Estimating Forward Looking Return Distribution With Generalized Recovery Theorem

Masatake Ito†
Graduate School of Science and Technology
Keio University, Yokohama, Japan
Tel: (+81) 45-566-1454, Email: masatakeitoh@gmail.com

Takuya Kiriu‡
Mitsubishi UFJ Trust Investment Technology Institute Co., Ltd.
Tokyo, Japan
Tel: (+81) 3-5572-0861, Email: kiriu@mtec-institute.co.jp

Norio Hibiki
Faculty of Science and Technology
Keio University, Yokohama, Japan
Tel: (+81) 45-566-1635, Email: hibiki@ae.keio.ac.jp

Abstract. Ross (2015) has showed "Recovery Theorem", which enables us to estimate real world distribution from state price. Under the assumption of time-homogeneous Markov economy, we can estimate forward looking return distribution from option price by applying the Recovery Theorem. However, it is not easy to obtain accurate estimates because it is necessary to solve an ill-posed problem in the estimation process. Kiriu and Hibiki (2016) has proposed a method to stabilize the solution by configuring the regularization term considering prior information, and showed the effectiveness of the method. Recently, Jensen *et al.* (2016) propose "Generalized Recovery Theorem" by relaxing the assumption of time-homogeneity of state price. By applying this new theorem, we can recover real world transition probabilities from a current state to all future states over different time horizons. Nevertheless, it is difficult to obtain appropriate estimates because it is necessary to solve an ill-posed problem as with the Recovery Theorem. In this research, we propose a new estimation method to stabilize the solution by giving prior information about a real world distribution, and a setting method of prior information using observed data. Furthermore, we verify the effectiveness of the proposed method through numerical experiments using hypothetical data.

Keywords: Generalized Recovery Theorem, probability distribution estimation, regularization method

1. INTRODUCTION

Ross (2015) has showed a theorem that enables us to drive a representative investor's risk preference and a real world distribution from risk neutral distribution under the assumption that there is a representative investor with time-separable utility. This theorem is named Recovery Theorem (RT). By applying RT under the assumption that state prices follow time-homogeneous Markov chain, we can recovery forward looking return distribution from option prices. However, it is pointed out that a recovered real world distribution is unstable because it is necessary to solve an ill-posed problem in the estimation process. Kiriu and Hibiki (2016) have extended the estimation method of Audrino *et al.* (2015) which use Tikhonov regularization, and proposed a method to stabilize the solution by configuring the regularization term considering risk neutral distribution as prior information. Furthermore, Kiriu and Hibiki (2016) have showed the effectiveness of the method through numerical experiments using hypothetical data.

Recently, Jensen *et al.* (2016) propose a new theorem, Generalized Recovery Theorem (GRT), by relaxing the assumption of time-homogeneity of state prices in RT. Compared with RT, it is expected that we can eliminate bias caused by the assumption of time-homogeneity of state prices by estimating real world distribution with GRT. Jensen *et al.* (2016) indicate we can apply GRT when we parameterize a representative investor's risk preference, and

† :Corresponding Author
‡: The views expressed in this paper are those of the authors and do not necessarily reflect the official views of Mitsubishi UFJ Trust Investment Technology Institute Co., Ltd.

then verify the predictive power of the expected return recovered from S&P500 option data.

In this research, we discuss the method to estimate a real world distribution with GRT. We propose a new estimation method to stabilize the solution by giving prior information about a real world distribution, because it is difficult for conventional method to estimate a real world distribution accurately. And then we show a method of setting prior information using the information contained in observed data. Furthermore, we verify the effectiveness of the proposed method through numerical experiments using hypothetical data.

We find the following two points through the numerical experiments. (1) We show a possibility that a real world distribution recovered with our proposed method is more accurate than the distributions both given as priori information and recovered with conventional estimation method. (2) Proposed method can improve the estimation accuracy by configuring prior information in accordance with the observed data.

