Problem Set 1

Akira Masuda

2020/1/19

`Course: MACS30100 Perspectives on Computational Modeling (Winter 2020)`
`Author: Akira Masuda (ID: alakira)`

Building models

Deviant aggressive behavior

1. **Theory I** focuses on people's experience and assumes that the experience would change their behavior. Thus, based on Theory I, implementing a policy of severe punishment for deviant aggressive behavior would prevent these behaviors. For example, increasing the number of police officers or making additional laws that prohibits certain behaviors would decrease them by affecting their experiences. In addition, praising one's good behavior also affect their experience and thus navigates their behavior to more favorable way. However, since those punishments or praises only affect the behavior of people who were already caught or done desirable deed, it does not affect those people who first attempt a deviant behavior. To prevent one's initial deviant conduct, we may implement additional early education which teaches individuals about the punishment and praising through virtual experience.

   Theory II assumes deviant aggressive behavior comes from personal emotion. It suggests a policy that approaches their emotions by ameliorating their behaviors. In concrete, counseling on those who had emotionally been hurt would recover their emotions and prevent them from conducting deviant behaviors. However, this would be only a temporary expedient. It would not be sufficient until we get rid of the root cause which makes them hostile. Therefore, additional social policy which changes social structure and mitigates stress factor which makes people angry is needed. One example could be increasing job market mobility which enables people to change their job if they feel uncomfortable.

   Theory III rooted in the problem of inequality. It suggests that people conduct deviant behavior based on rational calculation accounting how they would gain benefit by not doing the deviant action. If people are in the most oppressive environment, the benefit would be small and thus they would rationally be better off by conducting deviant behavior. Given the assumption, a policy which mitigates inequality and thus improves the most oppressive environment would mitigate the deviant behavior. For example, redistribution of wealth by tax and subsidy would fit in this rationale. Although it is impossible to result in the perfect equality and thus there always remains people who are under-represented and are discriminated under the system, we yet can achieve certain condition where those people gains enough benefit from the system and thus rationally choose not to conduct deviant behavior.

   Theory IV explains deviant behavior as social role. One would conduct a deviant behavior because that is what the individual is supposed to do in their social circumstances. One example is the behavior of members of mafia. Their behavior is navigated by the mafia organization's purpose. Their behavior is explained by division of labor of mafia among members. In this setting, approaching each individual would not work well because the structure of mafia still remains which reproduce similar behavior again. To address this root cause, policy has to focus on approaching the structure of deviant subculture. On the example of mafia, imposing a dissolution order of mafia organization would make it hard for mafia to gather and plan for deviant action. In addition, those subcultures often have hierarchical structure which is characterized by heterogeneous importance in the human network. The boss of higher-class member in mafia has more power and thus has more influence on other members. Given this characteristics of network, approaching those top-ranking member would be more efficient way to cope with the whole network.

Waiting until the last minute

1. a. From my experience, I usually have other tasks which are more urgent to do, so that certain things would be untouched until the last minute. Deciding whether to do certain things early or not is based on the expected time that I have to spend. As is often the case with prior prioritization, there are chances that it actually takes more time than I expected.

   b. Assume an individual has utility function:

$$U(L, S) = aL - bS$$

   where L is time allotted to leisure and S is time allotted to the task (e.g. study, work). a and b are constant term which is greater than or equal to 0. The individual has time T which he divides it into L and S, that is: $T = L + S$. Assuming that S cannot be divided, that is, the individual cannot work intermittently. The performance of the individual working on the task per time is $e(t)$, which is thought as efficiency and is changed by the time.
   Suppose that efficiency function is a monotonically increasing function:

$$\frac{\partial e(t)}{\partial t} \geq 0$$

   This assumption stands from the fact that when you start later, you would have more information to cope with the task. For example, when student start working for assignment later, he would start with more information about the assignment through colleague's questions or comments.
   The constrain is that he has to get the task done:

$$H = \int_{t'}^{t'+S(t')} e(t)dt$$

   where H is the amount of the task and t' is the time he starts the task $0 \leq t' \leq T - S$. Notice that the time allotted to the task (S) is a function of t'.
   Suppose the individual has perfect information so that he knows the shape of $e(t)$ and decides t' at the time of 0. Given above assumptions, the individual choose t' which maximizes his utility. From the time constrain and utility function:

$$\max U(t') = aT - aS - bS = aT - (a + b)S(t')$$

   Since a, b, and T are constant, $U(t')$ gets maximum when $S(t')$ gets minimum. Therefore, we have to solve $\min S(t')$ given the constrain of H. With some calculation, the optimal t' is the time which satisfy $t' + S(t') = T$, which means the individual decide to work at the latest period of the time range.

   c. Similar to the model of (b), an individual has its utility function but this time the value of leisure is different in time:

$$U(L, S) = \int_{t'}^{t'+L} l(t)dt + bS$$

where t' is the time the individual start to spend for leisure, L is the time of leisure, b is constant, S is the time of work. This time, L cannot be divided. $l(t)$ is a monotonically decreasing funtion of t that represents how the individual benefits from having leisure for one unit of time:

$$\frac{\partial l(t)}{\partial t} \leq 0$$

The difference from (b) is that the value of the time for leisure is decreasing by time, that is, the individual thinks the immediately following time as the most preferable time to spend on leisure. This can be justified by recent study of behavioral economics on Myopia bias.
As of (b), the constraint are on time: $T = L + S$, and the task: $S \cdot e = H$ where e is constant value that represents the efficiency of work. The individual has perfect information so that he knows the shape of $l(t)$ and decides t' at the time of 0.
Given above assumptions, the individual choose t' which maximizes his utility. Again, with some calculation, the optimal t' is that $t' = 0$, which means that the individual would have leisure first, and do the work at the end.

d.

