A new method for fast computation of cumulative distribution functions by fractional FFT

Ken’ichi Tanaka

1 School of Systems Information Science, Future University Hakodate, 116-2 Kamedanakanocho, Hakodate, Hokkaido 041-8655, Japan
E-mail ketanaka@fun.ac.jp

Received February 21, 2013, Accepted June 12, 2013

Abstract
We consider computation of cumulative distribution functions from their corresponding characteristic functions. We may use some known formulas with singular integrals for the computation. It is, however, difficult to speed up the computation with such formulas, because the fast Fourier transform (FFT) cannot be applied directly to them. Based on existing works for pricing of financial derivatives, we propose a fast method for the computation with fractional FFT and obtain accurate results on the entire real line.

Keywords cumulative distribution function, characteristic function, fractional FFT

1. Introduction

In this paper, we propose a fast method for computing cumulative distribution functions from their corresponding characteristic functions. Such computation is often needed in fields where stochastic modeling is employed such as physics, statistics, and finance, etc. In the field of finance, pricing of financial derivatives is one of major purposes, for which stochastic processes $X_t$ are used to model asset prices related to the financial derivatives.

For the pricing problem, it is fundamental to compute the distribution functions $F(x) = P(X_t \leq x)$. For some popular stochastic processes $X_t$ often employed for the modeling, the characteristic functions $\phi$ of $X_t$ can be obtained analytically, whereas the closed forms of the corresponding distribution functions $F$ are not available [1]. Even in such cases, $F$ can be numerically computed from $\phi$ by an inversion formula with the Fourier transform. For such computation, we may use some fundamental formulas with singular integrals [2]. It is, however, difficult to speed up the computation, because the fast Fourier transform (FFT) cannot be applied directly to them due to the singularity of the integrals.

Carr and Madan [3] pioneered in fast computation for derivative pricing using the FFT. Their methods have been used and developed by many authors (see [1] and the references therein). In particular, Chourdakis [4] improved efficiency of the computation using fractional FFT, which allows arbitrary width of equispaced grids for discretization of the Fourier transform. Carr and Madan’s methods and Chourdakis’s improvement can be applied to the computation of $F$ from $\phi$. Recently, Nakajima [5] proposed a method for the computation based on their works. His method is effective, but produces error not negligible for $F(x)$ at large $|x|$. Then, we propose a method to overcome this problem.

The rest of this paper is organized as follows. In Section 2, mathematical formulation of our problem is presented. In Section 3, we review some existing works with numerical examples. Our method is proposed and numerical examples for it are shown in Sections 4 and 5, respectively. Section 6 concludes.

2. Mathematical preliminaries

In the rest of this paper, we consider a one dimensional random variable $X$, dropping the subscript $t$. Let $f$ be the probability density function of $X$, let $F$ be the cumulative distribution function of $X$ and let $\phi$ be the characteristic function of $X$. Note that the following fundamental relations among them:

$$F(x) = \int_{-\infty}^{x} f(t) \, dt, \quad (1)$$

$$\phi(u) = E[e^{iuX}] = \int_{-\infty}^{\infty} f(x) \, e^{ixu} \, dx. \quad (2)$$

Furthermore, when $\phi$ is given, $f$ can be obtained by

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(u) \, e^{-ixu} \, du, \quad (3)$$

i.e., the inverse Fourier transform of $\phi$.

Applying some standard numerical integration formula to (3) and (1), we can compute $f(x)$ and $F(x)$. More precisely, for $x = x_n$ with $n = -N, \ldots, N - 1$, we can use a formula for (3) such as

$$f(x_n) \approx \frac{1}{2\pi} \sum_{m=-M}^{M-1} w_m \phi(u_m) e^{-iux_n}, \quad (4)$$

where $\{ w_m \}_{m=-M}^{M-1}$ and $\{ u_m \}_{m=-M}^{M-1}$ are some finite sequences. Furthermore, we can use the values $f(x_n)$ ($n = -N, \ldots, N - 1$) to approximate $F(x_n)$ in (1) similarly to (4). Then, in a naive manner, computational cost of $F(x)$ for a fixed $x = x_n$ is $O(MN)$. 

