

Grażyna Trzpiot

Akademia Ekonomiczna w Katowicach

REGRESJA KWANTYLOWA A ESTYMACJA VAR*

1. Wstęp

Koncepcja VaR jawiła się jako jedna z najważniejszych miar zagrożenia ryzyka rynkowego. Opisuje górną granicę strat w tym sensie, że przekraczamy ten wyznaczony poziom ze znanym, z góry ustalonym prawdopodobieństwem α , zazwyczaj wybieranym pomiędzy 1 a 5%. Bardziej specyficznie, warunkowa informacja, którą posiadamy w czasie t , wyznacza wartość VaR w momencie $t + h$ dla jednej jednostki inwestycji i jest ujemnym kwantylem rzędu α , dla $0 < \alpha < 1$, warunkowego rozkładu stopy zwrotu inwestycji

$$VaR_{t+h}^{\alpha} = -Q_{\alpha}(r_{t+h} | F_t) = -\inf_x \{x \in R : P(r_{t+h} \leq x | F_t) \geq \alpha\}, \quad (1)$$

gdzie $Q_{\alpha}(\cdot)$ zapisuje kwantyl rzędu α , r_t jest stopą zwrotu portfela w czasie t , F_t reprezentuje informacje dostępne w czasie t .

Własności i tym samym możliwości aplikacyjne VaR jako miary ryzyka rynkowego były wielokrotnie dyskutowane. Wyznaczanie wartości VaR dla zmieniających się α , dla $0 < \alpha < 1$, jest równoważne sprawdzaniu relacji dominacji stochastycznej pierwszego rzędu [2], ale pomija założenia o awersji do ryzyka przy rangowaniu projektów inwestycyjnych. Wykorzystujemy jedną ustaloną wartość prawdopodobieństwa α , w przeciwieństwie do dominacji stochastycznych, i dowolne inwestycje podlegają rangowaniu. VaR jest miarą ryzyka dla tego ustalonego prawdopodobieństwa α . VaR nie jest miarą koherentną, nie spełnia aksjomatu o subaddytywności, co jest sprzeczne z dywersyfikacją portfela.

Do implementacji jednowymiarowych miar bazujących na definicji VaR wykorzystywano estymatory wartości kwantyli odpowiednio na lewym ogonie rozkładu

* Praca zrealizowana w ramach grantu nr N111 003 32/0262.

stopy zwrotu dla określonych ustalonych momentów w czasie. Do najbardziej znanych podejść można zaliczyć takie metody, jak: historyczna symulacja wykorzystująca rozkład empiryczny bazujący na (możliwie przefiltrowanych) przeszłych obserwacjach, metody Monte Carlo czy metoda wariancji kowariancji, oraz parametryczne modele wykorzystujące rozkłady stopy zwrotu zawierające oraz uwzględniające dynamikę zmienności oraz modele bazujące na teorii wartości ekstremalnych w rozkładzie ogona rozkładu stopy zwrotu. Kolejnym rozszerzeniem metodologii było włączenie metod statystyki wielowymiarowej (*principal components analysis*) oraz wykorzystanie metod nieparametrycznych. Celem artykułu jest przedstawienie modelu estymacji VaR bazującego na kwantylowej regresji modelującej wartość ustalonego kwantyla w miejsce całego rozkładu.

2. Regresja kwantylowa

Rozpatrujemy problem estymacji wektora nieznanych parametrów (regresji) $\mathbf{b}$ dla próby niezależnych obserwacji na ciągu zmiennych losowych $Y_1, Y_2, \dots, Y_T$ zgodnych z rozkładem

$$P(Y_t < y) = F(y - x_t \mathbf{b}), \quad t = 1, \dots, T, \quad (2)$$

gdzie $\{x_t, t = 1, \dots, T\}$ jest wierszem w znanej macierzy obserwacji (o wymiarach $T \times K$) oraz rozkład F nie jest znany.

