Anchoring on Extreme Values

Harry Simon (hasimon@stanford.edu)
Symbolic Systems
Stanford, CA 94305 USA

Jocelyn Hickcox (jhickcox@stanford.edu)
Symbolic Systems
Stanford, CA 94305 USA

Gideon Weiler (gweiler@stanford.edu)
Symbolic Systems
Stanford, CA 94305 USA

Abstract

Anchoring bias occurs in numerical estimation when someone biases their estimation towards an initial number that has been presented to them. The anchor-and-adjust theory states that this bias is caused by people iteratively adjusting their estimates away from the initial anchor but failing to do so sufficiently and falling short of the correct value. This paper compares empirical evidence of anchoring effects in numerical estimation with the predictions of a resource-rational model of the anchor-and-adjust process. We show that although this model fits the data for reasonable anchoring values, it fails to accurately model anchoring at extreme values. Our results suggest that at extreme anchors, subjects reject the given anchor and self-generate their own initial estimate to adjust away from.

Keywords: Anchoring, Resource Rational

Introduction

Anchoring bias occurs when subjects place disproportionate weight on an initially considered piece of information when making a judgment. This causes subjects to bias their judgements towards this piece of information, which is called the anchor.

Subjects display anchoring bias even if the information they are given is irrelevant to the current task. Anchoring effects are important to understand, since they explain how behavioural data can deviate from strictly rational analysis of different problem domains in the behavioral sciences.

Although the causes behind anchoring bias is still an open question, a leading theory is the anchoring-and-adjustment model. In this process, subjects start with a reference point (the anchor) and adjust the value away from this reference point according to their knowledge of the domain. We say subjects are biased toward the anchor in that they fail to adjust sufficiently away from the initial value they consider. Studies have shown anchoring bias even when the anchor does not have anything to do with the question at hand.

Whether anchoring bias is a rational process is up for debate. As originally developed, anchoring was proposed as behavioural evidence of peoples deviation from rational choice theory (Tversky and Kahneman 1974). More recently, anchoring has been modeled as a resource-rational algorithm in which subjects produce optimal solutions given time and error costs. In this paper, we attempt to reproduce a resource-rational model that matches empirically observed anchoring data.

Our main interest lies in anchoring effects with extreme or implausible values. We wanted to determine whether the anchoring bias still exists at extreme values, and if so, to what extent. Furthermore, we wanted to test whether the reigning anchor-and-adjust model of anchoring matches empirical data at extreme values.

Background

For example, from (Tversky and Kahneman 1974), subjects were asked to evaluate 8! one of two ways. The first being: $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$; the second being: $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$. Subjects estimated the latter to be much higher than the former primarily. This is postulated because the subjects are presented with larger numbers first.

Moreover, this can be seen with arbitrary anchors as in (Ariely, Lowenstein, and Prelec 2003). Subjects were asked to give the last three digits of their social security numbers and then asked how much money they would take to listen to a sound clip for a certain amount of time. They found that anchoring effects persist even when generated randomly and when the anchor does not even relate to the question being asked.

Not only does the anchor not have to be related to the question being asked, the anchor can also be counterintuitive to such question. (Strack, Mussweiler 1997) ran experiments supporting such notion; they found that asking whether Mahatma Gandhi was 214 years old when he died still have anchoring effects on subjects.

Related Works

The following model is taken from "The Anchoring Bias Reflects Rational Use of Cognitive Resources" by (Lieder et al.). Estimating a numerical value involves a subjects uncertainty about quantity X given his knowledge K and is stated as the conditional probability $P(X \mid K)$, where a person estimates $P(X = x \mid K)$. Following Lieder et al., we approximate peoples estimation of $P(X = x \mid K)$ with an MCMC algorithm.

This model uses the following cognitive architecture with two mental operations. Operation one: take current estimate, and stochastically modify it to generate a new estimate. Operation two: compare posterior probability of new estimate to posterior of old one.

More concretely, a persons initial guess x_0 starts at the anchor value a, which is adjusted in a series of steps. At each step, a potential adjustment, δ, is proposed by sampling from a proposal distribution, which is a Gaussian centered around the current guess. Adjustments are either accepted i.e. $x_{t+1} = x_t + \delta$ or rejected i.e. $x_{t+1} = x_t$.