- 57 -
In practice, however, it is desirable to obtain the values of $F(x_n) \ (n = -N, \ldots, N - 1)$ as fast as possible. A promising steps to achieve effective computation are:

1. deriving $F$ directly from $\phi$ with the inverse Fourier transform, and

2. computing the transform with the FFT.

The difficulty in the step 1 is that $F$ does not have the usual Fourier transform on $\mathbb{R}$ because $\lim_{x \to \infty} F(x) = 1$. This can be remedied in some ways. For example, there is a known fundamental formula [2]

$$F(x) = \frac{1}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\phi(u)}{iu} e^{-ixu} \, du. \quad (5)$$

In this paper, some simple deformations $F_\rho$ of $F$ are treated such that they have closed forms of the Fourier transform $\psi_\rho$ expressed by $\phi$. These are described in Sections 3 and 4. The step 2 is computation of the sums:

$$F_\rho(x_n) \approx \frac{1}{2\pi} \sum_{m=-M}^{M-1} w_m \psi_\rho(u_m)e^{-iux_n} \quad (6)$$

for $n = -N, \ldots, N - 1$, which requires equispaced grids \( \{u_m\}_{m=-M}^{M-1} \) and \( \{x_n\}_{n=-N}^{N-1} \) such that $M = N$ and

$$u_m x_n = \frac{\pi mn}{N} \quad (7)$$

in order to apply the FFT directly to (6). This requirement can be relaxed by fractional FFT, as described in Sections 3. Note that, if these two steps are realized, the computational cost of $F(x_n) \ (n = -N, \ldots, N - 1)$ in the case $M = N$ is $O(N \log N)$.

3. Existing works

In this section, we review some existing works related to computation of distribution functions $F$, and specify the motivation of our work.

3.1 Nakajima’s idea based on Carr and Madan’s method and fractional FFT

Carr and Madan [3] introduced effective methods for derivative pricing using the FFT. Derivative prices can be mathematically described as expectations of some functions of random variables describing some financial markets. The probability distributions of the random variables are determined by some stochastic models. Since some popular models can give closed forms of the characteristic functions of the random variables [1], Carr and Madan developed a simple analytic expression for the Fourier transform of the derivative price, and used the FFT to numerically compute it. Moreover, Chourdakis [4] used fractional FFT, which is described in Section 3.2, for the computation of the Fourier transform to relax the requirement of equispaced grids such as (7) in Carr and Madan’s method, and improved efficiency of the computation.

Recently, Nakajima [5] proposed a method to compute distribution functions based on Carr and Madan’s method and fractional FFT. According to Carr and Madan’s idea, Nakajima considered a function $F_\rho$ defined by $F_\rho(x) = e^{-\rho x} F(x)$ for some $\rho > 0$, which corresponds to the step 1 in Section 2. Then, if $F$ decays exponentially as $x \to -\infty$ and $\rho$ is not so large, the function $F_\rho$ has the Fourier transform

$$\psi_\rho(u) = \int_{-\infty}^{\infty} F_\rho(x)e^{iu x} \, dx \quad (8)$$

with explicit form

$$\psi_\rho(u) = \frac{\phi(u + \rho)}{\rho} \frac{\rho}{u} \quad (9)$$

Therefore it follows that

$$F(x) = \frac{e^{\rho x}}{2\pi} \int_{-\infty}^{\infty} \psi_\rho(u)e^{-iux} \, du \quad (10)$$

Then, application of the trapezoidal formula to the integral in (10) yields

$$F(nh) \approx \frac{e^{\rho h}}{\pi} G_{\psi, h, N}(nh) \quad (11)$$

where $h, \tilde{h} > 0$ and

$$G_{\psi, h, N}(x) = \sum_{n=-N}^{N-1} h \psi_\rho(mh)e^{-imhx} \quad (12)$$

These correspond to (6) in the step 2 in Section 2, where $w_m = h, u_m = mh$ and $x_n = nh$, i.e., $h$ and $\tilde{h}$ are the width of the grids $\{u_m\}$ and $\{x_n\}$, respectively. In order to use the FFT directly for the computation of (12), $h$ and $\tilde{h}$ need to satisfy

$$\tilde{h} = \frac{\pi}{N} \quad (13)$$

due to (7). Then, in order to obtain accurate approximations of $F(nh) \ (n = -N, \ldots, N - 1)$, we need to use small $h$, which requires large $\tilde{h}$. This means that accurate computation with the FFT is possible only for a sparse grid $\{nh\}$. According to Chourdakis [4], Nakajima relaxed the requirement (13) by applying fractional FFT to (12).

