CALCULATION OF FUNCTIONALS BASED ON LARGE DIMENSION MATRIXES IN MAXIMAL LIKELIHOOD PROBLEMS

Gennady Medvedev
Institute of Mathematics and Informatics
Jan Dlugosz University

The problem of the likelihood function calculation is examined at parameter estimation of the stochastic process describing change of interest rates in the financial market. Such problem arises, when it is supposed, that process is not usual diffusion process, but possesses continuous derivatives. In this case the increments of process become correlated, and for the likelihood function evaluation it is necessary to invert a matrix of the high order equal to sample size. As is known the calculation of reciprocal matrixes of the high order either is impossible or results in essential mistakes of calculation. In paper the way to avoid this difficulty is offered.

One of the most important parameters of the financial market is so-called the riskfree interest rate. Its changes in time are usually generated by stochastic process. More often as such process one chooses a diffusion process. However the diffusion process, which is possessing independent increments and not differentiated with probability unit, not always adequately represents the real market changes of the interest rate. In this connection in [1] for modeling of the interest rate process $r(t)$ it is offered to use the stochastic process having the first derivative $r'(t)$, which is a diffusion process:

$$dr'(t) + 2a\, r'(t)\, dt + b[r(t) - \theta]\, dt = \sigma\, dW(t). \quad (1)$$

As the real interest rates are observed in discrete time the difference version of this equation is more convenient for practical application. This can be written down in the form

$$r_{k+2} = (e_1 + e_2)\, r_{k+1} - e_1 e_2\, r_k + \xi_1(k) + \xi_2(k), \quad (2)$$

where $r_k \equiv r(kh) - \theta$, h – an unit of discrete time; θ – stationary mean of process; $e_1 \equiv \exp[\lambda_1 h]$, $e_2 \equiv \exp[\lambda_2 h]$; λ_1, λ_2 – the roots of the characteristic equation

$$\lambda^2 + 2a\lambda + b = 0; \quad (3)$$

$\{\xi_1(k)\}$ and $\{\xi_2(k)\}$ – the sequences of normally distributed mutually independent random variables with zero expectation and such, that $\mathrm{var}[\xi_1(k)] = \sigma^2\gamma$, $\mathrm{var}[\xi_2(k)] = \sigma^2\delta$, but $\mathrm{cov}[\xi_1(k+1), \xi_2(k)] = \sigma^2\varepsilon$; $k = 0, 1, 2, \ldots$ The sequence of random variables generated by model (2), is similar to known process ARMA (2,1), but differs from it that there is a correlation between $\xi_1(k+1)$ and $\xi_2(k)$. Let's notice that the values λ_1, λ_2, γ, δ, ε are expressed in the analytical form (though also rather bulky) through parameters of initial model a and b of process $r(t)$ [2]. We shall mention them here for convenience:

$$\gamma + \delta = \frac{1}{(\lambda_2 - \lambda_1)^2}\left(\frac{(e_1^2 - 1)(e_2^2 + 1)}{2\lambda_1} + \frac{(e_2^2 - 1)(e_1^2 + 1)}{2\lambda_2} + 2\frac{1 - e_1^2 e_2^2}{\lambda_2 + \lambda_1}\right),$$

$$\varepsilon = \frac{1}{(\lambda_2 - \lambda_1)^2}\left(e_1\frac{1 - e_2^2}{2\lambda_2} + e_2\frac{1 - e_1^2}{2\lambda_1} - \frac{(e_1 + e_2)(1 - e_1 e_2)}{\lambda_2 + \lambda_1}\right).$$

Let's consider a problem of parameter estimation of model (1) by sample of observations of interest rates $\{r(kh), k = -1, 0, 1, 2, \ldots, n\}$. For this purpose the relation (2) is more convenient to rewrite in the form

$$y_{k+2} = \sigma\eta_{k+2}, \quad \text{where } y_{k+2} \equiv r_{k+2} - (e_1 + e_2)\, r_{k+1} + e_1 e_2\, r_k, \quad \sigma\eta_{k+2} \equiv \xi_1(k) + \xi_2(k).$$