2. GENERALIZED RECOVERY THEOREM

In this section, we summarize GRT, which Jensen *et al.* (2016) have showed. We assume an arbitrage free and complete market in a discrete time with a finite state multi-period model, and then define market state $s(=1,\cdots,s_0,\cdots,S)$ as r_s, which are underlying stock returns from current time and state, thus $r_{s_0}=0\%$. $\boldsymbol{\Pi}\equiv(\pi_{\tau,s})$ is a $T\times S$ spot state price matrix. $\pi_{\tau,s}$ denotes a state price of going from the current state s_0 to state s in τ periods. Similarly, we define a $T\times S$ risk neutral probability matrix $\boldsymbol{Q}\equiv(q_{\tau,s})$, a $T\times S$ real world probability matrix $\boldsymbol{P}\equiv(p_{\tau,s})$ and a $T\times S$ pricing kernel matrix $\boldsymbol{M}\equiv(m_{\tau,s})$.

We suppose a spot state price matrix $\boldsymbol{\Pi}$ is known because it can be estimated from option price. $q_{\tau,s}$ is derived from $\pi_{\tau,s}$ since $q_{\tau,s}$ and $\pi_{\tau,s}$ are related in the following formula.

$$q_{\tau,s}=\frac{\pi_{\tau,s}}{\sum_{s=1}^{S}\pi_{\tau,s}}\qquad(\tau=1,\cdots,T;s=1,\cdots,S)\qquad(1)$$

On the other hand, the relationship among $\pi_{\tau,s}$, $m_{\tau,s}$ and $p_{\tau,s}$ is expressed as

$$\pi_{\tau,s}=m_{\tau,s}p_{\tau,s}\qquad(\tau=1,\cdots,T;s=1,\cdots,S)\qquad(2)$$

Therefore, we need the values of not only state prices but also pricing kernel to obtain real world probabilities.

We can recover pricing kernel and real world probabilities from state prices with GRT by assuming that there is a representative investor with time-separable utility. Under this assumption, pricing kernel is given by

$$m_{\tau,s}=\delta^{\tau}\left(\frac{u_s}{u_0}\right)=\delta^{\tau}h_s\qquad(\tau=1,\cdots,T;s=1,\cdots,S)\qquad(3)$$

where $\delta\in(0,1]$ is a discount factor of the utility, $u_s>0$ is a marginal utility at state s, h_s is a normalized marginal utility so as to satisfy the condition that the marginal utility at state s_0 equals one. The following equation is obtained by substituting Equation (3) into Equation (2).

$$\pi_{\tau,s}=\delta^{\tau}p_{\tau,s}h_s\qquad(\tau=1,\cdots,T;s=1,\cdots,S)\qquad(4)$$

We define a S-dimensional diagonal matrix of marginal utilities $\boldsymbol{H}=\mathrm{diag}(h_1,h_2,\cdots,h_S)$ and T-dimensional diagonal matrix of discount factor $\boldsymbol{D}=\mathrm{diag}(\delta,\delta^2,\cdots,\delta^T)$. Equation (4) is represented using the following matrices.

$$\boldsymbol{\Pi}=\boldsymbol{DPH}\qquad(5)$$

Since $\boldsymbol{P}$ is a probability matrix, we can write $\boldsymbol{Pe}=\boldsymbol{e}$, where $\boldsymbol{e}=(1,\cdots,1)'$ is a vector of ones. The following expressions can be calculated, multiplying both sides of Equation (5) by $\boldsymbol{H}^{-1}\boldsymbol{e}$.

$$\boldsymbol{\Pi H}^{-1}\boldsymbol{e}=\boldsymbol{DPe}=\boldsymbol{De}\qquad(6)$$

$$\begin{bmatrix}\pi_{1,1}&\cdots&\pi_{1,S}\\\vdots&\ddots&\vdots\\\pi_{T,1}&\cdots&\pi_{T,S}\end{bmatrix}\begin{bmatrix}h_1^{-1}\\\vdots\\h_{s_0-1}^{-1}\\1\\h_{s_0+1}^{-1}\\\vdots\\h_S^{-1}\end{bmatrix}=\begin{bmatrix}\delta\\\vdots\\\delta^T\end{bmatrix}\qquad(7)$$

