LASER STATISTICS THROUGH ATMOSPHERIC TURBULENCE: A SIMULATION APPROACH

by

Tze Kuan Yan

B.S., University of Colorado, 1989

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Electrical Engineering
1991

This thesis for the Master of Science

degree by

Tze Kuan Yan

has been approved for the

Department of

Electrical Engineering

by

[redacted]

Arun K. Majumdar

[redacted]

Joe E. Thomas

[redacted]

Marvin F. Anderson

5-8-91

Date

ABSTRACT

Yan, Tze Kuan (M.S. Electrical Engineering)

Laser Statistics Through Atmospheric Turbulence: A Simulation Approach

Thesis directed by Professor Arun K. Majumdar

This thesis is based on maximum entropy method of constructing an unknown probability density function (PDF) from the high-order moments. Maximum entropy method is very useful in data reduction, and solving system of equations problems. This paper will look at the classical solution of solving the moment problems, and compare it with the new algorithm developed here, which is used to calculate the Lagrange multipliers. This method allows us to know how actually the Lagrange multipliers are related to moments of distribution. This new algorithm also applies to the infinite domain, semi-infinite domain, and finite domain.

The form and content of this abstract are approved. I recommend its publication.

Signed [redacted]

Arun K. Majumdar

DEDICATION

I wish to acknowledge the following people who made this study possible. My mentor, Dr. Arun K. Majumdar, of the Faculty of the Graduate School, Department of Electrical Engineering, University of Colorado at Denver, who was the source of my inspiration for this study. Dr. James Baker-Jarvis of NIST, who helped me to discover the new algorithm. Last, but not least, I am very grateful to my wife, and friends who assisted in typing this paper.

CONTENTS

CHAPTER I

INTRODUCTION

Laser irradiance fluctuations in real atmospheric turbulance as well as laboratory simulated turbulance have been studied for a long time. Nevertheless, it still remains as a puzzle for us to understand how the statistics of irradiance fluctuations behave in turbulant medium. In order to understand the nature of irridiance fluctuations, we have to know the form of probability density function (PDF), because probability density function is used to described the statistics of a fluctuating phenomenon like scintillation caused by atmospheric turbulance [10]. There are many ways we can calculate the probability density function. One of these methods is called polynomial expansion. It is a very popular method because it has been examined and studied by a lot of mathmaticians and scientists for more than a century. In recent years, a method called maximum entropy has been under intensive study. For problems outside the traditional domain of thermodynamics, the conceptual foundations of maximum entropy also have been debated for some time. This method becomes more meaningful in practical applications because computers have became more powerful today than before. From a practical point of view, the maximum entropy method is more useful because it use moments to describe the probability distribution. It is very easy for us to obtain the moments of the distribution from the experimental data.

This thesis is to use the technique of maximum entropy to reconstruct the probability density function from the finite moments. It is assumed that we know the physical data of signal information where we can calculate the lower-order and higher-order moments. The method of maximum entropy is a result of Bayesian theorem. With $P(A_i/A)$ being the probability calculated by maximum entropy (MAXENT). We will exam $P(A_i/A)$ very carefully later. Use of maximum entropy allows us to obtain and update the probability distributions as more information is available.

In 1948, Shannon [1] introduced the idea of maximum entropy to describe the information content of a signal. This idea was used by Jaynes [2] to analyze statistical physics in 1957. Since then, there are many outstanding application which used the method of maximum entropy. Mead and Papanicolaon [3] used MAXENT to solve integral equations. James Baker-Jarvis [4] use it to solve differential equations with constraint of boundary condition problems. Arun K. Majumdar has studied the general inverse scattering problem in the context of maximum entropy [5].

The uncertainty of a probability distribution can be measured by the information entropy. By maximizing the entropy subject to available information, we can obtain the probability distribution function that is least consistent with the knowledge of the data. So, a probability distribution for the current state of a system consistent with prior information can be obtained by MAXENT [6].

There are some restrictions that must be satisfied in Shannon's derived theory of maximum entropy [1].

1. The uncertainty of a probability distribution must be positive.
2. It must be an increasing function of the number of the degree of freedom (1/N).
3. Independent sources of uncertainty must be additive.

There are many areas of research that require the estimation of a probability density function (PDF) from the moment of a distribution, such as mine detection problems, moment problems, optical propagation problems. It is crucial in estimating the probability of successful detection in the mine detection problems. The probability of detection is quantified by a PDF for the noise level. Some of the research in the past has assumed a Gaussian distribution function is immersed into a sinusoidal signal. This is not always true, because the Gaussian distribution function is not the only noise exiting. This is why we need some kind of methods to calculate the probability density function from the moments which are obtained from the available data. The maximum entropy techniques for the moment problem that we are developing in this paper are very useful in this kind of problems.

CHAPTER 2

LASER STATISTICS IN ATMOSPHERIC TURBULENCE

Definition of Atmospheric Turbulence

Most energy from the optical waves will redistribute in atmospheric turbulence through small scale intensity scintillations. The log-intensity variance, $\delta^2_{\ln I}$, is used to describe the logarithm of the intensity of the wave. The variance usually increase along with the turbulence's strength. The turbulence saturates when the log-intensity variance is equal to 2.4. We can calculate the intensity scintillations in the saturation region when measured through aperture as small as $\sqrt{\lambda L}$ in diameter, where λ is the optic wavelength, L is the path length of propagation. The optic wave length varies inversely as the variance of the log-intensity scintillations. Because atmosphere is non-stationary, random fluctuations in density and refractive index are usually caused by atmosphere turbulence. Also the random fluctuations change according to different place and time. Since the variation of pressure is relatively small, we define the turbulence mainly as the fluctuations of temperature. In fact, the fluctuations of temperature is the prime reason to cause the fluctuation of the refractive index. The refractive index is defined as,

$$\delta_n = -79\, P\, \delta_T \,/\, (\theta - 1)\, T^2 \qquad (2.1)$$

Where T is the temperature, P is the pressure , and θ is the ratio of specific heat which is equal to 1.47 for air.

Velocity fluctuation is also another factor. Which is from large scale phenomena to affect the equilibrium of fluctuations. We define Lo the large scale, size, as the outer scale of turbulence. Where Lo = $\frac{2\pi}{K_o}$, K_o is a wave number. Also Lo is the inner scale of turbulence, where lo = $\frac{2\pi}{K_m}$

K_m is also a wave number.

