Experiment Design: In Aspect of Parameters

Sungji Wang, Woojin Han[1]

[1]Seoul National University, Seoul 151-747, Korea

May 28, 2024

Abstract

abcd

1 Notations on Decision Models

The extraction of the parameters from the decision made, an appropriate decision model is vital. All possible decision models can be described abstractly as follows. For the set of game states $g \in \mathcal{G}$, the model returns the temptation of each decision $i \in \mathcal{I}$. Note that the word expectation or probability is highly ambiguous to define, due to their broad usage. The probability can be discussed on the infinite trial of the same strategy. If the tactic is deterministic, the selection is fixed by definition to have an ensemble probability to the trivially distributed. If not, the decision distribution is widely spread with a certain temperature. So the return value of the decision model is called temptation, which can be considered as an ensemble distribution of the decisions from the same observation on g. Therefore, the evaluation of the tactic via experiment results can be driven as follows.

$$f : \mathcal{G} \times \mathcal{I} \to \mathbb{R} \tag{1}$$

$$\arg\min_{\Lambda} \left(-\sum_{j} \log\left(f(g_j, i_j); \Lambda\right) \right) = \Lambda_0 \tag{2}$$

The rock-bottom dialectics lie on Eq. 2. The function f is the tactic function, defined on $\mathcal{G}$ and $\mathcal{I}$ simultaneously, which are possible game states and its decision. $f(g, i; \Lambda)$ is the ensemble probability of the decision i made at state g, have parameters of Λ. Using the ideas of estimation parameters by maximum likelihood, the given experimental observation (g_j, i_j) from the patient elaborates the most probable parameters of the decisions. Moreover, if the return value on the decision state is cross-sectionally normalized, the intra-evaluation among the multiple decision models is also possible. In each, $\sum_{\mathcal{I}} f(g, i) = 1$, for all g, the evaluation of different function $f_1(;\Lambda_1), f_2(;\Lambda_2)$ can be mathametically written as $f_3(g, i; \Omega)$, where $\Omega = (\Lambda_1 \cup \Lambda_2)$ forms tactic function respectively. Not the abundant sources of decision models enlight the results. But also the parameters shall have their rigid, explainable character to represent the patient's character too. With the reasonable cutoff, for the total number of experimental observations from a number of rounds r, and total trial t, $|(g_j, i_j)| = N\ r \times t$, gives the ceiling of the number of parameters to use. Roughly the number of parameters should not exceed $N/10$, which covers over-fitting issues primarily. The sharpness of the fitted parameters can also be discussed if the parameter set is smooth enough. The small perturbation of the parameter, $\mathrm{d}\Lambda$ propagates the evaluation difference $\mathrm{d}\Sigma$. The approx amount of $\frac{\mathrm{d}\Sigma}{|\mathrm{d}\Lambda|}$, can be considered as the gradient value, the sharpness of the parameter set. However, the calculation of the gradient value is computationally costly. The recovery plot of the tactic function can therefore grab a place, such that the parameter obtained from simulated decision result (g, i) which is generated by the tactic function itself. In our research, the theoretical bound of the gradient is calculated (todo). The empirical proof of the sharpness is done by parameter recovery simulations. (todo)

2 Game Design Parameters

To increase the sharpness of the parameter fitting, the mathematical application has its limits. Since the index reaction is hard to distinguish when the given condition set g is not varied enough. This means that the clever design of the game boards can acquire more than enough information from patients. We have a total of 6 kinds of hyper-parameters that define the boards. $r, t, (n, m, \sigma_g, \sigma_0, l)$. The number of total rounds, the number of trials in rounds, and the parameters used to generate the boards. So, we have r, t and r different number of $(n, m, \sigma_g, \sigma_0, l)$ pairs correspond to each rounds. In $n \times n$ sized board, σ_g is the covariance size of the Gaussian processes applied, where the mean value is set to m. l is the length coefficient of the covariance matrix in the RBF kernel. After the arbitrary sample from the Gaussian process, the size of σ_0 white noise generates multiple trials from the same rounds. Each parameter has bounds from its nature. The number of rounds $r \times t < 200$ is controlled by the time given to the patient. And the more revealed game state has more evidence of parameters, the number of trials have its priority. $t = 20$, $r = 6$ is the optimal selection from experience which holds enough ratio with the board size $20/n^2$ either. We shall discuss the micro-managing of $(m, \sigma_g, \sigma_0, l)$ beyond the clarification of each parameter's role.

The mean value m can examine the prospect. The smaller the value, the bleaking outlook is expected. If the person playing as same as the board given in small m, he can be considered optimistic. As he is playing the same with the larger m, he might be inactive. With these predictable results, I suggest the challenging circumstances of m value should be experimented with. For example m value for six rounds given $(70, 60, 100, 30, 30, 70)$. Not only does the mean value fluctuate rapidly, the steady state observation might also be possible.

The length parameter l can examine the spatial exploration. With smaller l compared to 1, the unit pixel size, of the board seems to be rough since the reward is not highly correlated via spatial distance. Therefore, the larger l helps the patient explorer far to grasp the shape of the map. But with small l, the far exploration is not effective enough. In my point of view, the l definitely affects the human decision the most, but it does not seem like contains it's character a lot. However, the intellectual difference will highly determine the adjusting learning of l value. Just in case to prevent contamination on other signals, I would like to suggest fixing the l value near 1 2.

The σ_0 parameter hinders the exploit preference. Higher σ_0 makes the board data confused, the exploiting is not easily confirmed with its existence. But if the value exceeds some point, the confusion might push the patient too away from exploitation. I suggest an order of $m/10$ is a rational design, but various σ_0 might not be observed well from patients due to the statistical noise they have. I would like to run the small version for steps to glimpse the parameter nature we can seize.

3 Decision Models

3.1 Basic Gaussian Process UCB model

(todo) we all know this so far :)

3.2 Expectation z-value Model

The basic Gaussian Process UCB model has several different flaws while performing analysis of real data. The UCB model claims that the person should choose the farthest part which they have not explored yet, to gain more information. This seems logical, but the tactic does not count on the unrevealed region having a lack of chances, but to aim higher than expected. On real human explorations, they are tempted to be localized from their decisions, since they expect the nearby area to have rich information. For the first layer of game board analysis, it is quite logical to choose the Gaussian process. So, let's denote $\hat{m}, \hat{\sigma}(\mathcal{I}; g, \Lambda)$ as the mean value and the deviation value after performing GP at the board g, and some parameter Λ. At the basic GP UCB model, the temptation function can be written as an equation 3.

$$f(g, i; \Lambda \cup \{\beta, \tau\}) = \operatorname{soft\,max}_{\tau, \mathcal{I}} \left((\hat{m} + \beta\hat{\sigma}) (i; g, \Lambda) \right) \tag{3}$$

Which will explore enough until the maximal $\hat{\sigma}$ value indicates the exploration. Rationally, the exploitation-exploration rate should be controlled by the trials left either, not only the $\hat{\sigma}$ maximum value takes place. Above those problems, I suggest the expectation z model. The function can be written as equation 4.

$$f(g, i; \Lambda \cup \{E, \tau\}) = \operatorname{soft\,max}_{\tau, \mathcal{I}} \left(\frac{\hat{m} - E}{\hat{\sigma}} (i; g, \Lambda) \right) \tag{4}$$

The tactic maximizes the probability of the score being revealed larger than E every time. So the value of E is not the exact expectation value, which is $\hat{m}$ for the GP result, it can be called as desired value.

3.3 Meta RL

(todo) not familiar with this part, and also do not have enough time :(