We can find a solution of Equation (7) for $S\le T$ because we have T equations and S unknowns, which are $\delta,h_1^{-1},\cdots,h_{s_0-1}^{-1},h_{s_0+1}^{-1},\cdots,h_S^{-1}$. The solution can be obtained by minimizing the differences of both sides in Equation (7). This equation is linear except for δ^{τ}. To make the equation more simple, we linearize δ^{τ} around δ_0. Based on a Taylor expansion, we write $\delta^{\tau}\approx\delta_0^{\tau}+\tau\delta_0^{\tau-1}(\delta-\delta_0)=a_\tau+b_\tau\delta$ where $a_\tau=-(\tau-1)\delta_0^{\tau},b_\tau=\tau\delta_0^{\tau-1}$. Equation (7) becomes in the following:

$$\begin{bmatrix}-b_1&\pi_{1,1}&\cdots&\pi_{1,s_0-1}&\pi_{1,s_0+1}&\cdots&\pi_{1,S}\\\vdots&\vdots&\ddots&\vdots&\vdots&\ddots&\vdots\\-b_T&\pi_{T,1}&\cdots&\pi_{T,s_0-1}&\pi_{1,s_0-1}&\cdots&\pi_{T,S}\end{bmatrix}\begin{bmatrix}\delta\\h_1^{-1}\\\vdots\\h_{s_0-1}^{-1}\\h_{s_0+1}^{-1}\\\vdots\\h_S^{-1}\end{bmatrix}=\begin{bmatrix}a_1-\pi_{1,s_0}\\\vdots\\a_T-\pi_{T,s_0}\end{bmatrix}\qquad(8)$$

Rewriting Equation (8) in matrix form as

$$\boldsymbol{Bh}_{\delta}=\boldsymbol{a}_{\pi}.\qquad(9)$$

A discount factor of the utility and marginal utilities,

$\delta, h_1^{-1}, \cdots, h_{s_0-1}^{-1}, h_{s_0+1}^{-1}, \cdots, h_S^{-1}$, can be estimated by minimizing the differences of both sides in Equation (9). Specifically, the following optimization problem is solved.

$$\min_{\boldsymbol{h_\delta}} \quad \| \boldsymbol{B h_\delta} - \boldsymbol{a_\pi} \|_2^2 \qquad (10)$$

subject to $\quad 0 < \delta \leq 1 \qquad (11)$

$$h_s^{-1} > 0 \quad (s = 1, \cdots, s_0 - 1, s_0 + 1, \cdots, S) \qquad (12)$$

Finally, we can calculate a real world probability $p_{\tau,s}$ as

$$p_{\tau,s} = \frac{1}{\delta^\tau} h_s^{-1} \pi_{\tau,s} \qquad (\tau = 1, \cdots, T; s = 1, \cdots, S) \qquad (13)$$

3. ESTIMATION METHOD

In this section, we show the specific methodology to derive a real world probability matrix $\boldsymbol{P}$ from option prices with GRT. We discuss the estimation method under the assumption of $S \leq T$ hereafter.

3.1 Estimating State Prices

A state price matrix $\boldsymbol{\Pi}$ can be estimated by the method proposed by Breeden and Litzenberger (1978). The state price function $\pi(\tau, k)$ is expressed as

$$\pi(\tau, k) = \frac{\partial^2 c(\tau, k)}{\partial k^2} \qquad (\tau = 1, \cdots, T) \qquad (14)$$

where k is a strike price and $c(\tau, k)$ is the function of call option price. We can obtain $\pi_{\tau,s}$ numerically by discretizing $\pi(\tau, k)$.