The formula of spectrum is the same for velocity, temperature, and refractive index. It is defined by Talarslri as below [10] .

$$\phi_n(k) = 0.033\, C_n^2\, K^{-\frac{11}{3}} \exp\left(-\frac{K^2}{K_m^2}\right) \qquad (2.2)$$

Where

K is the spatial wave number.

K_m = 5.92/1o.

For small scales the equation is a good estimation if K >= K_m.

For large scales the equation is a poor estimation if K <= K_o.

Lo, lo, and C_n^2 are the three main parameters to describe the atmospheric turbulence, C_n^2 is defined as the structure parameter of refractive index. The intensity of the refractive index can be measured by C_n^2. Where l0^-17 <= C_n <= 10^-13, the unit is meters^(-2/3). The outer scale (large size) of atmospheric turbulence is Lo, above which the energy is to be introduced into the turbulence. It varies with the strength of the turbulence. That is 100 meters above the ground. The inner scale (small size) of atmosphere turbulence is lo, below which viscous dissipates and penetrates the spectrum. Turbulence energy is usually dissipated around 1mm above the ground.

Variance of Intensity Fluctuations

The variance of the log-intensity, $\delta_{\ln I}$, is usually measured from the experiment. So we can calculate the log amplitude

$$\delta_x = \delta_{\ln I} / 4 \qquad (2.3)$$

$\delta_{\ln I}^2$ has to be small for the spherical wave and the beam wave to be valid.

The value of $\delta_{\ln I}^2$ can be larger than 2.5 if the high value of C_n^2. is obtained in the 500 meters horizontal propagation path which is close to the ground. It is very different for $\delta_{\ln I}^2$ to be larger than 2.5 in vertical propagation paths because C_n^2 is constantly changing along the propagation path. From the experimental results, we can show that the spherical wave has a smaller variance than other kinds of waves.

Let Wo represents the radius of an optical beam at the transmitter and K is the optical wave number, $2\pi/\lambda$. If L << K Wo^2/2 , then $\delta^2_{\ln I}$ starts out equal to the variance of a plane wave. On the other hand, if L >> KWo^2/2 the variance will be approximately equal to the value of a spherical wave [10].

Relations between the PDF and the High Order Moments

To study phenomena like laser propagation through turbulence atmosphere, it is necessary to know the probability distribution function. However, it is difficult to measure PDF because of many technical problems such as random spikes and saturation of detector. As a result, there are many models have been proposed to obtain the probability density function of laser scintillation. Such as K distribution, I-K distribution, log-normal, exponential Bessel, Furutsn distribution [9]. No matter which model is used to calculate the PDF, we need to know the higher order moments of the random variable. The characteristic function of the random variable contains a series of higher order moments, which can be used to calculate the probability distribution. The higher order moments contain information to reconstruct the tail of probability density function. The information concerning the width and the lack of symmetry of the distribution can be obtained from the even-and odd order moments, if we use skewness to represent the measure and asymmetric nature, or flatness to measure the weight in the tail. It is very easy to obtain this information from Gaussian type distribution, but it is not the case when dealing with non-Gaussian type distribution. In order to understand the physical nature, we need to know the moments beyond 4th order.

skewness:

$$\Gamma_3 = \frac{<x^3>}{<x^2>^{\frac{3}{2}}} \tag{2.4}$$

excess:

$$\Gamma_4 = \frac{<x^4>}{<x^2>^2} - 3 \tag{2.5}$$

superskewness:

$$\Gamma_5 = \frac{<x^5>}{<x^3><x^2>} - 10 \tag{2.6}$$

super excess:

$$\Gamma_6 = \frac{<x^6>}{<x^2>^3} - 15 \tag{2.7}$$

hyperskewness:

$$\Gamma_7 = \frac{<x^7>}{<x^3><x^2>^2} - 105 \tag{2.8}$$

hyperexcess:

$$\Gamma_8 = \frac{<x^8>}{<x^2>^4} - 105 \tag{2.9}$$

For a Gaussian probability distribution, the odd-order central moments disappear; so Γ_3, Γ_5 and Γ_7 does not exist. We can obtain $\Gamma_4 = \Gamma_6 = \Gamma_8 = 0$ according to the above definition. Remember,Gaussian distribution is described by the first two moments only [5].

In this paper, it is assumed that we are able to obtain all of the high order moments that we need. Therefore, we can use this information to reconstruct the probability density function.

CHAPTER 3

THE INFORMATION ENTROPY

Definition of Entropy

Let us use a very simple example to induce the result, which will help us to understand the formalism of entropy. Consider the problem of tossing two coins. Each having possible outcomes R_1, R_2; therefore, the total number of possible outcomes are $R = R_1 * R_2$, or in this particular case, $2*2 = 4$. Nevertheless, the information contained in the two coins should be additive:

$$I(R_1, R_2) = I(R_1) + I(R_2) = \ln R \qquad (3.1)$$

So, we know the information must satisfy a logarithmic type of function.

Example: Information Content of a Morse Code

Consider an example below:

— — · · — · — · · · · — — — · · · · — · — · — —

Assume

N_1 = number of dashes

N_2 = number of dots

therefore

$N = N_1 + N_2$

(total number of dashes and dots)

$$R = \frac{N!}{N_1! * N_2!} \quad (3.2)$$

$I = Ln R$

$$= [Ln(N!) - Ln(N_1 !) - Ln(N_2!)] \quad (3.3)$$

Stirling's approximation:

$$N! \approx \sqrt{2 \pi n} * n^n * e^n \quad (3.4)$$

$$Ln N! = Ln(\sqrt{2\pi}) + 0.5*LnN + n*LnN - n \quad (3.5)$$

also

$$Ln(\sqrt{2\pi}) + 0.5*LnN << n*LnN - n \quad (3.6)$$

$$Ln N! \approx N(LnN - 1) \quad (3.7)$$

therefore

$$I = N(LnN - 1) - N_1(Ln N_1 -1) - N_2(LnN_2 -1) \quad (3.8)$$

remember the normalized information is

$$i = \frac{I}{N} \quad (3.9)$$

$$= LnN -1 - (N_1/N)*LnN_1 + N_1/N -(N_2/N)*LnN_2 + N_2/N$$

$$= LnN - (N_1/N)*LnN_1 -(N_2/N)*LnN_2$$

$$= -(N_1/N)*LnN_1 - (N_2/N)*LnN_2 + (N_1/N)*LnN + (N_2/N)*LnN \quad (3.10)$$

As a result

$$i = -\frac{N_1}{N}Ln\frac{N_1}{N} - \frac{N_2}{N}Ln\frac{N_2}{N} \quad (3.11)$$

Letting $P_i = N_i/N$, we can generalize the information entropy formula

$$S = \sum_{i}^{k} P_i \, LnP_i \quad (3.12)$$

an estimate of P_i can be obtained if we maximize the entropy [7].

