v = v^\star$. For the sake of generality, however, I will express this as an optimization problem.

In this context we want the residual $v - v^\star$ to be as small (in the sense of closeness to zero) as possible. Thus we need to apply some metric to this residual. For purposes that I won't expand here, minimizing the square of this residual is a practical choice:

$$\min_{k,E,C,K} (v - v^\star)^2 \tag{2}$$

For expressive convenience I will define $f = (v - v^\star)^2$.

4 The objective value space

As previously mentioned, this problem can be solved directly via analytic approaches. For demonstrative purposes I will instead be using graphical approaches.

Graphical approaches only work in low dimensionality space. Unfortunately we have a function of four variables that we wish to illustrate. This isn't an issue if we lump together parameters.

Choosing $m = kE$ as the maximum velocity, and $s = c/K$ as the saturation ratio, we can instead write

$$v = m\frac{s}{s+1} \tag{3}$$

This is a problem in two variables. Thus we can now sweep m and s, and plot a heat or height map of f.

Another observation. Both our original and lumped parameters are strictly non-negative values, and do tend to vary over order of magnitude. Thus it is possible and convenient to work in logarithmically rescaled parameters[1]. I defined $\alpha = \ln m$ and $\beta = \ln s$. I give these no particular names – just think of them as the magnitudes (exponents) of their corresponding lumped variables. Now,

$$v = \frac{e^{\alpha+\beta}}{e^{\beta}+1} \tag{4}$$

$$= e^{\alpha}\frac{1}{1+e^{-\beta}} \tag{5}$$

This new form is very nice for exploring the parameter space, as we simply need to vary α and β linearly. This is demonstrated in Figure 1.

Unfortunately the objective grows rapidly for large α and β. It's easier to see the overall trend by plotting $\ln f$, as shown in Figure 2. Since $\ln x$ is a monotonic function, this does not change the optimization problem.

Now there is an obvious 'valley' where the objective is minimized, corresponding to the 1D space of values for which combinations of the parameters solve our optimization problem. Already it should be apparent that our choice of lumped parameters was more than a convenience – for large β, the correct choice approaches $\alpha \approx 0$; on the other end, we have $\beta \approx -\alpha$.

This is consistent with our knowledge about the problem. When β is large, the system is highly saturated, and then all that matters is the maximum rate of reaction α. In the opposite regime, near the limit of mass-action kinetics, a decrease in saturation requires a commensurate increase in the maximum rate of reaction to achieve the target rate.

---

[1]There is some weirdness about zero, since the logarithm is not finite at this limit. In practice, this isn't terribly important, although it can lead to some numerical weirdness.
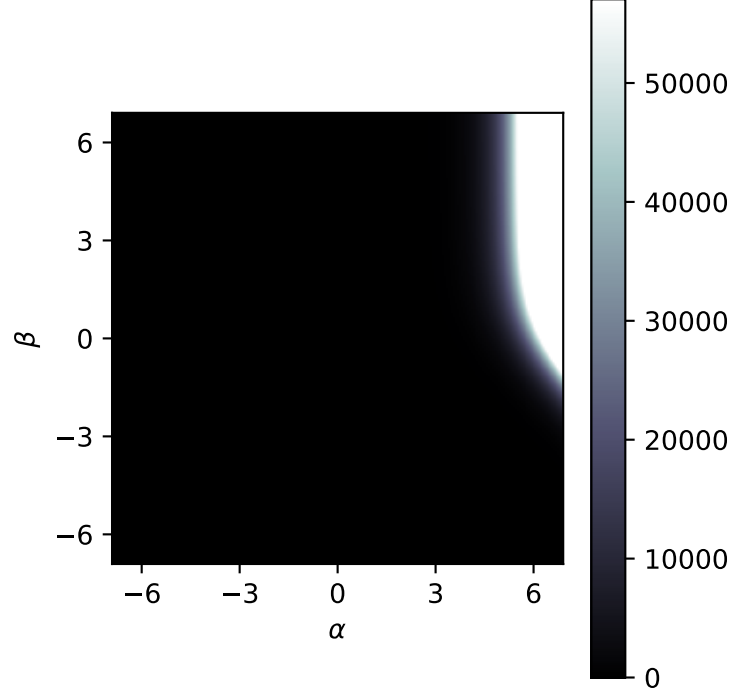
Figure 1: **Raw objective value sweep.**