3.2 Fractional FFT

Fractional FFT is developed by Bailey and Swartztrauber [6] to enable computation of the discrete Fourier transforms such as (12) with arbitrary $h$ and $\tilde{h}$. The fractional FFT is derived by regarding the sum in (12) as a circular convolution. Noting

$$2mn = m^2 + n^2 - (m - n)^2, \quad \text{for } n = -N, \ldots, N - 1 \text{ we have}$$

$$G_{\psi, h, N}(nh) = \sum_{m=-N}^{N-1} h \psi_\rho(mh)e^{-\pi i [m^2 + n^2 - (m - n)^2] \alpha}$$

$$= e^{-\pi i \alpha^2} \sum_{m=-N}^{N-1} y_m z_{n-m} \quad (14)$$

where $\alpha = \tilde{h}/(2\pi), \ y_m = \psi_\rho(mh)e^{-\pi i m^2 \alpha}$ and $z_m = e^{\pi i m^2 \alpha}$. Since $z_{-m} = z_m$, it suffices to consider $\{z_m\}_{m=0}^{2N}$. Note that the sum in (14) is a convolution but not circular convolution. Noting $2mn$ is a known fundamental formula [2], responses to the step 1 in Section 2. Then, if

$$y_m = 0 \quad (-2N \leq m < -N, N \leq m < 2N), \quad (15)$$

and $~ f$ need to satisfy

$$\tilde{h} = \frac{\pi}{N} \quad (13)$$

due to (7). Then, in order to obtain accurate approximations of $F(nh) \ (n = -N, \ldots, N - 1)$, we need to use small $h$, which requires large $\tilde{h}$. This means that accurate computation with the FFT is possible only for a sparse grid $\{nh\}$. According to Chourdakis [4], Nakajima relaxed the requirement (13) by applying fractional FFT to (12).

3.2 Fractional FFT

Fractional FFT is developed by Bailey and Swartztrauber [6] to enable computation of the discrete Fourier transforms such as (12) with arbitrary $h$ and $\tilde{h}$. The fractional FFT is derived by regarding the sum in (12) as a circular convolution. Noting

$$2mn = m^2 + n^2 - (m - n)^2, \quad \text{for } n = -N, \ldots, N - 1 \text{ we have}$$

$$G_{\psi, h, N}(nh) = \sum_{m=-N}^{N-1} h \psi_\rho(mh)e^{-\pi i [m^2 + n^2 - (m - n)^2] \alpha}$$

$$= e^{-\pi i \alpha^2} \sum_{m=-N}^{N-1} y_m z_{n-m} \quad (14)$$

where $\alpha = \tilde{h}/(2\pi), \ y_m = \psi_\rho(mh)e^{-\pi i m^2 \alpha}$ and $z_m = e^{\pi i m^2 \alpha}$. Since $z_{-m} = z_m$, it suffices to consider $\{z_m\}_{m=0}^{2N}$. Note that the sum in (14) is a convolution but not circular, i.e., we cannot regard $\{z_m\}_{m=0}^{2N}$ as 2N-periodic such as $z_{m+2N} = z_m$. Then, we need to convert this sum into a form of a circular convolution. A way for this conversion is to extend $\{y_m\}_{m=-N}^{N-1}$ and $\{z_m\}_{m=0}^{2N}$ to length $4N$ as follows:

$$y_m = 0 \quad (-2N \leq m < -N, N \leq m < 2N), \quad (15)$$
\[ z_m = e^{\pi i (m - 4N)\alpha} \quad (2N < m < 4N), \quad (16) \]

and extend \( \{ z_m \}_{m=0}^{4N-1} \) to 4N-periodic one for \(-4N \leq m < 0\). Then, we have the circular convolution

\[ G_{\psi, h, N}(n\tilde{h}) = e^{-\pi i n^2 \alpha} \sum_{m=-2N}^{2N-1} y_m z_{n-m}. \quad (17) \]

Finally, we can use the ordinal FFT to compute (17) with computational cost \( O(N \log N) \).