The Linear problem with one constraint

Given

$$S = \sum_{i}^{k} P_i \, LnP_i \quad (3.13)$$

where

$$\sum_{i}^{k} P_i = 1 \quad (3.14)$$

is the normalized constraint for k points.

If

$$S = -\sum_{i}^{K} P_i \, LnP_i - \lambda(\sum_{i}^{K} P_i - 1) \qquad (3.15)$$

taking the variation of s with respect to P_i

$$\frac{\delta S}{\delta P_i} = -\sum_{i}^{k} [\, LnP_i - 1 - \lambda \,] \qquad (3.16)$$

then set Equation (3.16) equal to zero.

$$Ln\, P_i + 1 - \lambda = 0 \qquad (3.17)$$

therefore

$$P_i = \exp(\lambda - 1) \qquad (3.18)$$

To find λ , we need to use the constraint condition.

$$\sum_{i} P_i = K * \exp(\lambda - 1) = 1 \qquad (3.19)$$

where

$$K = 1/\exp(\lambda - 1)$$

$$= 1/P_i \qquad (3.20)$$

or

$$P_i = 1/K \tag{3.21}$$

The Linear Problem with Many Constraints

The constraints from an experiment are usually given in term of expectation values.

$$< f_k > = \sum_i P_i \ f_{ki} \tag{3.22}$$

where

$$\sum_i^k P_i = 1 \tag{3.23}$$

therefore

$$S = -\sum_i [P_i LnP_i - \lambda_0(P_i - 1) - \sum_k \lambda_k(f_{ki}P_i - <f_k>)] \tag{3.24}$$

$$\frac{\delta S}{\delta P_i} = -\sum_i [Ln\, P_i - \lambda_0 - \sum_k \lambda_k f_{ki}] \tag{3.25}$$

then set equation (3-25) equal to zero, as a result

$$P_i = \exp(-\lambda_0 - \sum_k \lambda_k f_{ki}) \tag{3.26}$$

from normalization constraint condition, we obtain

$$\exp(-\lambda_0) = \frac{1}{\sum_i \exp(-\sum_k \lambda_k f_{ki})} \tag{3.27}$$

then the probability is

$$P_i = \frac{\exp(-\sum_k \lambda_k f_{ki})}{\sum_i \exp(-\sum_k \lambda_k f_{ki})} \tag{3.28}$$

CHAPTER 4

THE MOMENT PROBLEM

History of the Moment Problem

From the classical point of view, estimation of an associated with moment problem can be calculated from the moments of the function. Maximum entropy is one of the techniques used for these type of problems. In order to obtain the least biased estimate of a probability density function, we need to maximize the entropy of a probability distribution and also consider the constraint that is the prior information. This process can be done with the maximum entropy method by use of Lagrange multipliers. It is not an easy job to calculate the Lagrange multipliers because they always involves a system of simultaneous non-linear equations. Since non-linear equations are not unique, the calculation always give multiple solutions to the same systems of equations. Therefore, this kind of process requires us to understand the physical behavior of the nature of the system, and generate a method to guess the initial conditions. Such a step is very important because Lagrange multipliers are just a set of number from a non-linear mathematical equations and they have no meaning by themselves. As a result, we need to understand the relationship between the Largrange multipliers and the physical nature. In this paper, the main object is to find an algorithm that is used to calculate the Lagrange multipliers explicitly.

Let us define a positive square integral function, v(x), so the moments of the function v(x) are given:

$$< x^n > = \int_{-\infty}^{\infty} v(x)\, x^n \, dx \tag{4.1}$$

where n = 0,1,2,3,4, ...

Ideally we should have infinite numbers of moments, but in real life, only finite moments can be considered. As I mention above, such a system of equations may be unique; it will give multiple solutions to the same system of equations. Therefore, we have to introduce another constraint that is the function that will at least decay as fast as exp(-|x|) for a large x.

Maximum entropy is a very powerful alternative method to solve the moment problem, where we have successfully used in optics problems, data reduction application and physics problems.

In this paper, I am going to take a closer look at the classical method that is used to calculate moment problems by solving a system of simultaneous non-linear equations, a job that is not always easy. The new iterative technique that we are developing in this paper, which is easier to handle. I also compare the moment expansion technique and cumulant expansion technique.

Bayesian Theorem

Suppose that an experiment is repeated a large number, N times, resulting in both A and B. $A \cap B$ n_{11} times; A and not B, $A \cap \overline{B}$, n_{21} times, B and not A, $\overline{A} \cap B$, n_{12} times, and neither $\overline{A}$ or $\overline{B}$, n_{22} times. I present these results in table 1.1.

Table 4.1 Table for Event A and B

	A	$\overline{A}$
B	n_{11}	n_{12}
$\overline{B}$	n_{21}	n_{22}

Note that $N = n_{11} + n_{12} + n_{21} + n_{22}$.