5 The gradient value space

Most optimization algorithms are, at some level, performing a gradient descent. Even if gradient descent can't result in a global solution, the gradient can still provide useful local information.

In our case, the gradient is just the two partial derivatives $\partial f/\partial \alpha$ and $\partial f/\partial \beta$. Since f is hard to visualize, I again instead look at the gradient on the logarithm of f, which is just

$$\frac{\partial \ln f}{\partial \alpha} = \frac{1}{f}\frac{\partial f}{\partial \alpha} \tag{6}$$

and

$$\frac{\partial \ln f}{\partial \beta} = \frac{1}{f}\frac{\partial f}{\partial \beta} \tag{7}$$

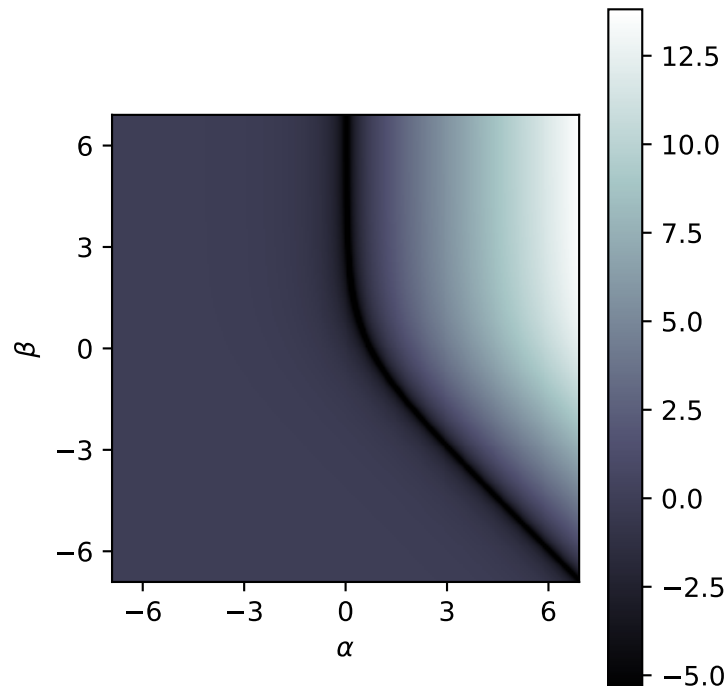These terms are depicted in Figures 3 and 4.

Figure 2: **Logarithmically scaled objective value sweep.**

What we see is consistent with the analysis of Figure 2. Gradients, unfortunately, are somewhat hard to interpret. Another way to look at this is by just computing the direction of the optimal local[2] gradient descent, as shown in Figure 5.

From this depiction we can clearly see that there are four regimes. At the top these correspond to moving 'right' or 'left' in the parameter space, while at the bottom, the best local search direction is diagonal (up-right or down-left). This is easier to see by looking at a flow diagram, as in Figure 6.

The purposes of this exercise are as follows:

- To develop an intuition about the parameter space, and how that corresponds to the objective space.

- To show how a transformation can simplify both the visualization and interpretation of the problem.