3.2 Proposed Method

The condition number of 21×21 matrix $\boldsymbol{B}$ estimated using S&P500 option data from 2000 to 2016 is a very large value of 2.7×10^{17} in average. Therefore, optimization problem (10) - (12) is ill-posed. The ill-posed problem has the bad characteristics that the solution is highly sensitive to a noise because of low independency of equations. It is generally difficult to derive an accurate solution from the ill-posed problem. Consequently, the solutions are not stably obtained, and a real world distribution estimated by a conventional estimation method is inappropriately distorted.

In this research, we propose a new estimation method to stabilize the solution by giving prior information about a real world distribution. Specifically, we reformulate the object function as

$$\| \boldsymbol{B h_\delta} - \boldsymbol{a_\pi} \|_2^2 + \zeta_1 \left(\delta - \bar{\delta}\right)^2 + \zeta_2 \sum_{s=1, s \neq s_0}^{S} \left(h_s^{-1} - \bar{h}_s^{-1}\right)^2 \qquad (15)$$

where ζ_1 and ζ_2 are regularization parameters, $\bar{\delta}$ is prior information of a discount factor of the utility and $\bar{h}_s^{-1}$ are those of marginal utilities. The first term denotes the original objective function. The second and third terms are regularization terms, which penalize the differences between the estimates and prior information. When $\bar{\delta}$, $\bar{h}_s^{-1}$ are given, we rewrite Equation (13) into

$$\bar{p}_{\tau,s} = \frac{1}{\bar{\delta}^\tau} \bar{h}_s^{-1} \pi_{\tau,s} \qquad (\tau = 1, \cdots, T; s = 1, \cdots, S) \qquad (16)$$

Therefore, the proposed method is regarded as the way of solving a problem by giving prior information about a real world distribution.

3.3 Priori Information

We show a method of setting prior information using the information contained in observed data though various kinds of prior information are examined.

3.3.1 The Discount Factor of the Utility

We examine the method of using a risk-free discount factor implied in state prices as the priori information about the discount factor of the utility. Since the sum of state prices equals the risk-free discount factor, $\bar{\delta}$ can be set as in the following optimization problem.

$$\min_{\delta} \quad \sum_{\tau=1}^{T} \left(\bar{\delta}^\tau - \sum_{s=1}^{S} \pi_{\tau,s} \right)^2 \qquad (17)$$

The discount factor of the utility equals the risk-free discount factor when the representative investor is risk neutral. These values do not generally coincide with each other, but we suppose the difference is small. We also set $\bar{\delta}$ as the specified point δ_0 in a Taylor expansion.

3.3.2 Marginal Utility

We suppose a CRRA utility as a representative investor's risk preference. We substitute $\delta = \bar{\delta}, h_s^{-1} = (1 + r_s)^{\gamma_R}$ into the optimization problem (10) - (12) and solve for γ_R. We set the solution as prior information of the marginal utility. In this method we attempt to find a more accurate solution by setting prior information with respect to representative investor's risk preferences estimated under the assumption of parametric utility. We can flexibly express the representative investor's risk preferences implied in observed option data since the non-parametric solution is obtained.

4. NUMERICAL EXPERIMENTS

In this section we examine the estimation accuracy of the proposed method by comparing a preset hypothetical real world probability matrix $\boldsymbol{P}^H$ with a real world probability matrix $\boldsymbol{P}^E$ estimated from data with noise. The reason we conduct numerical experiments by using hypothetical data is that we do not know a true real world probability matrix from real option data and it is difficult to verify the accuracy of estimation.