The vector of random variables ($\eta_2\ \eta_3\ \ldots\ \eta_n$) is normally distributed with zero expectation and a covariance matrix $\Sigma_n = (\Sigma_{ij})$ with elements

$$\Sigma_{ii} = \gamma + \delta;\quad \Sigma_{ij} = \varepsilon, \text{ if } |i-j| = 1;\quad \Sigma_{ij} = 0, \text{ if } |i-j| > 1;\quad 1 \le i, j \le n. \tag{4}$$

Therefore logarithmic function of likelihood will look like:

$$n \ln\sigma^2 + \ln\det\Sigma_n + Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n/\sigma^2, \tag{5}$$

where $Y_n^{\mathrm{T}} \equiv (y_1, y_2, \ldots, y_n)$. Minimization (5) on σ^2 gives an estimate $\hat{\sigma}^2 = Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n/n$ (up to parameters a and b).

Substitution of the estimate $\hat{\sigma}^2$ in logarithmic function of likelihood (5) results in a problem of minimization of expression

$$(\det\Sigma_n)^{1/n}(Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n) \tag{6}$$

by parameters a and b.

At great size of sample n direct calculation of expression (6) is inconvenient and implies enough large computing errors.

The purpose of this paper in specifying the recurrent way of calculation (6) which is not resulting the large errors, peculiar to usual procedures of calculation of determinants of the high order and inverse matrixes of the high order.

Let's enter designations: $D_n \equiv \det\Sigma_n$; $Q_n \equiv Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n$; $o_n^{\mathrm{T}} \equiv (0\ 0\ \ldots\ 0\ \varepsilon)$ – n-vector-row, $o_1 = \varepsilon$; $v_n \equiv o_n^{\mathrm{T}}\Sigma_n^{-1}o_n$; $q_n \equiv Y_n^{\mathrm{T}}\Sigma_n^{-1}o_n$; $\mu_{k+1} \equiv (\gamma + \delta - o_k^{\mathrm{T}}\Sigma_k^{-1}o_k)^{-1}$.

The proposition 1. If a matrix Σ_n it is composed of the elements determined by formulae (4) the recurrent relations take place

$$\begin{gathered} D_k = (\gamma+\delta)D_{k-1} - \varepsilon^2 D_{k-2},\quad k > 1,\ D_1 = \gamma + \delta,\ D_0 = 1; \\ Q_n = Q_{n-1} + \mu_{n-1}(y_n - q_{n-1})^2,\ Q_1 = y_1^2/(\gamma+\delta); \\ q_k = \varepsilon\mu_k(y_k - q_{k-1}),\ k > 1,\ q_1 = \varepsilon\mu_1 y_1; \\ \mu_k = D_{k-1}/D_k;\ \mu_k = (\gamma + \delta - v_{k-1})^{-1},\ k > 1,\ \mu_1 = (\gamma+\delta)^{-1}; \\ v_k = \varepsilon^2(\gamma+\delta - v_{k-1})^{-1},\ k > 1,\ v_1 = \varepsilon^2(\gamma+\delta)^{-1}. \end{gathered} \tag{7}$$

Proof. D_n is known the Jacobi determinant. The recurrent formula of its calculation is known [3]. For the proof of other recurrent formulae we shall take advantage of representation of a inverse matrix for matrix A_n set in the block form [4]:

$$A_n^{-1} = \begin{pmatrix} A_{n-k} & a_{nk} \\ a_{kn} & A_k \end{pmatrix}^{-1} =$$

$$= \begin{pmatrix} [A_{n-k} - a_{nk}A_k^{-1}a_{kn}]^{-1} & -A_{n-k}^{-1}a_{nk}[A_k - a_{kn}A_{n-k}^{-1}a_{nk}]^{-1} \\ -[A_k - a_{kn}A_{n-k}^{-1}a_{nk}]^{-1}a_{kn}A_{n-k}^{-1} & [A_k - a_{kn}A_{n-k}^{-1}a_{nk}]^{-1} \end{pmatrix}. \tag{8}$$