A proposed adjustment is always accepted if it makes the estimate more probable:

$$P(X = x_t + \delta \mid K) > P(X = x_t \mid K) \tag{1}$$

Otherwise the adjustment is accept with probability:

$$\alpha = \frac{P(X = x_t + \delta \mid K)}{P(X = x_t \mid K)} \tag{2}$$

Or the proportion between the posterior probability of the adjusted estimate and the unadjusted estimate.

The number of iterations is determined by resource-rational optimization. Producing a new estimate incurs a cost due to the time it takes to generate and assess the new value. Because the time needed to generate each new estimate is constant, there is a linear time cost for the number of iterations the model goes through. However, there is also a cost in generating an incorrect guess. As Lieder points out, this cost decreases geometrically for each additional iteration. The optimal number of adjustments can be found by minimizing the sum of the time cost and the error cost. This is called the optimal speed-accuracy tradeoff. Lieder et al. have found that the optimal number of adjustments for anchored estimations tends to be 29 and 8 for unanchored estimations. We use these findings as constant values for the number of iterations our MCMC model performs.

In sum, we approximate subjects numerical estimates by modelling an MCMC algorithm that iteratively adjusts away an initial guess until it is no longer cost-effective to do so.

The Model

Our model is a simplified Church version of the model used in the Lieder 2012 paper. Using an MCMC model as outlined in the Algorithms for Inference chapter in Probabilistic Models of Cognition, we set the initial state to be one of the anchor values used in the experiment and the number of iterations to be 29. The target distribution is a gaussian distribution approximating the unanchored estimations given by the experiment subjects. The proposal function generates a proposed adjustment from a Gaussian distribution centered on the current estimation and with a variance of 10 (value taken from (Lieder 2012)). Both the target distribution and the proposal distribution probabilities are determined using the gaussian probability distribution of the current estimate from the target distribution and the new estimate from the old estimate, respectively.

Experiments

Following the work of Strack, Mussweiler et al, 1997 we looked at how subjects were anchored when estimating the age of Gandhi's death(a figure widely known, but not at that level of specificity). We asked the two following questions:

1. Did Gandhi die before or after the age of (*anchor*)?

2. At what age did Gandhi die?

We tested the following anchors: 1, 20, 40, 60, 100, 120, 140, 220, 400. We also ran an experiment with no anchor. The experiments were ran on Mechanical Turk with 35 subjects per anchor.

The goal of the two scenarios was to test how different anchoring values – high and low – impact the cognitive act of numerical estimation. We chose *reasonable* anchors, even though we were primarily investigating *unreasonable* anchors. The *reasonable* anchors were: 1, 20, 40, 60, and 100 (this plausible at the very least.) The *unreasonable* anchors were: 120, 140, 220, 400.

Experimental Data

Figure 1 shows the anchor and the respective age predictions. The blue line shows a regression based on just the intercept. The red line shows the linear regression based on the anchor. Table 1 shows the anchors and their respective means. Figure 1 shows all of the anchor values plots the subject data. These results support previous findings that anchors do skew subjects decision making. The anchor had a significant effect on the age predictions with: $p < .0001$. This data backs the notion that anchoring biases do exist.

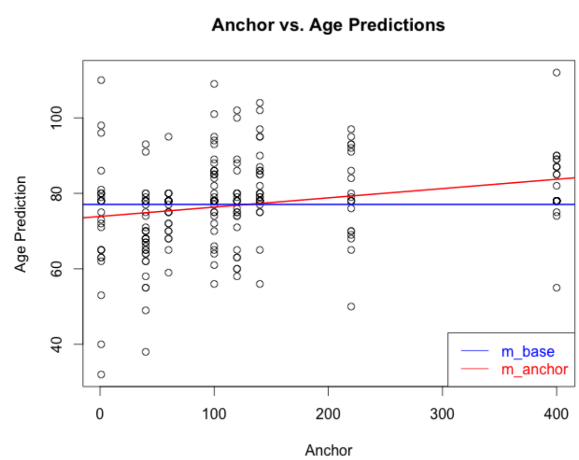

Figure 1: Anchor vs. Age Predictions

Model data

Figure 2 shows the model predictions. The blue line shows the model based on all the anchors. The red line shows the data excluding anchors 220 and 400. The figure shows that

Table 1: Anchor Means

Anchor	Mean
None	78.0
1	73.12
40	69.82
60	75.21
100	79.9
120	76.4
140	82.28
220	79.52
400	81.39

there is a constant steady increase for the age predictions as the anchor increases. The model predicts that an anchor of 400 will predict an age of mean 343.5. The model based on all the anchors over-predicts for the larger anchors.