Ciąg wartości $\{y_t, t = 1, \dots, T\}$ to obserwacje w próbie losowej pobranej z populacji o rozkładzie $\mathbf{Y}$ mającym dystrybuantę F . Wówczas kwantyl rzędu α , dla $0 < \alpha < 1$, może być wyznaczony jako rozwiązanie zadania:

$$\min_{\beta \in \mathbb{R}} \left\{ \sum_{t \in \{t: y_t \geq \beta\}} \alpha |y_t - \beta| + \sum_{t \in \{t: y_t < \beta\}} (1 - \alpha) |y_t - \beta| \right\}. \quad (3)$$

Wiemy, że obserwacje odstające są trudne do wykrycia w przypadku klasycznej regresji, więc niejednoznaczność prowadzi do rozszerzenia klasycznego modelu na model kwantylowy [5].

Zapiszemy jako $\{x_t, t = 1, \dots, T\}$ ciąg K wektorów (wierszy) macierzy obserwacji, zakładamy, że $\{y_t, t = 1, \dots, T\}$ jest losową próbą procesu regresji $u_t = y_t - x_t \mathbf{b}$ mającym dystrybuantę F . Wówczas kwantyl regresji rzędu α , dla $0 < \alpha < 1$ jest zdefiniowany jako rozwiązanie zadania:

$$\min_{\beta \in \mathbb{R}} \left\{ \sum_{t \in \{t: y_t \geq x_t \beta\}} \alpha |y_t - x_t \beta| + \sum_{t \in \{t: y_t < x_t \beta\}} (1 - \alpha) |y_t - x_t \beta| \right\}. \quad (4)$$

W przypadku $K = 1$ oraz $x_t = 1$ dla wszystkich t , model (4) redukuje się do zadania (3). Najmniejszy błąd bezwzględny jest wówczas równy medianie. Zadanie (4) ma zawsze rozwiązanie; w przypadku rozkładów ciągłych rozwiązanie jest jednoznaczne.

Zadanie minimalizacji związane z regresją kwantylową jest równoważne następującemu zadaniu programowania liniowego:

$$\min\{\alpha \mathbf{1}' \mathbf{r}^+ + (1 - \alpha) \mathbf{1}' \mathbf{r}^-\} \quad (5)$$

przy ograniczeniach

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\mathbf{b} + \mathbf{r}^+ + \mathbf{r}^- \\ (b, \mathbf{r}^+, \mathbf{r}^-) &\in R^K \times R_+^{2T}, \end{aligned}$$

gdzie $\mathbf{1}$ jest wektorem jednostkowym T -wymiarowym.

Klasyczny model regresji liniowej $E(y_i | x_i) = b_0 + x_i b_1$ zastąpiony został regresją kwantylową, zmienną objaśnianą jest kwantyl rzędu α : $Q_\alpha(y_i | x_i) = b_0(\alpha) + x_i b_1(\alpha)$. Jeżeli $b(\alpha)$ nie zależy od α , to model kwantylowy redukuje się do modelu warunkowej średniej ze stałą wariancją błędu dopasowania. Jeżeli $b(\alpha)$ zależy od α , to model specyfikuje wariancję, taką że kwantyl rozkładu y_i zależy od x_i . To podejście zawiera również heteroskedastyczne modele, w których wariancja zależy od zmiennych niezależnych. Model kwantylowy jednak odbiega od heteroskedastycznej specyfikacji.

3. Estymacja VaR – modele z regresją kwantylową

Szacowanie wartości VaR może być w sposób naturalny powiązane się z wykorzystaniem regresji kwantylowej. Aby estymować warunkowe wartości kwantyli, szereg czasowy konkretnego kwantyla należy modelować z wykorzystaniem wszystkich dostępnych informacji. Nie ma założeń co do typu rozkładu stopy zwrotu. Podstawowa idea tego podejścia to modelowanie warunkowego kwantyla rzędu α :

$$Q_\alpha(r_t | X_t) = -\text{VaR}_t \quad (6)$$

jako funkcji informacji zawartych w zmienności $X_t \in F_{t-1}$; r_t jest stopą zwrotu portfela w momencie t , F_t reprezentuje informacje dostępne w czasie t .

$$\text{VaR}_t = -g_\alpha(X_t; \beta_t), \quad (7)$$

gdzie $g_\alpha(\cdot; \cdot)$ – wektor parametrów β zależy od α .