More convenient representation of blocks of this inverse matrix can be received, using the statement from [4]: if square matrixes B, A, P are nonsingular, a matrix, inverse to $B = A + XPY$, is representable as

$$B^{-1} = A^{-1} - A^{-1}X(P^{-1} + YA^{-1}X)^{-1}YA^{-1}.$$

Therefore in the presentation (8)

$$[A_{n-k} - a_{nk}A_k^{-1}a_{kn}]^{-1} = A_{n-k}^{-1} + A_{n-k}^{-1}a_{nk}(A_k - a_{kn}A_{n-k}^{-1}a_{nk})^{-1}a_{kn}A_{n-k}^{-1}.$$

For receiving of the recurrent form of the inversion of a matrix $\Sigma_n = A_n$ it is convenient to take $A_{n-k} = \Sigma_{n-1}$, $a_{nk} = o_{n-1}$, $a_{kn} = o_{n-1}^{\mathrm{T}}$, $A_k = \gamma + \delta$. Then

$$\Sigma_n^{-1} = \begin{pmatrix} \Sigma_{n-1}^{-1} + \mu_n \Sigma_{n-1}^{-1} o_{n-1} o_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} & -\mu_n \Sigma_{n-1}^{-1} o_{n-1} \\ -\mu_n o_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} & \mu_n \end{pmatrix},$$

where $\mu_n = (\gamma + \delta - o_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} o_{n-1})^{-1}$.

Further representing Y_n in the block form $Y_n^{\mathrm{T}} = (Y_{n-1}^{\mathrm{T}} \;\; y_n)$ gives

$$Y_n^{\mathrm{T}} \Sigma_n^{-1} Y_n = Y_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} Y_{n-1} + \mu_n (Y_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} o_{n-1})^2 - 2\mu_n y_n (Y_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} o_{n-1}) + \mu_n y_n^2.$$

Applying the accepted designations for Q_n and q_n, we have

$$Q_n = Q_{n-1} + \mu_{n-1}(y_n - q_{n-1})^2.$$

Use of the block form for Σ_n^{-1}, Y_n^{T} and e_n^{T} gives

$$Y_n^{\mathrm{T}} \Sigma_n^{-1} o_n = \varepsilon \mu_n y_n - \varepsilon \mu_n Y_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} o_{n-1}, \text{ i. e. } q_n = \varepsilon\mu_n(y_n - q_{n-1}).$$

Now we shall take advantage of the block form for Σ_n^{-1}, to calculate $o_n^{\mathrm{T}} \Sigma_n^{-1} o_n$. Let's notice, that the first $n - 1$ components of a vector o_n are zero, and last is equal ε. Therefore

$$o_n^{\mathrm{T}} \Sigma_n^{-1} o_n = \varepsilon^2 (\gamma + \delta - o_{n-1}^{\mathrm{T}} \Sigma_{n-1}^{-1} o_{n-1})^{-1},$$

so $v_n = \varepsilon^2 (\gamma + \delta - v_{n-1})^{-1}$, $v_1 = \varepsilon^2 (\gamma + \delta)^{-1}$. Furthermore $v_n = \varepsilon^2 \mu_n$.

Let's notice, that the recurrent relation for D_n can be written down as

$$\frac{D_k}{D_{k-1}} = \gamma + \delta - \varepsilon^2 \frac{D_{k-2}}{D_{k-1}}, \quad \text{or} \quad \frac{D_{k-1}}{D_k} = \frac{1}{\gamma + \delta - \varepsilon^2 \dfrac{D_{k-2}}{D_{k-1}}}, \qquad \frac{D_0}{D_1} = \frac{1}{\gamma + \delta},$$

and these relations just also determine μ_k, i.e. we have $\mu_k = D_{k-1}/D_k$, or $D_k = \left(\prod_{i=1}^{k} \mu_i \right)^{-1}$. It finishes the proof of the proposition.