---

[2]Note that this is not the *global* best search direction, which would corrspond to snapping directly to the curved line of solutions. In gradient descent, you're only allowed to look at the slope of the ground at your feet, so to speak.
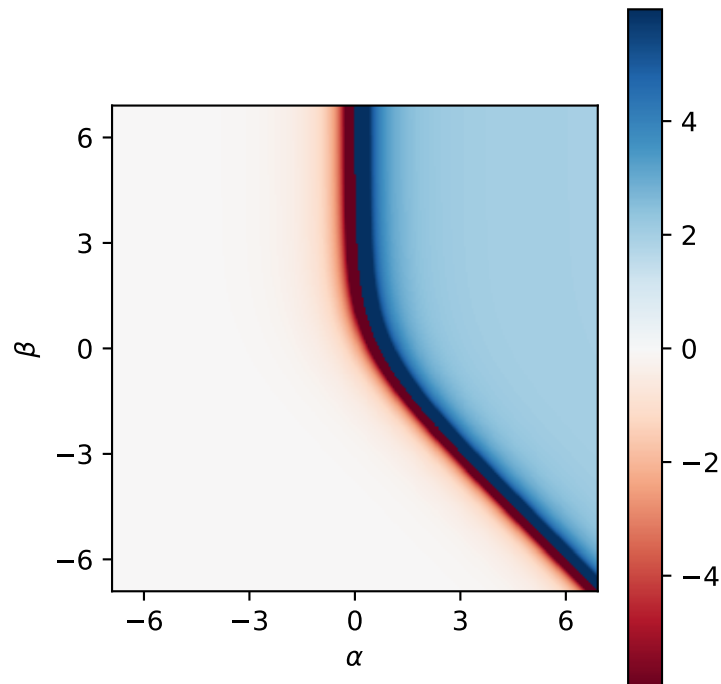
4

Figure 3: **Derivative of objective with respect to α.**

- To demonstrate that α and β are a logical parameter basis for these problems.

6   Parameter space exploration

α and β are not the only possible choices for lumped parameters in this model, but they probably are the best. Remember, α corresponds to the maximum rate of reaction, while β corresponds to the degree of saturation. A third parameter, $\gamma = \alpha + \beta$, corresponds to the rate of reaction in the limit of mass-action kinetics. When divided by the enzme concentration, this term is often called the 'catalytic efficiency'[3]. Any two of α, β, and γ are sufficient to define this system. Which is the 'right' choice?

Let's define a few things.

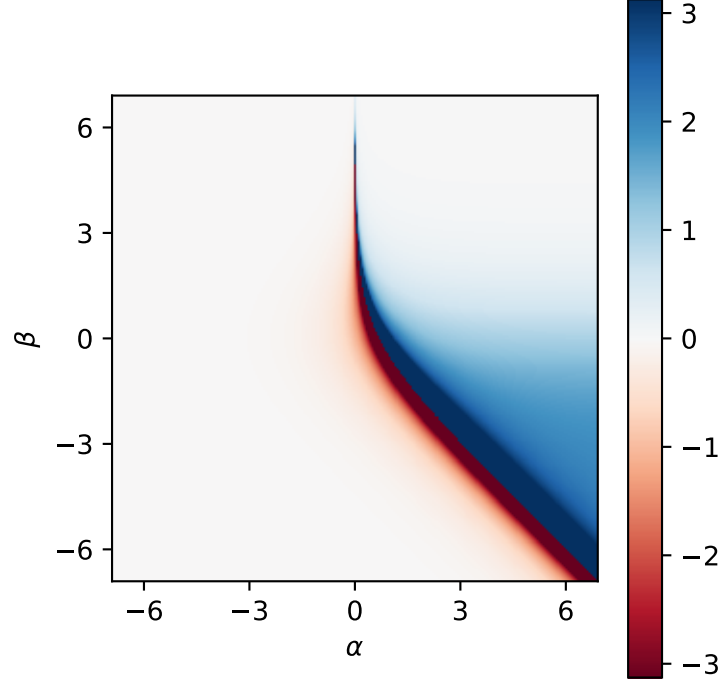[3]For reasons, I'll add, that I don't really understand.