Dobry przykład trafnej, dobrej informacji oraz postaci analitycznej funkcjonału powinien być bliski aproksymacji kwantyla w populacji. Koenker i Bassett [5] uogólnili regresję liniową poprzez przeniesienie uwagi z warunkowej średniej na warunkowe kwantyle. Jak pokazano w pracy Koenker i Portnoy [6], kwantyl z próby można wyznaczyć jako:

$$\min_{\beta \in R} \left\{ \sum_{r_i \geq \beta} \alpha |r_i - \beta| + \sum_{r_i < \beta} (1 - \alpha) |r_i - \beta| \right\}. \quad (8)$$

Rozwijając klasyczne zadanie regresji liniowej, Koenker i Bassett [5] zdefiniowali estymator regresji kwantylowej rzędu α jako:

$$\hat{\beta}(\alpha) = \min_{\beta \in R^k} \left\{ \sum_{r_i \geq x_i, \beta} \alpha |r_i - x_i \beta| + \sum_{r_i < x_i, \beta} (1 - \alpha) |r_i - x_i \beta| \right\}, \quad (9)$$

gdzie x_i jest nielosowym wektorem. Podstawowym założeniem modelu liniowej regresji kwantylowej przeniesionej do opisu rynku jest zależność $r_i = x_i' \beta_\alpha + \mu_{i,\alpha}$. Zauważmy, że rozkład błędu nie jest wyspecyfikowany. Jedynym założeniem modelu jest postać warunkowej funkcji rozkładu kwantyla zapisana następująco $Q_\alpha(r_i | x_i) = x_i' \beta_\alpha$ (czyli kwantyl stopy zwrotu portfela r_i jest liniową funkcją ekspozycji stylu) oraz $Q_\alpha(\mu_{i,\alpha} | x_i) = 0$.

Naturalnym rozszerzeniem funkcji celu do ogólnego możliwie też nieliniowego przypadku była propozycja zapisana w pracy Engle'a i Manganelli [4]:

$$\min_{\beta \in R^k} \left\{ \sum_{r_i \geq -VaR_i} \alpha |r_i + VaR_i| + \sum_{r_i < -VaR_i} (1 - \alpha) |r_i + VaR_i| \right\}. \quad (10)$$

W pracy tej przedstawiono również zgodność i asymptotyczną normalność nieliniowej kwantylowej regresji dla szeregów czasowych.

W równaniu $VaR_i = -g_\alpha(X_i; \beta_i)$ mamy związek z warunkowym odchyleniem standardowym. Model CAViaR (*conditional autoregressive VaR*) wykorzystuje wartość VaR_{i-1} jako zmienną objaśniającą x_i , aby adaptacyjnie objąć zależność od przeszłej zmienności i średniej. Funkcja r_{i-1} zawiera odniesienie do warunkowego kwantyla stopy zwrotu. W bazowym modelu proces adaptacyjny nie uczy się na bieżącej wartości stopy zwrotu, jak to jest w modelu o absolutnej symetrii CAViaR

$$VaR_i = \beta_0 + \beta_1 VaR_{i-1} + \beta_2 |r_{i-1}|. \quad (11)$$

Taki zapis pozwala na to, aby współczynnik β_1 był różny od jedności oraz wprowadza wprost odpowiedź na wartość kwantyla w procesie stopy zwrotu, traktując w sposób symetryczny efekt wartości ekstremalnych oraz zmienność.

Założenie symetryczności wpływów jest osłabione w modelu zakładającym asymetryczne współczynniki nachylenia

$$VaR_i = \beta_0 + \beta_1 VaR_{i-1} + \beta_2 \max[r_{i-1}, 0] + \beta_3 \max[-r_{i-1}, 0], \quad (12)$$

który pozwala na predykcję, wykorzystując odpowiednio dodatnie lub ujemne stopy zwrotu, jak również obejmuje efekt dźwigni.