Using recurrent formulae (7) it is possible to construct computing procedure of logarithmic function of likelihood (6) enough simply and conveniently.

Let's consider a problem of convergence of the received recurrent relations. The basic recurrent procedure is a calculation of value v_k as through it are expressed D_k and μ_k. For the analysis of convergence it is convenient to enter values $\omega_k \equiv v_k/(\gamma + \delta)$ and $\rho \equiv \varepsilon/(\gamma + \delta)$. Then for ω_k the simple recurrent relation is received

$$\omega_k = \rho^2/(1 - \omega_{k-1}), \quad \omega_0 = 0. \tag{9}$$

Its properties depend on value only one parameter ρ. Let's notice, that this parameter has the sense of a correlation coefficient and consequently on absolute value never exceeds units. Limiting value ω_k at $k \to \infty$ it is determined by the equation $\omega_\infty^2 - \omega_\infty + \rho^2 = 0$, having two roots. The greater root is a unstable limit, therefore recurrent sequence $\{\omega_k\}$ converges to a limit

$$\omega_\infty = \left(1 - \sqrt{1 - 4\rho^2}\right)/2.$$

At $\rho \le 0{,}5$, that is usually carried out in practice, monotonous convergence takes place $\omega_k \to \omega_\infty$, and ω_k increases from zero up to ω_∞. To characterize process of convergence in this case it is convenient by examining the ratio ω_k/ω_∞, which determines a degree of convergence for everyone k. In tab. 1 the values ω_k/ω_∞ are submitted for various $\rho \le 0{,}5$. Apparently from this table for 9 iterations convergence on 90% for all $\rho \le 0{,}5$ is guaranteed, and for $\rho \le 0{,}4$

convergence up to 99 % is reached in all for 4 iterations. The slowest convergence is observed for critical value $\rho = 0{,}5$. In this case convergence on 95 % is reached for 19 iterations, and on 99 % – for 94 iterations.

Table 1

Values of a parameter of convergence ω_k/ω_∞ (in percentage) for various ρ

k	$\rho = 0{,}1$	$\rho = 0{,}2$	$\rho = 0{,}3$	$\rho = 0{,}4$	$\rho = 0{,}49$	$\rho = 0{,}5$
1	98,99%	95,83%	90,00%	80,00%	59,95%	50,03%
2	99,99%	99,82%	98,90%	95,24%	78,89%	66,71%
3	100,00%	99,99%	99,88%	98,82%	87,64%	75,05%
4	100,00%	100,00%	99,99%	99,71%	92,37%	80,05%
5	100,00%	100,00%	100,00%	99,93%	95,15%	83,39%

The analysis of explicit expression for $\rho \equiv \varepsilon/(\gamma + \delta)$ as functions of parameters a, b and h shows, that $\rho \in (0, 1/4)$, and the maximal value $\rho = 1/4$ is accepted in a limiting case, when $h \to 0$, $a \to 0$, $b \to 0$. Then from tab. 1 follows, that in a problem examined by us the recurrent procedure (9) converges for 4 iterations.

As it was noted earlier [5], the equation (1) can be considered as expanded the Vasiček model in the sense that model (1) and the Vasiček model will generate processes with the constant (not dependent on r) volatility σ and a variance $\sigma^2/2k = \text{var}[r(t)]$. In the literature there are results of parameter estimation of the Vasiček model for real financial processes [6–9]. In tab. 2 these estimations and results of calculations on their basis of roots of the characteristic equations used above, and also values $\gamma + \delta$, ε, ρ are resulted. (Roots λ_1 and λ_2 of characteristic equations (3) can be complex. It turns out for models from [8–9]. In this case in the table real and imaginary parts of these roots are resulted with use of a designation $\lambda_1 = \alpha - i\beta$, $\lambda_2 = \alpha + i\beta$.)