Then it follows that :

$$P(A) = \frac{n_{11} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \qquad (4.2)$$

$$P(B) = \frac{n_{11} + n_{12}}{n_{11} + n_{12} + n_{21} + n_{22}} \qquad (4.3)$$

$$P(A/B) = \frac{n_{11}}{n_{11} + n_{12}} \qquad (4.4)$$

$$P(B/A) = \frac{n_{11}}{n_{11} + n_{12}} \quad (4.5)$$

$$P(B \cap A) = \frac{n_{11}}{n_{11} + n_{12} + n_{21} + n_{22}} \quad (4.6)$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{n_{11}}{n_{11} + n_{12}} \quad (4.7)$$

or

$$P(AB/C) = P(A/BC) * P(B/C) \quad (4.8)$$

also

$$P(A/B) + P(\bar{A}/B) = 1 \quad (4.9)$$

We can prove

$$P(A/BC)*P(B/C) = P(B/AC)*P(A/C) \quad (4.10)$$

as a result, we have Bayesian theorem:

$$P(A/BC) = \frac{P(B/AC)*P(B/C)}{P(A/C)} \quad (4.11)$$

Let us consider C as our prior information about A before we receive information B. We treat P(A/C) as our prior probability of A as long as we know C. If we also know the new information about B, we can use this new information to update the posterior probability of P(A/BC). We can see now how powerful the Baysian theorem is. It allows us to study and forecast future events from our limited information, and it also allows us to add more information, when it becomes available, in the learning process.

Sometimes, we might encounter multivariate sample space. The Bayesian rule can be rewritten in another form. Assume sample space S is a union of mutually exclusive subsets. That is, say $S = B_1 \cup B_2 \cup \ldots \cup B_k$, when $B_i \cup B_k = \phi$, where i and k are not equal. Then any subset A of S can be written as :

$$A = A \cap S$$

$$= A \cap (B_1 \cup B_2 \cup \ldots \cup B_k)$$

$$= (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k) \qquad (4.12)$$

so

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_k)$$

$$= P(B_1) * P(A/B_1) + P(B_2) * P(A/B_2) + \ldots$$

$$+ P(B_k) * P(B_k)$$

$$= \sum_{i=1}^{k} P(B_i) * P(A/B_i) \qquad (4.13)$$

a conditional probability of the form $P(B_i/A)$ can be evaluated as :

$$P(B_i / A) = \frac{P(A/B_i)}{P(A)}$$

$$= \frac{P(B_i) * P A/B_i)}{\sum_{i=1}^{k} P(B_i)*P(A/B_i)} \qquad (4.14)$$

Derivation of Maximum Entropy Methods

If we define the information entropy

$$S = - \int_{-\infty}^{\infty} v(x) \, Lnv(x) \, dx \qquad (4.15)$$

where S is a function of a continuously varying random variable x [1].

As in chapter 3, we use the variation method to maximize s, In order to obtain the most probable distribution function, the constraints can be carried into the entropy formalism by Lagrange multipliers.

$$\overline{S} = - \int_{-\infty}^{\infty} v(x) \, Lnv(x) \, dx - \text{normalized constraint} \qquad (4.16)$$

where

$$\text{normalized constraint} = (\lambda_0 - 1)\int_{-\infty}^{\infty} v(x)dx + \sum_{n=1}^{k} \lambda_n \int_{-\infty}^{\infty} v(x)\, x^n dx$$

k is an even integer, and λ_n are Lagrange multipliers.

If we take $\delta\overline{S}$ with respect to δV and set it equal to zero, then we have an expression for the probability density function.

$$V(x) = \exp[- \sum_{n=0}^{k} \lambda_n x^n] \tag{4.17}$$

let

$$\phi(k) = \int_{-\infty}^{\infty} [\exp(ikx)\, V(x)]\, dx$$

$$= \int_{-\infty}^{\infty} \exp(ikx) \exp[- \sum_{n=0}^{k} \lambda_n x^n]\, dx \tag{4.18}$$

Where we have assumed $\lambda_n > 0$.

If we integrate $\phi(k)$ by parts and use of boundary conditions at infinity, then we will have a differential equation for the Fourier transform of the probability density function.

$$\phi(k)(ik - \lambda_1) - \sum_{n=2}^{k} n\, \lambda_n i^{-(n-1)} \frac{\partial^{(n-1)}\phi(k)}{\partial k^{(n-1)}} = 0 \tag{4.19}$$

assume

$$\phi(k) = \sum_{n=0}^{\infty} \frac{(ik)^n <x^n>}{n!} \tag{4.20}$$

and if we substitute $\phi(k)$ back into Equation (4.19), then arrange all co-efficients according to their powers. As a result, we will have a system of linear equations.

$$\sum_{n=1}^{k} n <x^{j+n-2}> \lambda_n = (j-1) <x^{j-2}> \tag{4.21}$$

for j = 1,2 ,3 ,4 , ... , m. (m must be an even integer).

The reason why m must be an even integer will be explained later.

Put into a matrix form, we will have

$$\begin{vmatrix} <x^0> & <x^1> & <x^2> & <x^3> & \ldots \\ <x^1> & <x^2> & <x^3> & <x^4> & \ldots \\ <x^2> & <x^3> & <x^4> & <x^5> & \ldots \\ <x^3> & <x^4> & <x^5> & <x^6> & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{vmatrix} \begin{matrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \\ 4\lambda_4 \\ \ldots \end{matrix} = \begin{matrix} 0 \\ <x^0> \\ 2<x^1> \\ 3<x^2> \\ \ldots \end{matrix} \tag{4.22}$$

let us define the above equation as

$$\overline{\overline{x}}\, \vec{\beta} = \vec{d} \tag{4.23}$$

therefore

$$\bar{\bar{x}} = \begin{vmatrix} \langle x^0\rangle & \langle x^1\rangle & \langle x^2\rangle & \langle x^3\rangle & \dots \\ \langle x^1\rangle & \langle x^2\rangle & \langle x^3\rangle & \langle x^4\rangle & \dots \\ \langle x^2\rangle & \langle x^3\rangle & \langle x^4\rangle & \langle x^5\rangle & \dots \\ \langle x^3\rangle & \langle x^4\rangle & \langle x^5\rangle & \langle x^6\rangle & \dots \\ \dots & \dots & \dots & \dots & \dots \end{vmatrix} \tag{4.24}$$

$$\vec{\beta} = (\lambda_1\,,\ 2\lambda_2\,, 3\lambda_3\,,\ 4\lambda_4\,,\ \dots,\ k\lambda_k)^t \tag{4.25}$$

(t denotes transpose).

and

$$\vec{d} = (0\,, \langle x^0\rangle\,, 2\langle x^1\rangle\,, 3\langle x^2\rangle\,, \dots\,, (k-1)\langle x^{k-2}\rangle)^t \tag{4.26}$$

Ideally, we can solve the infinite numbers of linear equation of (4.22) to obtain Lagrange multipliers, λ_n. If we know all λ_n, then we can calculate λ_0 by using the following formula.

$$V(0) = \exp(-\lambda_0) = \frac{\langle x^0\rangle}{A} \tag{4.27}$$

where

$$A = \int_{-\infty}^{\infty} \exp[-\sum_{i=1}^{k} \lambda_i x^i]\, dx \tag{4.28}$$

so we have finally completed the solution of V(x)

$$V(x) = \exp[-\sum_{i=0}^{k} \lambda_i x^i]\, dx \tag{4.29}$$

here, we have assumed that $<x^0> = 1$.