Figure 4: **Derivative of objective with respect to β.**

$$a = \ln k \tag{8}$$
$$b = \ln E \tag{9}$$
$$c = \ln C \tag{10}$$
$$d = \ln K \tag{11}$$

These four new variables are just logarithms of our original independent parameters, and thus themselves are an independent parameter basis. Now,

$$\alpha = a + b \tag{12}$$
$$\beta = c - d \tag{13}$$
$$\gamma = a + b + c - d \tag{14}$$

Already we see something critical – α and β are independent with respect to the underlying independent model parameters, while γ collides with both. In practical terms, we can modify a or b to modify α without disrupting β, and
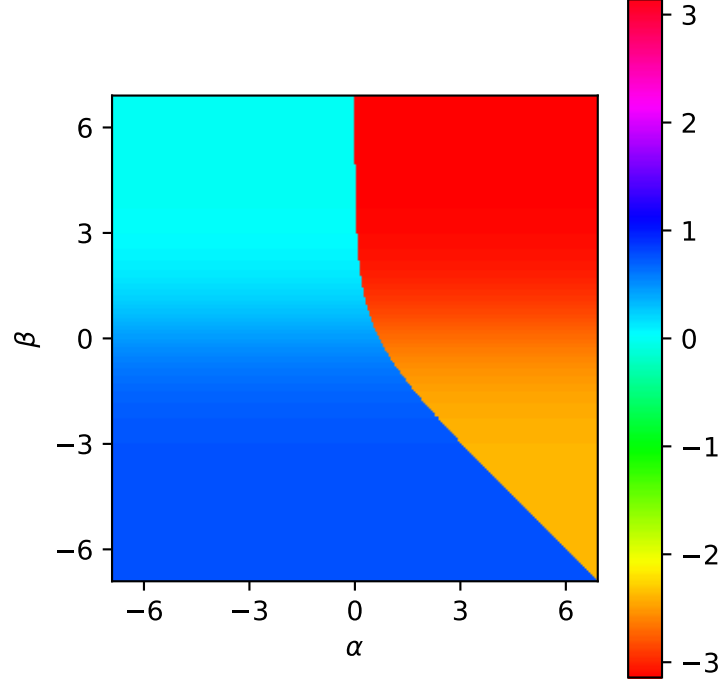
6

Figure 5: **Best local search direction.** The color at each point indicates the direction (in radians, with 0 corresponding to 'right' and moving counter-clockwise) of the direction of gradient descent. Roughly, 'red' corresponds to 'left', 'cyan' corresponds to 'right', 'yellow-green' to 'down', and 'blue-violet' to 'up'.

vice-versa for c and d. Thus, even though they represent lumped quantities, we can safely claim to modify α and β as if they were independent, as it's trivial to resolve that into changes into the underlying model parameters. This isn't the case with γ and α or β.

Regardless, let's bully ahead with γ and see what happens. I'm going to choose α and γ as my lumped model parameters to explore, since they both correspond to rates (but in different model regimes). This gives us the model equation