Pośredni model GARCH(1,1) dla procesu CAViaR zaproponowany w pracy Engle'a i Manganelli [4]:

$$VaR_i = (\beta_0 + \beta_1 VaR_{i-1}^2 + \beta_2 r_{i-1}^2)^{1/2} \quad (13)$$

będzie zadowalający dla stopy zwrotu określonej jako

$$r_i = \mu_i + \varepsilon_i = \mu_i + \sigma_i z_i \quad (14)$$

GARCH(1,1) ze średnią zero i warunkową wariancją.

4. Testowanie dynamicznej postaci kwantyla

Zdefiniujemy ciąg prognoz wartości VaR: $H_t = I(r_t < -\text{VaR}_t)$ efektywnych względem F_{t-1} , jeżeli

$$E(H_t | F_{t-1}) = \alpha, \quad (15)$$

odnoszący się do iteracyjnych wartości oczekiwanych oraz implikujący, że H_t jest nieskorelowany z dowolną funkcją w zbiorze dostępnych informacji w czasie $t - 1$. Jeżeli zachodzi równanie (15), to zmienność VaR pojawia się zarówno z warunkowymi, jak i bezwarunkowymi prawdopodobieństwami. Wówczas prognozy zarówno wartości VaR, jak i H_t nie mogą być poprawione.

Równanie (15) jest silniejsze niż warunkowa zbieżność i sugeruje, że dowolne $x_{t-1} \in F_{t-1}$ jest nieskorelowane z H_t . Ta warunkowa zmienność wartości VaR jest najistotniejsza. Jako $\{\text{VaR}_t\}_{t=1}^T$ zapiszemy ciąg zmiennych losowych niezależnych i o takim samym rozkładzie takich, że

$$\text{VaR}_t = \begin{cases} P, & \text{dla prawdopodobieństwa } 1 - \alpha \\ -P, & \text{dla prawdopodobieństwa } \alpha \end{cases} \quad (16)$$

Dla P dostatecznie dużego i zależnego warunkowo od VaR ciąg zmienności wskazuje na dobrą warunkową zbieżność dla VaR. W celu zoperacjonalizowania równania (15) można zapisać zależność H_t od zmiennych w F_{t-1} np. jako:

$$H_t = \lambda_0 + \sum_{i=1}^p \beta_i H_{t-i} + \beta_{p+1} \text{VaR}_t + \mu_t, \quad (17)$$

gdzie przy założeniu hipotezy zerowej H_0 : $\lambda_0 = \lambda$ oraz $\beta_i = 0$, dla $i = 1, \dots, p+1$. W notacji wektorowej mamy $H - \lambda \mathbf{1} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$

$$\text{dla} \quad \mathbf{u}_t = \begin{cases} -\alpha, & \text{dla prawdopodobieństwa } 1 - \alpha \\ 1 - P, & \text{dla prawdopodobieństwa } \alpha \end{cases},$$

gdzie $\beta_0 = \lambda_0 - \lambda$ oraz $\mathbf{1}$ jest wektorem jednostkowym. Przy założeniu prawdziwości hipotezy zerowej danej równaniem (15) regresor nie ma siły wyjaśniającej, czyli H_0 : $\boldsymbol{\beta} = 0$. Ponieważ zmienne objaśniające nie są skorelowane ze zmienną objaśnianą, przy założeniu prawdziwości hipotezy zerowej oraz z centralnego twierdzenia granicznego wynika, że estymator $\hat{\boldsymbol{\beta}}_{LS} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{H} - \lambda \mathbf{1})$ ma rozkład $N(0, (\mathbf{X}'\mathbf{X})^{-1}(\lambda(1 - \lambda)))$, zatem można zapisać test dynamiki kwantyla jako:

$$DQ = \frac{\hat{\boldsymbol{\beta}}_{LS}' \mathbf{X}' \mathbf{X} \hat{\boldsymbol{\beta}}_{LS}}{\lambda(1 - \lambda)} \quad (18)$$

i jest to statystyka o rozkładzie χ^2_{p+2} .