Table 2

Numerical values of parameters of models described in papers [6–9]

Parameter	CKLS (1992)	Bali (1999)	Aït-Sahalia (1996)	Aït-Sahalia (1999)
θ	0,0866	0,0642	0,0891	0,0717
σ	0,020	0,0077	0,0467	0,0224
a	0,5	0,5	0,5	0,5
k	0,1779	0,0436	0,8584	0,2610
h	1/12	1/12	1/365	1/12
λ_1 or α	– 0,7685	–0,9543	–0,5	–0,5
λ_2 or β	–0,2315	–0,0457	0,7800	0,1049
Lg($\gamma + \delta$)	–3,3572	–3,4496	–7,8642	–3,4497
Lg(ε)	–4,3311	–4,0517	–8,4662	–4,0518
ρ	0,1062	0,24992	0,2500	0,24996

CKLS (1992): the annualized one-month U.S. Treasury bill yield from June 1964 to December 1989 (306 observations). Ait-Sahalia (1996): the 7-day Eurodollar deposit spot rate, daily from 1 Jun 1973 to 25 Feb 1995 (5505 observations). Bali (1999): annualized one-month U.S. Treasury bill yield from June 1964 to December 1996 (390 observations). Ait-

Sahalia (1999): the Federal Reserve System funds data monthly from January 1963 to December 1998.

As in real problems the recurrent procedure (9) practically converges for 4 iterations, in formulas for calculation of logarithmic function of likelihood it is possible to use limiting values of sizes D_k, μ_k, q_k and ω_k that are calculated recurrently when the size of sample n is great enough.

Let's designate

$$\mu \equiv \mu_\infty = \frac{\omega_\infty}{(\gamma+\delta)\rho^2} = \frac{1-\sqrt{1-4\rho^2}}{2(\gamma+\delta)\rho^2}, \quad \varepsilon\mu = \frac{1-\sqrt{1-4\rho^2}}{2\rho}.$$

Then

$$\det\Sigma_n = D_n = \left(\prod_{j=1}^{n}\mu_j\right)^{-1} \approx 1/\mu^n = \left(\frac{2\rho^2(\gamma+\delta)}{1-\sqrt{1-4\rho^2}}\right)^n = [(\gamma+\delta)(1+\sqrt{1+4\rho^2})/2]^n.$$

$$q_k = \sum_{j=1}^{k}(-1)^{k-j} y_j \prod_{i=j}^{k}(\varepsilon\mu_i) \approx \sum_{j=1}^{k}(-1)^{k-j} y_j (\varepsilon\mu)^{k-j+1} = \varepsilon\mu\sum_{j=0}^{k-1} y_{k-j}(-\varepsilon\mu)^j.$$

$$Q_n \equiv Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n = \sum_{k=1}^{n}\mu_k(y_k - q_{k-1})^2 \approx \mu\sum_{k=1}^{n}\left(\sum_{j=0}^{k-1} y_{k-j}(-\varepsilon\mu)^j\right)^2.$$

Thus we receive an approximation logarithmic function of likelihood as

$$(\det\Sigma_n)^{1/n}(Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n) \approx \sum_{k=1}^{n}\left(\sum_{j=0}^{k-1} y_{k-j}(-\varepsilon\mu)^j\right)^2, \tag{10}$$

which is essentially easier, than the formula (6). However expression (10) yet is not completely ready for calculations.

Let's notice, that by definition $r_k = r(kh) - \theta$. It means, that

$$y_{k+2} = r_{k+2} - (e_1 + e_2)\, r_{k+1} + e_1 e_2\, r_k =$$
$$= r((k+2)h) - (e_1 + e_2)\, r((k+1)h) + e_1 e_2\, r(kh) - \theta(1-e_1)(1-e_2).$$

Thus in observable values are $r(kh)$, $k = -1, 0, 1, 2, \ldots, n$. It means, that expression (10) implicitly includes one more unknown parameter of model θ, which needs to be estimated. It directly is not connected to parameters of model a and b and consequently can be estimated irrespective of them, however this estimation will depend on a matrix Σ_n, i.e. finally, from estimations of parameters a and b.