- From the model of (b), it can be predicted that even the efficiency increased, it does not change the individual's behavior of waiting until the last minute. With the notation from (b), even $e(t)$ shifted upward by exogeneous factor (e.g. gaining knowledge), because it does not change its monotonically increasing characteristic, the results would not change. (This can be show by the same calculation did in (b))

- Also from the model of (b), the behavior of waiting until the last minutes would not change by the amount of task. While it would increase the total time that the individual should spend on the task ($S(t')$) and decreases t', the time is still the latest among the possible starting time.

- From the model of (c), the time he starts doing the work depends on the shape of $l(t)$. If we relax our assumption on $l(t)$ and assume that it is some increasing function, the individual would start woriking at the beginning. This implies that not the property of work but the property of leisure changes his behavior. Therefore, if we want him to start his work early, we have to impose some better benefit which he would gain only from the later leisure.

- Also from the model (c), any change on magnitude of $\frac{\partial l(t)}{\partial l}$ does not change the behavior as long as it is monotonically decreasing. This implies that even one's perception is close to neutral, he would do the task in the last minute. However, at the point where the function flips to increasing function, the individual's behavior changes drastically, from the last minute to the first minute. Therefore this model predicts that we mostly observe these two types of behavior.

Selecting and fitting a model

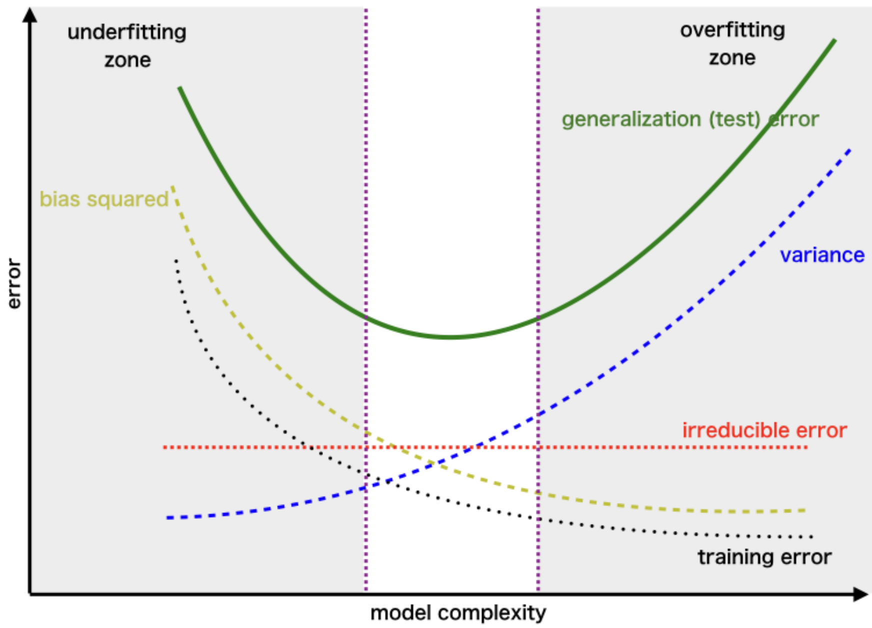
1. a. When the sample size n is extremely large and the number of predictors p is small, the risk of overfitting is relatively small. Therefore, a flexible method preferable since it might perform better than an inflexible method.

   b. When the number of predictors p is extremely large and the number of observations n is small, there is a huge risk of overfitting when using flexible method. Therefore, an inflexible method is preferable than a flexible method.

   c. When the relationship between the predictors and response is highly non-linear, a flexible method would fit well with the data since it has more degrees of freedom. Therefore, a flexible method is preferable than an inflexible method.

   d. When the variance of the error terms, i.e. $\sigma^2 = Var(e)$, is extremely high, it means that there are a lot of noise which might cause overfitting. Therefore, an inflexible method is preferable than a flexible method.

2.



a.

- **Bias**: The bias curve monotonically decreases as the model complexity increases. This is because the model fits well to the observations as it becomes more complex function.

- **Variance**: The variance curve monotonically increases as the model complexity increases. This is because that the more complex (flexible) model becomes, the more prediction becomes closer to each observation, that is, the prediction changes drastically with different observation set.

- **Training Error**: The training error curve monotonically decreases as the model complexity increases. As the model complexity increases, the model would become more flexible to fit the data, meaning that it reduce more prediction errors with the same dataset.

- **Test Error**: The test error curve would first decreases as the model complexity increases, but at certain point, it starts to increase. When the model complexity is relatively low, the model fits well with the test data like it fits well with the training data. However, when the model becomes too complex, it would adapt with the training data too much, which is known as 'overfitting', the model would not fit well with the test data.

- **Irreducible Error**: The irreducible error is constant, which is shown in a horizontal line because this error ($Var(\epsilon)$) cannot be mitigated by model complexity. This line is under the test error curve because the test error consists of reducible error and irreducible error.