5. Podsumowanie

W pracy omówiono modele estymacji VaR wykorzystujące regresję kwantylową. Nie wyjaśniono w niej następujących wcześniejszych podejść: podejścia regresji kwantylowej wykorzystywanego przez Chernozhukova i Umantseva [3] do modelowania wartości VaR bez weryfikacji zachowania modelu w terminach zmienności, estymacji wielookresowej wartości VaR podejmowanej przez Taylora [8] w kontekście zmian stopy kursów walutowych, który zapisał równanie (7) w postaci funkcji liniowej zależnej od zmienności estymatorów oraz horyzontu inwestycji. Efektywność tego modelu, jak to często ma miejsce w literaturze, jest weryfikowana bezwarunkowo. W artykule przedstawiono podstawowe zadanie regresji kwantylowej oraz model CAViaR (*conditional autoregressive VaR*) wraz z testem dynamicznej postaci kwantyla.

Literatura

- [1] Artzner P., Delbaen F., Eber J.-M., Heath D., *Coherent Measures of Risk*, „Mathematical Finance” 1999, 9, s. 203-228.
- [2] Bawa V.S., *Safety-First, Stochastic Dominance, and Optimal Portfolio Choice*, „Journal of Financial and Quantitative Analysis” 1978, 13, s. 255-271.
- [3] Chernozhukoy V., Umantsev L., *Conditional Value-at-Risk: Aspects of Modeling and Estimation*, „Empirical Economics” 2001, 26, 271-292.
- [4] Engle R.F., Manganelli S., *CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*, „Journal of Business and Economic Statistics” 2004, 22, s. 367-381.
- [5] Koenker R., Bassett G., *Regression Quantiles*, „Econometrica” 1978, 46, s. 33-50.
- [6] Koenker R., Portnoy S., *Quantile Regression*, Working Paper 97-0100, University of Illinois at Urbana-Champaign 1997.
- [7] Koenker R., Zhao Q., *Conditional Quantile Estimation and Inference for ARCH Models*, „Econometric Theory” 1996, 12, s. 793-813.
- [8] Taylor J.W., *A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns*, „Journal of Derivatives” 1999, 7, s. 64-78.
- [9] Trzpiot G., *O nieparametrycznych metodach estymacji VaR i ETL*, [w:] *Modelowanie preferencji a ryzyko '05*, Prace Naukowe AE w Katowicach, AE, Katowice 2006, s. 229-238
- [10] Trzpiot G., Ganczarek A., *Risk on Polish Energy Market*, [w:] *Dynamic Econometric Models*, red. Z. Zieliński, vol. 6, Nicolaus Copernicus University 2004, s. 127-136.
- [11] Trzpiot G., Ganczarek A., *Value at Risk Using the Principal Components Analysis on the Polish Power Exchange*, [w:] *From Data and Information*

Analysis to Knowledge Engineering, red. M. Spiliopoulou, R. Kruse, Ch. Bogelt, A. Nurnberger, W. Gaul, Proc. 29th Annual GFKL Conference, University of Magdeburg, March 9-11 2005, Springer-Verlag, Heidelberg-Berlin 2006, s. 550-557.

- [12] Trzpiot G., Jaguś F., *Koncepcja PCVaR w analizie ryzyka inwestycji portfelowej na GPW w Warszawie*, [w:] *Modelowanie preferencji a ryzyko'05*, Prace Naukowe AE w Katowicach, AE, Katowice 2006, s. 239-250.

A QUANTILE REGRESSION APPROACH TO ESTIMATING VAR

Summary

This paper discusses a quantile regression as a model for estimating downside market risk measures – VaR. Firstly, the author presents a problem of quantile regression defined in statistics. Secondly, we made in a few steps applications of this approach to estimate VaR accordingly to different assumptions for time series of returns. The paper also describes a dynamic quantile statistical test.