Let's receive an explicit dependence of expression (10) from θ also we minimize it on this parameter. We shall present vector Y_n as

$$Y_n \equiv R_n - \theta(1-e_1)(1-e_2)\mathbf{1}_n,$$

where $\mathbf{1}_n$ – vector composed of units, and $R_n^{\mathrm{T}} \equiv (R_{n1}, R_{n2}, \ldots, R_{nn})$ – vector not dependent on θ with components $R_{nk} \equiv r(kh) - (e_1 + e_2)\, r((k-1)h)$. Then it is possible to write down

$$Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n = (R_n - \theta(1-e_1)(1-e_2)\mathbf{1}_n)^{\mathrm{T}}\Sigma_n^{-1}(R_n - \theta(1-e_1)(1-e_2)\mathbf{1}_n).$$

Minimization of this expression on θ gives estimate

$$\hat{\theta} = \frac{1}{(1-e_1)(1-e_2)}\frac{\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}R_n}{\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n}.$$

This estimation is unbiased, $E[\hat{\theta}]=\theta$, and its variance is

$$\mathrm{var}[\hat{\theta}]=\frac{1}{\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n}\left(\frac{\sigma}{(1-e_1)(1-e_2)}\right)^2.$$

Submitting value $\hat{\theta}$ in expression for Y_n instead of θ, we receive new representation for Y_n, already independent from unknown parameter.

$$Y_n=\left(I-\frac{\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}}{\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n}\right)R_n, \tag{11}$$

where I – an unit matrix.

Using representation (11) in expression (6), we receive expression for logarithmic function of likelihood through a vector R_n, determined in observable sizes $r(kh)$, $k=-1, 0, 1, 2, \ldots$:

$$(\det\Sigma_n)^{1/n}(Y_n^{\mathrm{T}}\Sigma_n^{-1}Y_n)=(\det\Sigma_n)^{1/n}\left(R_n^{\mathrm{T}}\Sigma_n^{-1}R_n-\frac{(\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}R_n)^2}{\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n}\right). \tag{12}$$

For simplification of calculation of this expression again we shall take advantage of recurrent procedures.

The proposition 2. The following recurrent relations take place

$$\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n=\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}\mathbf{1}_{n-1}+\mu_n(1-\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}o_{n-1})^2,\quad \mathbf{1}_1^{\mathrm{T}}\Sigma_1^{-1}\mathbf{1}_1=\mu_1;$$

$$\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}o_n=\varepsilon\mu_n(1-\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}o_{n-1})=\sum_{m=1}^{n}(-1)^{n-m}\prod_{k=m}^{n}(\varepsilon\mu_k),\qquad \mathbf{1}_1^{\mathrm{T}}\Sigma_1^{-1}o_1=\rho;$$

$$\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}R_n=\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}R_{n-1}+\mu_n(1-\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}o_{n-1})(R_{nn}-o_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}R_{n-1}),$$

$$\mathbf{1}_1^{\mathrm{T}}\Sigma_1^{-1}R_1=\mu_1R_{n1};$$

$$o_n^{\mathrm{T}}\Sigma_n^{-1}R_n=\varepsilon\mu_n(R_{nn}-o_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}R_{n-1})=\sum_{m=1}^{n}(-1)^{n-m}R_{nm}\prod_{k=m}^{n}(\varepsilon\mu_k),$$

$$o_1^{\mathrm{T}}\Sigma_1^{-1}R_1=\rho R_{n1}.$$

The proof of these relations practically repeats proofs of similar relations of the proposition 1 and consequently here is not resulted.