4.1 Procedure of Analysis

Figure 1 represents an overview of the analysis. Firstly, we give the hypothetical real world probability matrix $\boldsymbol{P}^H$, the discount factor of the utility δ^H, marginal utility h_s^H as true values. We calculate $\boldsymbol{\Pi}^H$ in Equation (5). Given market data with noise, we generate the state price matrix $\boldsymbol{\Pi}^N$, which is expressed as

$$\pi_{\tau,s}^N = \pi_{\tau,s}^H(1+\sigma e_{\tau,s}) \qquad (\tau = 1,\cdots,T; s = 1,\cdots,S) \quad (18)$$

where σ is a noise parameter. $e_{\tau,s}$ follows an independent and standard normal distribution. We obtain estimated values of the discount factor δ^E, marginal utilities h_s^E and the real world probability matrix $\boldsymbol{P}^E$ from $\boldsymbol{\Pi}^N$ with proposed or conventional method. Finally, we evaluate the estimation accuracy by measuring the differences between $\boldsymbol{P}^H$ and $\boldsymbol{P}^E$.

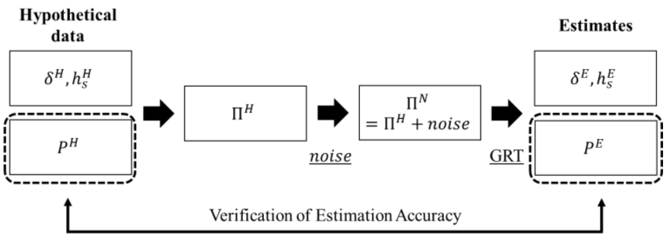


Figure 1: Summary of the analysis

4.2 Setting

4.2.1 State and Period

Market return is defined by underlying asset returns from state s_0 at period 0. We provide 21 returns in total placed by 5% symmetrically from the return of 0%. Specifically, $r_1 = -50\%, r_{s_0} = r_{16} = 0\%, r_{21} = 50\%$. For simplicity, we suppose $T = S = 21$.

4.2.2 Evaluation Criteria

The estimation accuracy is evaluated by the Kullback-Leibler divergence (KL divergence), which measure the difference between two probability distributions. The KL divergence of the estimated distribution $\boldsymbol{P}^E$ from the hypothetical distribution $\boldsymbol{P}^H$ is expressed as

$$D_{KL}(\boldsymbol{P}^E \parallel \boldsymbol{P}^H) = \sum_{\tau=1}^{T}\sum_{s=1}^{S} p_{\tau,s}^E \ln\left(\frac{p_{\tau,s}^E}{p_{\tau,s}^H}\right). \quad (19)$$

We evaluate the ratio of KL divergence of the estimated distribution $D_{KL}(\boldsymbol{P}^E \parallel \boldsymbol{P}^H)$ to KL divergence of the risk neutral distribution $D_{KL}(\boldsymbol{Q}^E \parallel \boldsymbol{P}^H)$. We develop the analysis with 100 sets of random numbers and calculate the average of the ratio $\bar{D}_{KL}$. $\bar{D}_{KL} < 1$ shows that the estimation accuracy is improved by a risk adjustment with GRT. On the other hand, $\bar{D}_{KL} > 1$ shows that the estimation accuracy is degraded by the risk adjustment.

We set $\zeta_1 = 10^{-2}$ for simplicity because it is little sensitive to the estimation accuracy.

4.2.3 Hypothetical Data

The hypothetical real world probability matrix $\boldsymbol{P}^H$ is generated based on S&P500 historical daily return data from January 3, 1950 to December 30, 2016. Firstly, we set a reference date and then calculate returns using rolling windows with the size of 30, 60,⋯, 630 calendar days. Then we generate a matrix based on the number of state transitions from the return reference date. A return less than or equal to -47.5% (greater than or equal to 47.5%) is assigned to state 1 (state 21). The reference date is rolling on a daily basis from January 3, 1950. Finally, we sum up all the matrices, and adjust each element of the matrix so that it can be a stochastic matrix, which sum of row elements is equal to one.