### 3.3 Numerical examples

We show numerical examples for the method based on (11) and the fractional FFT. We treat the standard normal distribution \( N(0,1) \) and the gamma distribution \( Ga(2,1) \). Their distribution functions \( F_N \) and \( F_{Ga} \), and characteristic functions \( \phi_N \) and \( \phi_{Ga} \) are as follows:

\[
F_N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} \, dt, \quad \phi_N(u) = e^{-u^2/2},
\]

\[
F_{Ga}(x) = \begin{cases} 
0 & (x < 0), \\
1 - (1 + x)e^{-x} & (x \geq 0),
\end{cases}
\]

\[
\phi_{Ga}(u) = (1 - iu)^{-2}.
\]

As for the parameters in (11), we set

\[ N = 2048, \quad h = \frac{5}{N}, \quad \tilde{h} = \frac{10}{N}, \quad \rho = 0.9. \quad (18) \]

That is, the interval on which \( \phi_N \) and \( \phi_{Ga} \) are evaluated is \([-5,5]\), and the interval on which \( F_N \) and \( F_{Ga} \) are computed is \([-10,10]\). Programs for these computations are written in C with double-precision floating-point arithmetic. The results are shown by Figs. 1 and 2. These results show that the accuracy is relatively bad at large \( x_n = n\tilde{h} \), i.e., at the tails of the distributions.

We can give theoretical explanation for these results. The error of the approximation (11) is estimated as

\[ |F(n\tilde{h}) - \frac{e^{ho n\tilde{h}}}{2\pi} G_{\psi, h, N}(n\tilde{h})| = \frac{e^{ho n\tilde{h}}}{2\pi} E_{\psi, h, N}(n\tilde{h}) \quad (19) \]

is the error of the trapezoidal formula. We can consider two causes for the bad error as follows:

1. The error \( E_{\psi, h, N}(n\tilde{h}) \) in (20) may become large when \( |n\tilde{h}| \gg \pi/h \), because \( F_{\rho}(x) \) is not periodic whereas \( G_{\psi, h, N}(x) \) in (12) has period \( 2\pi/h \).

2. Even if \( E_{\psi, h, N}(n\tilde{h}) \) is sufficiently small for any \( n \), the increasing factor \( e^{ho n\tilde{h}} \) in (19) may make the total error worse for large \( n \).

Then, we propose a method to remove these causes in Section 4.

### 4. Proposed method

In this section, we propose a method which enables fast and accurate computation of distribution functions \( F \) on the entire real line \( \mathbb{R} \). In the following, we write down an algorithm and then show its validity.

### 4.1 Proposed method

Suppose that the characteristic function \( \phi \) is analytic on the region \( D_d \) defined by \( D_d = \{ z \in \mathbb{C} \mid \text{Im}[z] < d \} \) for some \( d > 0 \). Then, our method is described as follows:

**Step 1.** For given \( \phi \), define \( \zeta \) by

\[ \zeta(u) = \left( \begin{array}{c}
1 \left( \phi(u)/u \right) - \frac{\pi}{2\sinh \left( \frac{\pi u}{2} \right)} \\
\text{i} \phi'(0)
\end{array} \right) \quad (u \neq 0), \quad (u = 0). \quad (21) \]

**Step 2.** For \( \varepsilon \) with \( 0 < \varepsilon < \min\{d, 1\} \), choose \( \rho \) with \( 0 < \rho \leq \min\{d, 1\} - \varepsilon \) and define \( \gamma \) by

\[ \gamma(u) = \frac{\zeta(u - \rho \text{i}) + \zeta(u + \rho \text{i})}{2}. \quad (22) \]

**Step 3.** Consider approximation of the inverse Fourier transform of \( \gamma \) as

\[ \Gamma_{N, h}(n\tilde{h}) = \frac{1}{2\pi} \sum_{m=-N}^{N-1} h\gamma(mh)e^{-imn\tilde{h}}, \quad (23) \]

and apply the fractional FFT to computation of (23) for \( n = -N, \ldots, N - 1 \).