Using the property of fast convergence of recurrent procedures for enough high n it is possible to find the approached values of the values determining a relation (12):

$$\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}o_n=\sum_{m=1}^{n}(-1)^{n-m}\prod_{k=m}^{n}(\varepsilon\mu_k)\approx\varepsilon\mu\sum_{m=1}^{n}(-\varepsilon\mu)^{n-m}\approx\frac{\varepsilon\mu}{1+\varepsilon\mu},$$

$$\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n\approx\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}\mathbf{1}_{n-1}+\mu\left(1-\frac{\varepsilon\mu}{1+\varepsilon\mu}\right)^2\approx\frac{n\mu}{(1+\varepsilon\mu)^2},$$

$$\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}R_n\approx\mathbf{1}_{n-1}^{\mathrm{T}}\Sigma_{n-1}^{-1}R_{n-1}+\frac{\mu}{1+\varepsilon\mu}\sum_{m=1}^{n}(-\varepsilon\mu)^{n-m}R_{nm}\approx\frac{\mu}{(1+\varepsilon\mu)^2}\sum_{m=1}^{n}R_{nm},$$

$$R_n^{\mathrm{T}}\Sigma_n^{-1}R_n\approx\frac{\mu}{1-(\varepsilon\mu)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}R_{ni}R_{nj}(-\varepsilon\mu)^{|j-i|}.$$

Now we use the received results in expression (12) for logarithmic function of likelihood that gives the following approximate formula convenient for calculation of minimizing function

$$(\det\Sigma_n)^{1/n}\left(R_n^{\mathrm{T}}\Sigma_n^{-1}R_n-\frac{(\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}R_n)^2}{\mathbf{1}_n^{\mathrm{T}}\Sigma_n^{-1}\mathbf{1}_n}\right)\approx$$

$$\approx\frac{1}{1-(\varepsilon\mu)^2}\sum_{i,j=1}^{n}R_{ni}R_{nj}(-\varepsilon\mu)^{|j-i|}-\frac{1}{(1+\varepsilon\mu)^2}\frac{1}{n}\left(\sum_{m=1}^{n}R_{nm}\right)^2. \qquad (14)$$

Let's remind that in the formula (14) components of vector R_n are determined by equality $R_{nk}\equiv r(kh)-(e_1+e_2)\,r((k-1)h)$, where $r(kh)$, $k=-1, 0, 1, 2, \ldots$, are market observations of the interest rates. The formula (14) is represented more simple for calculations rather than initial expression (6) for logarithmic function of likelihood.

REFERENCES

1. Medvedev G. A. The Processes with Dependent Increments as Mathematical Models of the Interest Rate Processes // Proc. of the 10-th Intern. AFIR Symposium, Tromsø, 2000, 483–506.
2. Медведев Г. А. Математические модели финансовых рисков. Часть 1. Риски из-за неопределенности процентных ставок. Минск: БГУ, 1999.
3. Мишина А. П., Проскуряков И. В. Высшая алгебра. М: Физматгиз, 1962.
4. Хорн Р., Джонсон. Ч. Матричный анализ. М.: «Мир», 1989.
5. Medvedev G. A. The Asset Pricing When the Interest Rates Are Differentiable Stochastic Processes // Proc. of the 11-th Annual Intern. AFIR Symposium. Vol. 2. Toronto, 2001. P. 517–536.
6. CKLS: Chan, K. C., Karolyi, G. A., Longstaff, F. A., and Sanders A. S. An Empirical Comparison of Alternative Rate Models // Journ. of Finance, 1992, **47**, 1209–1227.
7. Bali, T. G. An Empirical Comparison of Continuous Time Models of the Short Term Interest Rate // Journ. of Futures Markets, 1999, **19**, 777–797.
8. Aït-Sahalia, Y. Nonparametric Pricing of Interest Rate Derivative Securities // Econometrica, 1996, **64**, 527–560.
9. Aït-Sahalia, Y. Transition densities for interest rate and nonlinear diffusions // Journ. of Finance, 1999, **54**, 1361–1395.