The hypothetical discount factor of the utility δ^H and marginal utility h_s^H are generated based on S&P500 historical return data and option data. We calculate a state price matrix $\boldsymbol{\Pi}$ from an option data on March 1, 2017 as shown in the section 3.1. The solutions of the following optimization problem under the constraints (11) and (12) are applied to hypothetical data δ^H, h_s^H.

$$\min_{\delta, h_s} \parallel \boldsymbol{\Pi} - \boldsymbol{DPH} \parallel_2^2 \quad (20)$$

Figure 2 represents the obtained pricing kernel $m_{\tau,s}^H$. Theoretically, pricing kernel decreases monotonically when a representative investor is risk averse. However, it is pointed out that the pricing kernel estimated from real data is not a monotonic function, which partially rises as shown in a lot of previous studies (Jackwerth (2000) and Rosenberg and Engle (2002), etc.). The form of the hypothetical pricing kernel coincides with the form of pricing kernel estimated empirically in previous studies.

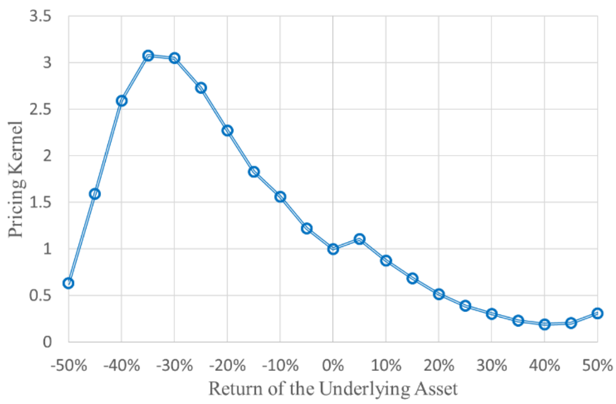
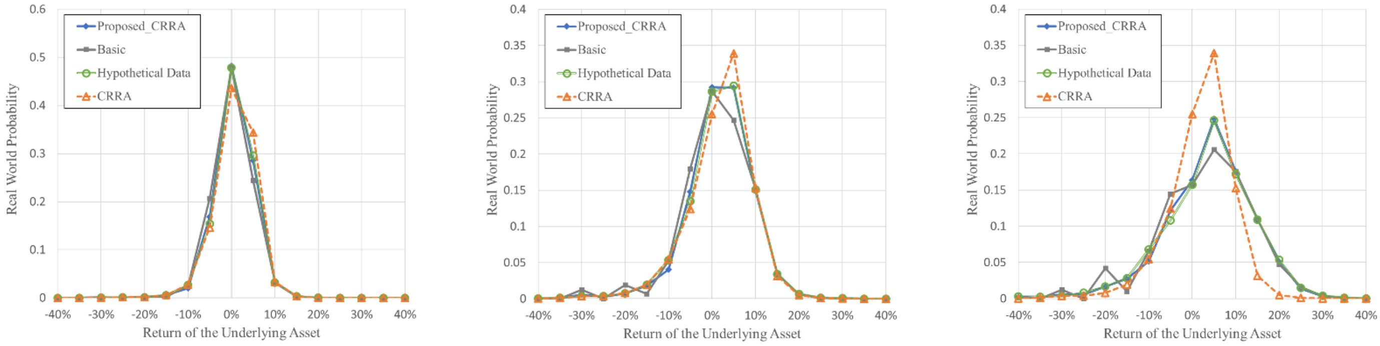


Figure 2: Hypothetical Pricing Kernel



(a) Real world distribution (a month) (b) Real world distribution (three months) (c) Real world distribution (six months)

Figure 3: Estimates of the distribution （$\sigma = 1\%, \zeta_1 = 10^{-2}, \zeta_2 = 10^{-4.4}$）

4.2.4 Perspective for Evaluation

We compare KL divergence $\bar{D}_{KL}(\boldsymbol{P}^E \parallel \boldsymbol{P}^H)$ of the following five types of distributions, and examine the effectiveness of the proposed method. "RND" is a risk neutral probability distribution calculated with Equation (1). "Basic" is the real world probability distribution estimated with the optimization problem shown by Jensen *et al.* (2016). "CRRA" is the real world distribution calculated with Equation (16) by using $\bar{\delta}, \bar{h}_s^{-1}$ estimated as shown in the section 3.3. "Proposed_CRRA" and "Proposed_RND" are the distributions calculated by the proposed method with the distribution "CRRA" and a risk neutral probability distribution "RND" as prior information, respectively. Setting the risk neutral distribution as prior information corresponds to $\bar{h}_s^{-1} = 1 (s = 1, \cdots, S)$ which is the same as the prior information used in Kiriu and Hibiki (2016).