$$v = \frac{1}{e^{-\alpha} + e^{-\gamma}} \tag{15}$$

The derivation of this form, and its graphical exploration, are left as exercises
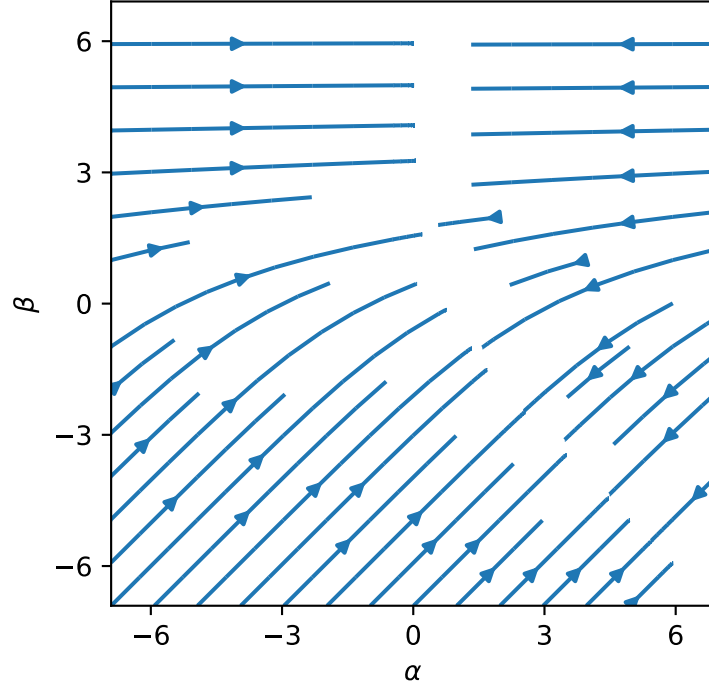
Figure 6: **Gradient descent flow diagram.**

to the reader[4]. I do think this form is quite suggestive, given its symmetric nature.

Now, we can claim to modify these two values independently, but how do we resolve this into real parameter values if their underlying definitions are in fact interdependent? This relation is expressed in matrix-vector form as

$$\underbrace{\begin{bmatrix} \alpha \\ \gamma \end{bmatrix}}_{=y} = \underbrace{\begin{bmatrix} +1 & +1 & 0 & 0 \\ +1 & +1 & +1 & -1 \end{bmatrix}}_{=\begin{bmatrix} a_1 & a_2 \end{bmatrix}^{\mathrm{T}}=A} \underbrace{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}}_{=x} \tag{16}$$

This new matrix form $y = Ax$ will simplify our analysis. Our question, more explicitly, is this: given values of y, how can we find x? This is an underdetermined problem – which is good, but means we need some way to

---

[4]Meaning that I've lost interest in writing any more code. However I promise you that it has an interesting shape.

8

manage the degrees of freedom. Taking the Moore-Penrose pseudoinverse of A, we can find an x for any y as

$$x = A^\dagger y \tag{17}$$

where

$$A^\dagger = \frac{1}{2} \begin{bmatrix} +1 & 0 \\ +1 & 0 \\ -1 & +1 \\ +1 & -1 \end{bmatrix} \tag{18}$$

Now, what exactly is going on here? First off, let's note that the second column in $A^\dagger$ corresponds to β. That's pretty interesting, considering that we haven't brought up β at all in this problem. It fell out naturally.

How about that first vector? Interestingly, the dot product of this vector with the vector of coefficients defing γ is zero – they're orthogonal! So, on a per-column basis, these vector represent independent means of modifying the underlying independent parameters to obtain a change in one of α or β without changing the value of the other.

If you follow that – and I know it's difficult – then it's worth reexamining the case where we choose α and β as our lumped parameters, as in the previous section. Then

$$A^\dagger = \frac{1}{2} \begin{bmatrix} +1 & 0 \\ +1 & 0 \\ 0 & +1 \\ 0 & -1 \end{bmatrix} \tag{19}$$

These columns are just our definitions of α and β, reemphasizing their intrinsically independent nature[5].

This begins to get at the parsimonious perturabtion approach I take in my larger parameter estimtion system. In short – regardless of the basis taken, it's possible to explore the parameter space in a fashion that perturbs combined parameters of interest (like a maximum rate of reaction) while holding others as close to constant as possible.

What I haven't answered – and what is, in truth, a lingering question, is which parameter space is the most natural. The ideal parameter space would reorient the system such that perturbation to the parameters would most reliably and rapidly yield net movement towards the global (true) solution. Apart from empirically testing alternatives, I'm not sure which choice is best. For my current parameter estimation system, I use the equivalents of α and β, though γ-like parameters in place of β-like would be interesting to test.

---

[5]Perhaps it would be more appropriate to start calling these orthogonal.