Now, Let us look at another alternative. If we define $\phi(k)$ in term of cumulants instead of moments, use it to expand the Fourier transform of the probability distribution, we will have a different result.

Let

$$\phi(k) = \exp[\sum_{n=0}^{k} \frac{(ik)^n c_n}{n!} \tag{4.30}$$

if we substitute it into Equation (4.19), expand this equation, and then identifying the k power of the coefficients. We will have another form of linear system of equations.

$$\begin{vmatrix} 1 & c_1 & c_1^2 + c_2 & c_1^3 + 3c_1c_2 + c_3 & \dots \\ 0 & c_2 & 2c_1c_2 + c_3 & 3c_1^2c_2 + 3c_1c_3 + 3c_2^2 + c_4 & \dots \\ 0 & c_3 & 2c_1c_3 + 2c_2^2 + c_4 & 3c_1^2c_3 + \dots + 9c_2c_3 + c_5 & \dots \\ 0 & c_4 & 2c_1c_4 + 6c_2c_3 + c_5 & 3c_1^2c_4 + \dots + 9c_3^2 + c_6 & \dots \\ & \dots & \dots & \dots & \dots \end{vmatrix}$$

$$\begin{vmatrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \\ 4\lambda_4 \\ \ldots \end{vmatrix} = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \ldots \end{vmatrix} \tag{4.31}$$

If we define it as

$$\bar{\bar{C}}\,\bar{\beta} = \bar{a} \tag{4.32}$$

where $\bar{\bar{C}}$ is equal to the matrix below

$$\begin{vmatrix} 1 & c_1 & c_1^2 + c_2 & c_1^3 + 3c_1c_2 + c_3 & \ldots \\ 0 & c_2 & 2c_1c_2 + c_3 & 3c_1^2c_2 + 3c_1c_3 + 3c_2^2 + c_4 & \ldots \\ 0 & c_3 & 2c_1c_3 + 2c_2^2 + c_4 & 3c_1^2c_3 + \ldots. \quad + 9c_2c_3 + c_5 & \ldots \\ 0 & c_4 & 2c_1c_4 + 6c_2c_3 + c_5 & 3c_1^2c_4 + \ldots \quad + 9c_3^2 + c_6 & \ldots \\ & \ldots & \ldots \quad .. & \ldots & \ldots \end{vmatrix}$$

$$\vec{\beta} = (\lambda_1, 2\lambda_2, 3\lambda_3, 4\lambda_4, \ldots, k\lambda_k)^t \tag{4.33}$$

and

$$\bar{a} = (0, 1, 0, 0, \ldots)^t \tag{4.34}$$

Moments and Cumulants

Since we can either use moment method or cumulant method to expand the Fourier transform, we should know the relationship between them in order to understand how they affect each other.

From Equations (4.20) and (4.30)

$$\phi(k) = \sum_{n=0}^{k} \frac{(ik)^n <x^n>}{n!} = \exp\left[\sum_{n=0}^{k} \frac{(ik)^n c_n}{n!} \right] \quad (4.35)$$

$$\text{lhs} = \{ 1 + <x^1>t + \frac{<x^2>t^2}{2!} + \dots + \frac{<x^r>t^r}{r!} + \dots \} \quad (4.36)$$

where $t = (ik)^r$

$$\text{rhs} = \exp\{ 1 + c_1 t + \frac{c_2 t^2}{2!} + \dots + \frac{c_r t^r}{r!} + \dots \}$$

$$= \exp(\frac{c_1 t}{1!}) \exp(\frac{c_2 t^2}{2!}) \dots \exp(\frac{c_r t^r}{r!}) \dots$$

$$= \{1 + \frac{c_1 t}{1!} + \frac{c_1 t^2}{2!} + \dots \} \{1 + \frac{c_2 t^2}{2!} + \frac{1}{2!}(\frac{c_2^2 t^4}{2!^2}) + \dots \} \dots$$

$$\{1 + \frac{c_r t^r}{r!} + \frac{1}{2!}(\frac{c_r^2 t^{2r}}{r!^2}) + \dots \} \dots \quad (4.37)$$

It is worth noting that the very tedious process of writing down the explicit relation for particular values of n may be shortened considerably. In fact, differentiating Equation (4.35) with respect to c_j . we have

$$\frac{t^j}{j!}\left(1 + \langle x^1\rangle t + \frac{\langle x^2\rangle t^2}{2!} + \ldots + \frac{\langle x^r\rangle t^r}{r!} + \ldots\right)$$

$$= \frac{\partial\langle x^1\rangle t}{\partial c_j} + \ldots + \frac{1}{r!}\frac{\partial\langle x^r\rangle t^r}{\partial c_j} + \ldots \qquad (4.38)$$

hence identifying the power of t

$$\frac{\partial\langle x^r\rangle}{\partial c_j} = \binom{r}{j}\langle x^{r-j}\rangle \qquad (4.39)$$

In particular, j = 1

$$\frac{\partial\langle x^r\rangle}{\partial c_1} = r\langle x^{r-1}\rangle \qquad (4.40)$$

therefore, given any $\langle x^r\rangle$ in term of the c's, we can write down successively those of lower orders by a differentiation.

$$\langle x^1\rangle = c_1 \qquad (4.41)$$

$$\langle x^2\rangle = c_2 + c_1^2 \qquad (4.42)$$

$$\langle x^3\rangle = c_3 + 3c_2c_1 + c_1^3 \qquad (4.43)$$

$$\langle x^4 \rangle = c_4 + 4c_3c_1 + 3c_2^2 + 6c_2c_1^2 + c_1^4 \qquad (4.44)$$

$$\langle x^5 \rangle = c_5 + 5c_4c_1 + 10c_3c_2 + 10c_3c_1^2 + 15c_2^2c_1 + 10c_2c_1^3 + c_1^5 \qquad (4.45)$$

$$\langle x^6 \rangle = c_6 + 6c_5c_1 + 15c_4c_2 + 15c_4c_1^2 + 10c_3^2 + 60c_3c_2c_1 + 20c_3c_1^3$$
$$+ 20c_3c_1^3 + 15c_2^3 + 45c_2^2c_1^2 + 15c_2c_1^4 + c_1^6$$
$$(4.46)$$

The above moments are regular moments.