We attempt to examine the results from the following perspective for evaluation in section 4.4.

(1) When the KL divergence of "Proposed_CRRA" or "Proposed_RND" is less than that of "RND", which equals one, the result means that the estimation accuracy of the distribution is improved by risk adjustment with the proposed method.

(2) When the KL divergence of "Proposed_CRRA" or "Proposed_RND" is less than that of "Basic", it indicates that the proposed method can estimate a real world distribution more accurately than the conventional method.

(3) When the KL divergence of "Proposed_CRRA" is less than that of "CRRA" (that of "Proposed_RND" is less than that of "RND"), it suggests that the distribution calculated by the proposed method is more accurate than the distribution given as prior information.

(4) When the KL divergence of "Proposed_CRRA" is less than that of "Proposed_RND", we can estimate the distribution more accurately by giving prior information using observed data.

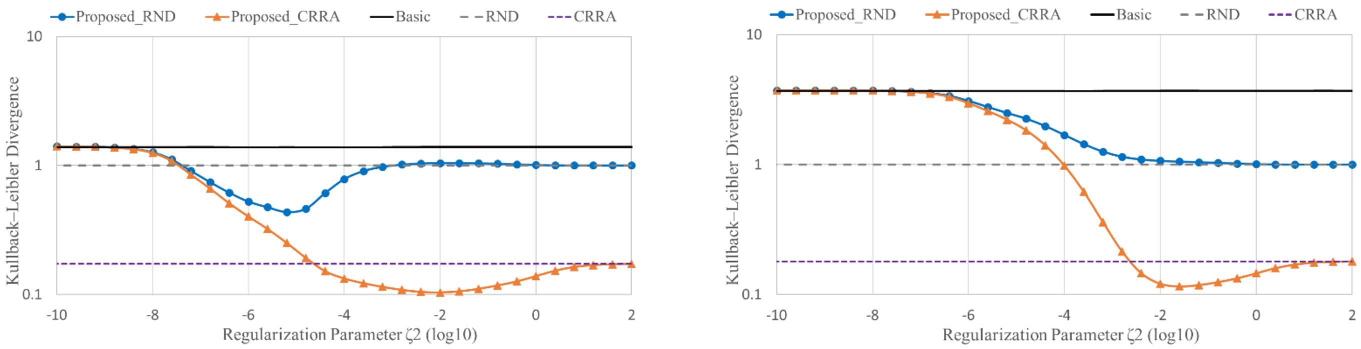
4.3 Examples of the Estimates

The examples of real world distributions are shown in Figure 3. Fig 3 (a), (b), and (c) show distributions in a month, three months, and six months. The distributions estimated by the conventional method are inappropriately distorted. It is because it is necessary to solve the ill-posed problem in the estimation process, and the estimated values are affected heavily by noise. On the other hand, the distributions estimated by the proposed method with "CRRA" as prior information are close to the hypothetical distributions. It is because it is expected that the regularization terms diminish the effect of noise.

4.4 Base Case

Figure 4 shows the KL divergence of five types of distributions as shown in section 4.2.4 for different values of ζ_2 where $\zeta_2 = 10^{-10}, 10^{-9.6}, \cdots, 10^2$. The smaller the KL divergence is, the higher the estimation accuracy is. The KL divergences of "Proposed_CRRA" and "Proposed_RND" approach asymptotically that of "Basic" as ζ_2 gets smaller because the solution estimated by the proposed method in the case of $\zeta_1, \zeta_2 = 0$ corresponds with the solution estimated by the conventional method. The KL divergence of the distribution estimated by the proposed method asymptotically approaches that of the distribution given as priori information as ζ_2 gets larger since a solution estimated by the proposed method in the case of $\zeta_1 \to \infty$ and $\zeta_2 \to \infty$ coincides with $\bar{\delta}$ and $\bar{h}_s^{-1}$ given as prior information. Therefore, as ζ_2 gets smaller, the KL divergence of "Proposed_CRRA" asymptotically approaches that of "CRRA". The KL divergence of "Proposed_RND" becomes close to that of " RND".