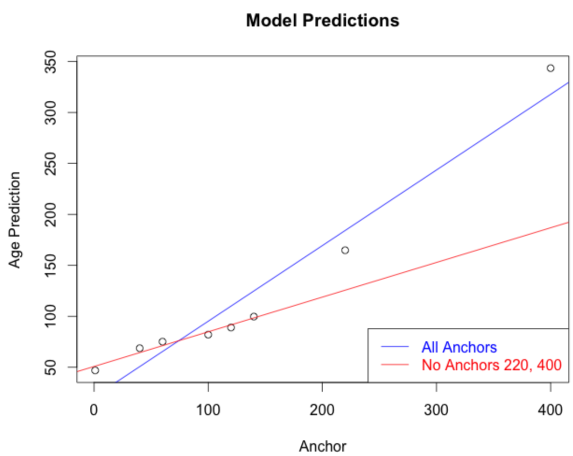

Figure 2: Model Predictions

Figure 3 shows the model predictions when it employs directional adjustments. The model using directional adjustments is more accurate than the model that does not employ directional adjustments. Yet, for the more *unreasonable* anchors, it still over predicts the experimental data. The mean for anchor of 400 is 288.7.

Figure 4 shows the model predictions based on all anchors. The blue line shows the model without directional adjustment. The black line shows the model with directional adjustment. The red line shows the linear regression based on the experimental data.

Figure 5 shows the model predictions without anchors 220 and 400. The red line is the linear regression for the experimental data. The blue line is the linear regression for the model data without directional adjustment. The black line is the model with directional adjustment. Both model predictions increase at a much steeper slope than the experimental

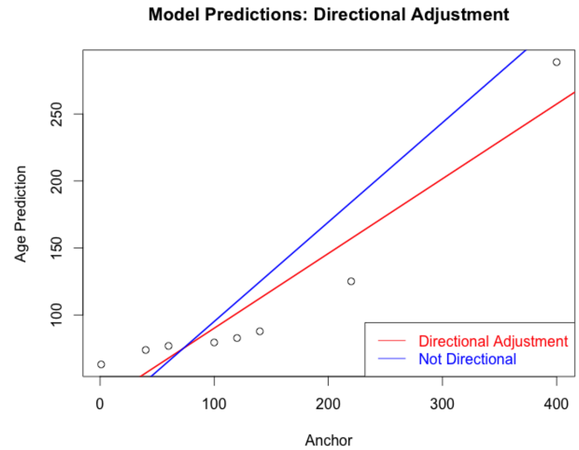

Figure 3: Model Predictions: Directional Adjustment

data. We see that the model with directional adjustments is more accurate.

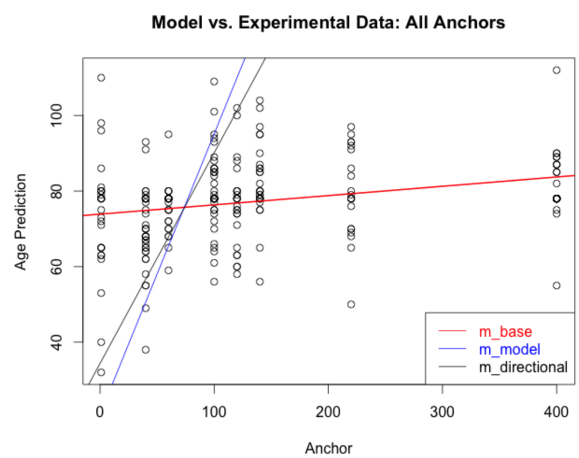

Figure 4: Model vs. Data: All Anchors

Model Results

Overall, the model seems to be a good fit for *reasonable* anchors. The model begins to diverge from the data for the more *unreasonable* anchors such as 220 and 400. The model predicts that the subjects would predict a much larger estimate (heavily skewed by the anchor) than what the data suggests.