Conversely

$$\frac{c_1\, t^1}{1!} + \ldots + \frac{c_r\, t^r}{r!} + \ldots = \mathrm{Ln}\left(1 + \frac{\langle x^1 \rangle\, t^1}{1!} + \ldots + \frac{\langle x^r \rangle\, t^r}{r!} + \ldots\right) \qquad (4.47)$$

we will have the first 6 terms as the following

$$c_1 = \langle x^1 \rangle \qquad (4.48)$$

$$c_2 = \langle x^2 \rangle - \langle x^1 \rangle^2 \qquad (4.49)$$

$$c_3 = \langle x^3 \rangle - 3\langle x^2 \rangle \langle x^1 \rangle' + 2\langle x^1 \rangle^3 \qquad (4.50)$$

$$c_4 = \langle x^4 \rangle - 4\langle x^3 \rangle \langle x^1 \rangle - 3\langle x^2 \rangle^2 + 12\langle x^2 \rangle \langle x^1 \rangle^2 - 6\langle x^1 \rangle^4 \qquad (4.51)$$

$$c_5 = \langle x^5\rangle - 5\langle x^4\rangle \langle x^1\rangle - 10\langle x^3\rangle \langle x^2\rangle + 20\langle x^3\rangle \langle x^1\rangle^2 + 30\langle x^2\rangle^2\langle x^1\rangle - 60\langle x^2\rangle \langle x^1\rangle^3 + 24\langle x^1\rangle^5 \tag{4.52}$$

$$c_6 = \langle x^6\rangle - 6\langle x^5\rangle \langle x^1\rangle - 15\langle x^4\rangle \langle x^2\rangle + 30\langle x^4\rangle \langle x^1\rangle^2 - 10\langle x^3\rangle^2 + 120\langle x^3\rangle \langle x^2\rangle \langle x^1\rangle - 120\langle x^3\rangle\langle x^1\rangle^3 + 30\langle x^2\rangle^3 - 270\langle x^2\rangle \langle x^1\rangle^2 + 360\langle x^2\rangle \langle x^1\rangle^4 - 120\langle x^1\rangle^6 \tag{4.53}$$

Analysis of the MAXENT Algorithm

From Equation (4.22), we can see the systems of linear equation is infinite. It is impossible to solve such a system of linear equation. There are many ways to approximate the solution to this kind of system of equation. One of this ways is to truncate the system of linear equation into a even K system. (K must be an even number). The truncated system will include the higher order moments, which is not the case for Gaussian distribution. It only requires two moments to describe its characteristics. Therefore, maximum entropy becomes a problem of non-linear nature. In this case, we need to intelligently guess the higher the higher moments in order to solve the system for the Lagrange multipliers, except for the Gaussian distribution. We can obtain the Lagrange multipliers for Gaussian distribution by using two linear equations. Higher order moments do not require here because they do not influence the solution.

The algorithm I am developing here can improve the approximation to maximum entropy by partitioning the equations into two matrix systems

$$\overline{\overline{A}}\vec{\beta} = \vec{d_1} \tag{4.54}$$

$$\overline{\overline{B}}\vec{\beta} = \vec{d_2} \tag{4.55}$$

The second system has H equations. (H is an even number). Assume the $<x^{k+2}>$ is the highest moment that is going to be used, and all the higher moments after $<x^{k+2}>$ are set to zero. Arrange the equations completely, so we solve moments instead of Lagrange multipliers. B is an H by H matrix of Lagrange multipliers. And beta is an H by 1 matrix of moments. From the first equation, we can approximate the values of moments in A and use these values to solve for the Lagrange multipliers. Then, I insert the calculated Lagrange multipliers into the second system of equations to calculate the new moments. After that , I can substitute the new moments values from Equation (4.55) to update the a matrix A in Equation (4.54). The moments and the Lagrange multipliers will finally converge in a few iterations. The above method will only work based on two assumptions. The first assumption is to have an educated initial guess. The second assumption requires that the Lagrange multipliers are consistent with the integrability of the distribution.

There is no difference between the Equation (4.22) and the system of linear equations obtained from non-linear, least square approximation. Nevertheless, in the least square approach, the development of moment matrix is based on the assumption that the underlying probability distribution is constant. We also need a series expression of V(x) in terms of expression function. In MAXENT approach, no such assumptions have been made. In addition to this, the maximum entropy method is more flexible, and we feed additional constraint information into system of equations without any trouble.

When I used maximum entropy algorithm, I found out some interest points, if I used ten or more moments in calculating the Lagrange multipliers. The numerical solutions for Lagrange multiplier were not as good as I wanted them to be. They were very noisy. This might due to the ill-condition matrix, because the values of the elements in the first column and the last column are to much different. The values of the first column are too small, but the values of the last column are too big. The highest moment I ever used in my algorithm is sixth moment. It seemed to give me some good results. Finally, the numerical values of Lagrange multipliers calculated from moment method and cumulant method were very close to each other.