Firstly, we examine the results of the case of $\sigma = 1\%$. The KL divergence of "Basic" is larger than that of "RND", which shows that the estimation accuracy gets worse by



(a) $\sigma = 1\%$ (b) $\sigma = 5\%$

Figure 4: Relationship between KL divergences and regularization parameters

risk adjustment. It is because the solution is easily affected by noise due to the ill-posed problem. On the other hand, the KL divergences of "Proposed_CRRA" and "Proposed_RND" are less than that of "RND" when ζ_2 is larger than 10^{-7}, which indicates that risk adjustment by the proposed method can improve the estimation accuracy. The KL divergence of "Proposed_CRRA" is less than that of "CRRA". It suggests the proposed method can estimate the distribution more accurately than the distribution given as prior information. Furthermore, the KL divergence of "Proposed_CRRA" is less than that of "Proposed_RND" at any ζ_2. We show that the proposed method can estimate the distribution more accurately by giving prior information using observed data.

In addition, we investigate the results of the case of $\sigma = 5\%$. The KL divergence of "Proposed_RND" is not less than that of "RND", which shows that "Proposed_RND" cannot estimate an accurate distribution. On the other hand, "Proposed_CRRA" can estimate the accurate distribution by selecting an appropriate regularization parameter as in the case of $\sigma = 1\%$. However, the range of ζ_2 when noise is large becomes narrower than when noise is small in which the KL divergence of "Proposed_CRRA" is less than that of "CRRA". Thus, it is particularly important to select an appropriate regularization parameter for a large noise.

5. CONCLUSION

In our paper, we discuss the method to estimate a forward looking return distribution from option prices with GRT. It is necessary to solve an ill-posed problem in order to estimate a real world distribution with GRT. Then we propose a new estimation method to stabilize the solution under the given prior information with respect to the real world distribution, and the method of setting prior information using observed data. Furthermore, we verify the effectiveness of the proposed method through numerical experiments using hypothetical data.

We find the following two points. (1) There is possibility that a real world distribution recovered with our proposed method is more accurate than the distribution given as priori information and recovered with conventional estimation method. (2) Proposed method can improve the estimation accuracy by configuring prior information in accordance with the observed data.

Future issues are mainly as follows, (1) developing the selection criteria of appropriate parameters, and (2) estimating a real world distribution from option data with our proposed method and verifying the predictive power.

REFERENCES

Audrino, F., R. Huitema and M. Ludwig (2015), An empirical analysis of the Ross recovery theorem, *Available at SSRN 2433170.*

Breeden, D. T. and R. H. Litzenberger (1978), Prices of state-contingent claims implicit in option prices, *Journal of business*, 621-651.

Kiriu, T. and N. HIbiki (2016), Estimating forward looking distribution with the Ross recovery theorem, *Available at http://www.ae.keio.ac.jp/lab/soc/hibiki/profile_2/Kiriu&Hibiki_20161027.pdf.*

Jackwerth, J. C. (2000), Recovering risk aversion from option prices and realized returns, *The Review of Financial Studies*, 13(2), 433-451.

Jensen, C. S., D. Lando and L. H. Pedersen (2016), Generalized recovery. *Available at SSRN 2674541.*

Rosenberg, J. V. and R. F. Engle, (2002), Empirical pricing kernels. *Journal of Financial Economics*, 64(3), 341-372.

Ross, S. (2015), The recovery theorem, *The Journal of Finance*, 70, 615-648.