We attempted to improve our models predictions at extreme anchors by modifying the way proposed adjustments are generated. For anchors like 400, we increased the variance of the proposal function from 10 to 100 and then 1000 to capture the intuition that the adjustments people consider become greater when the anchor is a round number. The estimates generated by this proposal function for extreme an-

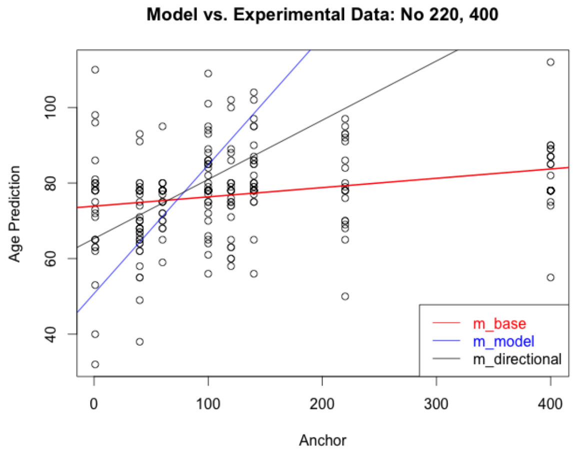

Figure 5: Model vs. Data without Anchors 220, 400

chors were better than the ones before, but they were still outside the range of plausible estimates.

There is evidence that people determine the direction they must adjust in prior to generating adjustments (Simmons et al.), so our second improvement changed the proposal function to sample from the gaussian distribution only in the direction that would produce more likely estimates. This again improved our models estimations at extreme anchors but still did not produce estimates that were close to those from the experiment.

One plausible reason for the divergence between the model and the experimental data is that subjects may self-generate anchors if the anchor presented is too *unreasonable*. For example, for an anchor of 400, the subject might discard this initial anchor and self-generate their own anchor to adjust away from. The distance of an extreme anchor from the correct value can still influence the anchoring bias by biasing the self-generated anchor the subject chooses. In general, the data fits a situation where people discard extreme anchors and self-generate an anchor of around 90.

Future Work

As for moving forward, we could change our experiment design. Instead of asking for a specific age, we could ask for a range of ages. Presumably, the anchors will affect the bounds of the ranges given. Asking for ranges instead of specific ages would allow us to capture how uncertain the subject is and how the anchors affect such uncertainty.

Conclusion

By comparing the predictions of the anchor-and-adjust model to empirical results at extreme anchor values, we have found the anchor-and-adjust model does not work for anchors with extreme values. It is unclear whether the subjects accept the absurd value and make adjustments with greater magnitude, or whether the anchor is discarded and a self-generated anchor used in its place. Moreover, if the anchor is discarded, it is unclear how the new value is generated and how the discarded value impacts the new anchor, if at all. We have shown our model of anchor-and-adjust matches empirical data within reasonable range, but deviates at extreme values. This is an avenue for further research to more fully understand anchoring effects.

Acknowledgments

We would like to thank the wonderful staff of Psych 204: Noah Goodman, Michael Henry Tessler, and Desmond Ong.

References

- Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics*, 118(1), 73–106.
- Lieder, F., Griffiths, T., Huys, Q., Goodman, N., “The Anchoring Bias Reflects Rational Use of Cognitive Resources”.
- Lieder, F., Griffiths, T.L., & Goodman, N.D. (2012). Burn-in, bias, and the rationality of anchoring. In P. Bartlett, F.C.N. Pereira, Leon Bottou, Chris J.C. Burges, & K.Q. Weinberger (editors). *Advances in Neural Information Processing Systems 25*, Lake Tahoe, USA.
- Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: do people adjust from provided anchors? *Journal of personality and social psychology*, 99(6), 917–932.
- Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology*, 73(3), 437.
- Tversky, A. and Kahneman, D., (1974). “Judgment under uncertainty: Heuristics and biases,” *Science*, vol. 185, no. 4157, pp. 1124-1131.