Application in Infinite Domain

Let us consider the classical problem. As we mentioned earlier, Gaussian distribution is one of the special functions that we only use two moments to describe it. So, the Lagrange multipliers can be calculated by two linear equations. Put it into the matrix form .

$$\begin{vmatrix} \langle x^0\rangle & \langle x^1\rangle \\ \langle x^1\rangle & \langle x^2\rangle \end{vmatrix} \begin{vmatrix} \lambda_1 \\ 2\lambda_2 \end{vmatrix} = \begin{vmatrix} 0 \\ \langle x^0\rangle \end{vmatrix} \tag{4.56}$$

using Cramer's rule, we obtain

$$\lambda_1 = \frac{\langle x^0\rangle \langle x^1\rangle}{\langle x^2\rangle \langle x^0\rangle - \langle x^1\rangle^2} \tag{4.57}$$

$$\lambda_2 = \frac{\langle x^0\rangle^2}{2\langle x^2\rangle \langle x^0\rangle - 2\langle x^1\rangle^2} \tag{4.58}$$

$$\exp(-\lambda_o) = \frac{\langle x^0 \rangle}{\int_{-\infty}^{\infty} \exp(-\sum_{n=1}^{k} \lambda_n x^n)\, dx} \tag{4.59}$$

Example 1: a distribution of Gaussian type.

Let

$$f(x) = \exp(-x^2) \tag{4.60}$$

the moments are

$$\langle x^n \rangle = \int_{-\infty}^{\infty} \exp(-x^2)\, x^n\, dx = \Gamma(\frac{n+1}{2}) \tag{4.61}$$

for n = 0, 2, 4, ... even numbers.

the moments are equal to zero for n = 1, 3, 5, ... odd numbers.

$$\langle x^0 \rangle = 1$$

$$\langle x^1 \rangle = \langle x^3 \rangle = \langle x^5 \rangle = 0$$

$$\langle x^2 \rangle = 0.5$$

$$\langle x^4 \rangle = 0.75$$

$$\langle x^6 \rangle = 1.875$$

therefore, the Lagrange multipliers are calculated from the bellowing equation

$$\begin{vmatrix} 1 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.75 \\ 0.5 & 0 & 0.75 & 0 \\ 0 & 0.75 & 0 & 1.875 \end{vmatrix} \begin{matrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \\ 4\lambda_4 \end{matrix} = \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \end{matrix} \qquad (4.62)$$

we get the solutions for this system.

$$\lambda_2 = 1$$

$$\lambda_1 = \lambda_3 = \lambda_4 = 0$$

in this case, the solution is exact.

In fact, if we set the higher moments equal to zero, we can still get the same result because a Gaussian distribution is completely determined by its first and second moments.

therefore

$$V(x) = \exp(-x^2) \tag{4.63}$$

This expression is plotted on Figure 4.1.

Application in Semi-Infinite Domain

We define the moments

$$<x^n> = \int_0^\infty v(x)\, x^n\, dx \tag{4.64}$$

and the Laplace transform of the probability density function of the Fourier transform

$$\phi(k) = \int_0^\infty v(x) \exp(ikx)\, dx = \sum_{n=0}^{\infty} \frac{(-k)^n <x^n>}{n!} \tag{4.65}$$

we can calculate the Lagrange multipliers from the following equation

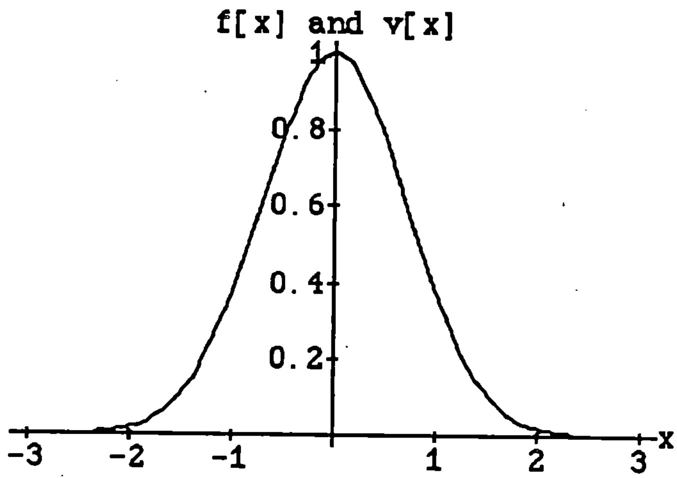

Figure 4.1

Comparison of

f[x] and v[x]

$$\begin{vmatrix} \langle x^0\rangle & \langle x^1\rangle & \langle x^2\rangle & \langle x^3\rangle & \dots \\ \langle x^1\rangle & \langle x^2\rangle & \langle x^3\rangle & \langle x^4\rangle & \dots \\ \langle x^2\rangle & \langle x^3\rangle & \langle x^4\rangle & \langle x^5\rangle & \dots \\ \langle x^3\rangle & \langle x^4\rangle & \langle x^5\rangle & \langle x^6\rangle & \dots \\ \dots & \dots & \dots & \dots & \dots \end{vmatrix} \begin{matrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \\ 4\lambda_4 \\ \dots \end{matrix} = \begin{matrix} v(0) \\ \langle x^0\rangle \\ 2\langle x^1\rangle \\ 3\langle x^2\rangle \\ \dots \end{matrix}$$

(4.66)

Example 1: a distribution function with initial conditions.

Let

$$V(x) = \exp(-x) \tag{4.67}$$

the moments are given as

$$\langle x^n \rangle = \int_0^\infty \exp(-x)\, x^n\, dx = n! \tag{4.68}$$

where

$\langle x^0\rangle = 1$

$\langle x^1\rangle = 1$

$\langle x^2\rangle = 2$

$\langle x^3\rangle = 6$

$\langle x^4\rangle = 24$

substitute these values into Equation (4.66)

$$\begin{vmatrix} 1 & 1 & 2 \\ 1 & 2 & 6 \\ 2 & 6 & 24 \end{vmatrix} \begin{matrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \end{matrix} = \begin{matrix} 1 \\ 1 \\ 2 \end{matrix}$$

(4.69)

we get

$\lambda_1 = 1$

$\lambda_2 = \lambda_3 = 0$

again the maximum entropy is exact.

$$V(x) = \exp(-x) \tag{4.70}$$

This function is plotted in Figure 4.2.

Now let us look at another application of Equation (4.66) which yields an inexact solution.