**Step 4.** As approximations of the values \( F(n\tilde{h}) (n = -N, \ldots, N - 1) \), compute

\[ \left( \cosh(\rho n\tilde{h}) \right)^{-1} \Gamma_{N, h}(n\tilde{h}) + \frac{1}{2} \left( \tanh(n\tilde{h}) + 1 \right). \]

The analyticity of \( \phi \) on \( D_d \) is necessary for the existence of the inverse Fourier transform of \( \gamma \). Note that the last step contains multiplication of the decay factor \( \left( \cosh(\rho n\tilde{h}) \right)^{-1} \) to \( \Gamma_{N, h}(n\tilde{h}) \), which removes the causes 1 and 2 of the bad errors at the tail of the distributions pointed out in Section 3.3. Validity of the method is shown in Section 4.2.
4.2 Validity of the proposed method

Validity of the proposed method follows from the following two propositions. For the distribution function $F$, set $\tilde{F}(x) = F(x) - (\tanh x + 1)/2$ and $\tilde{F}_\rho(x) = \cosh(\rho x)\tilde{F}(x)$. Proof of Proposition 2 is straightforward and then omitted.

**Proposition 1** The Fourier transform of the function $\tilde{F}(x)$ exists and is written in the form (21).

**Proposition 2** The Fourier transform of the function $\tilde{F}_\rho(x)$ exists and is written in the form (22).

**Proof of Proposition 1** Set $T(x) = (\tanh x + 1)/2$ and let $\tau$ denote the Fourier transform of $T'(x)$. Then it follows that $\tau(u) = \frac{\pi u}{2 \sinh \left( \frac{\pi u}{2} \right)}$. (24)

Let $H$ be the Heaviside function

$$H(x) = \begin{cases} 1 & (x \geq 0), \\ 0 & (x < 0), \end{cases}$$

and set $Z(x) = F(x) - H(x)$ and $W(x) = T(x) - H(x)$. Then, $Z$ and $W$ have the Fourier transforms $\eta$ and $\theta$, respectively, which are written in the forms

$$\eta(u) = \frac{1 - \phi(u)}{iu}, \quad \theta(u) = \frac{1 - \tau(u)}{iu}. \quad (25)$$

Therefore, the Fourier transform of $\tilde{F}(x) = F(x) - T(x) = Z(x) - W(x)$ coincides with

$$\eta(u) - \theta(u) = \frac{\phi(u) - \tau(u)}{u} \quad (26)$$

for $u \neq 0$. Thus we obtain (21).

(QED)

5. Numerical examples

Using the proposed method, we treat again the normal distribution $N(0,1)$ and the Gamma distribution $Ga(2, 1)$. We used the same parameters as (18). The results are shown by Figs. 3 and 4, respectively. From these results, we can expect that the proposed method enables accurate computation of distribution functions on the entire real line.

Theoretical support for these results is given in a similar manner to (19) and (20). The total error of the proposed method is

$$\left( \cosh(\rho n\tilde{h}) \right)^{-1} \left| \cosh(\rho n\tilde{h})\tilde{F}(n\tilde{h}) - F_{N_{X_{\lambda}}}(n\tilde{h}) \right|. \quad (27)$$

This is a product of the decay factor $\left( \cosh(\rho n\tilde{h}) \right)^{-1}$ and the error of the trapezoidal formula for the inverse Fourier transform. Therefore the total error is suppressed even if $|n\tilde{h}|$ is large. We emphasize that it is on the entire real line that the proposed method achieves enough accuracy, although it produces worse error for some $x = n\tilde{h}$ than the existing method based on (11).

6. Concluding remarks

Using the fractional FFT, we have proposed a fast method for computing distribution functions $F$ from their corresponding characteristic functions $\phi$ which are analytic on the region $D_\phi$. Our method enables accurate computation on the entire real line, which is supported by the numerical examples and the theoretical estimate.

As future works, we may consider optimal error control based on rigorous error analysis of our method, accelerating convergence of the method, or modification of the method for $\phi$ with weaker analyticity, and so on.

Acknowledgments

The author gives special thanks to Prof. Masaaki Sugihara for his variable comments about this work. This work is supported by JSPS KAKENHI Grant Number 24760064.

References