Example 2:

Let

$$V(x) = x \exp(-x) \tag{4.71}$$

the moments are given as

$$\langle x^n \rangle = \int_0^{\infty} x \exp(-x)\, x^n \, dx = (n+1)! \tag{4.72}$$

where

$\langle x^0 \rangle = 1$

$\langle x^1 \rangle = 2$

$\langle x^2 \rangle = 6$

$\langle x^3 \rangle = 24$

$\langle x^4 \rangle = 120$

$\langle x^5 \rangle = 720$

$\langle x^6 \rangle = 5040$

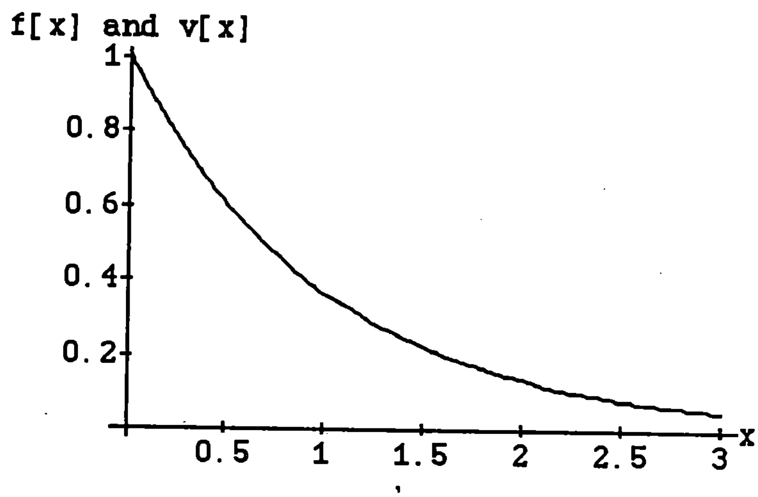

Figure 4.2

Comparison of

f[x] and v[x]

substitute these values into Equation (4.66)

$$\begin{vmatrix} 1 & 2 & 6 & 24 \\ 2 & 6 & 24 & 120 \\ 6 & 24 & 120 & 720 \\ 24 & 120 & 720 & 5040 \end{vmatrix} \begin{matrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \\ 4\lambda_4 \end{matrix} = \begin{matrix} 0 \\ 1 \\ 4 \\ 18 \end{matrix} \quad (4.73)$$

the solutions are

$\lambda_1 = -3$

$\lambda_2 = 1.5$

$\lambda_3 = -0.22222$

$\lambda_4 = 0.0104167$

therefore the function can be approximated as

$$V(x) = \exp(-2.75 + 1.5\,x - 0.3\,x^2 + 0.222\,x^3 - 0.0104167\,x^4) \quad (4.74)$$

This expression is plotted in Figure 4.3.

Application in Finite Domain

The domain of the problem of the probability function is a finite region $a =< x =< b$. In this case, we define the moments

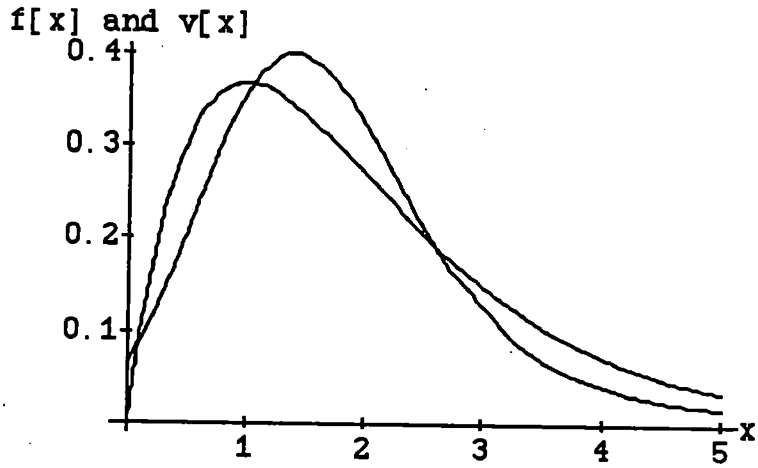

Figure 4.3
Comparison of
f[x] and v[x]

$$< x^n > \;=\; \int_a^b v(x)\, x^n\, dx \tag{4.75}$$

and the Fourier transform of the probability function

$$\phi(k) \;=\; \int_a^b v(x)\exp(ikx)\,dx \;=\; \sum_{n=a}^{b} \frac{(-k)^n < x^n >}{n!} \tag{4.76}$$

if we require that the probability distribution is not zero at the initial and the end point, then we must use the boundaries conditions to calculate the Lagrange multipliers. Substituting these conditions into the Equation (4.22). then we have

$$\begin{vmatrix} <x^0> & <x^1> & <x^2> & <x^3> & \dots \\ <x^1> & <x^2> & <x^3> & <x^4> & \dots \\ <x^2> & <x^3> & <x^4> & <x^5> & \dots \\ <x^3> & <x^4> & <x^5> & <x^6> & \dots \\ \dots & \dots & \dots & \dots & \dots \end{vmatrix} \begin{matrix} \lambda_1 \\ 2\lambda_2 \\ 3\lambda_3 \\ 4\lambda_4 \\ \dots \end{matrix} \;=\; \begin{vmatrix} V(a) - V(b) \\ <x^0> + aV(a) - bV(b) \\ 2<x^1> + a^2V(a) - b^2V(b) \\ 3<x^2> + a^3V(a) - b^3V(b) \end{vmatrix} \tag{4.77}$$

Example 1:

Let

$$V(x) \;=\; x \exp(-x^2) \tag{4.78}$$

the moments are defined as

$$< x^n > \; = \int_0^{\infty} x \exp(-x^2)\, x^n \, dx \; = \frac{1}{2}\Gamma\left(\frac{n+2}{2}\right)$$

(4.79)

where

$<x^0>$ = 0.5

$<x^1>$ = 0.44311134

$<x^2>$ = 0.5

$<x^3>$ = 0.66467

$<x^4>$ = 1

$<x^5>$ = 1.6616755

$<x^6>$ = 3

using the Equation (4.76), we can approximate the Lagrange multipliers

λ_1 = -10.2

λ_2 = 12.415

λ_3 = -5.653

λ_4 = 0.969

therefore, the function can be approximated as

$$V(x) = \exp(-3.5 + 10.2\,x - 12.415\,x^2 + 5.653\,x^3 - 0.969\,x^4)$$

(4.80)

This function is plotted at Figure 4.4.

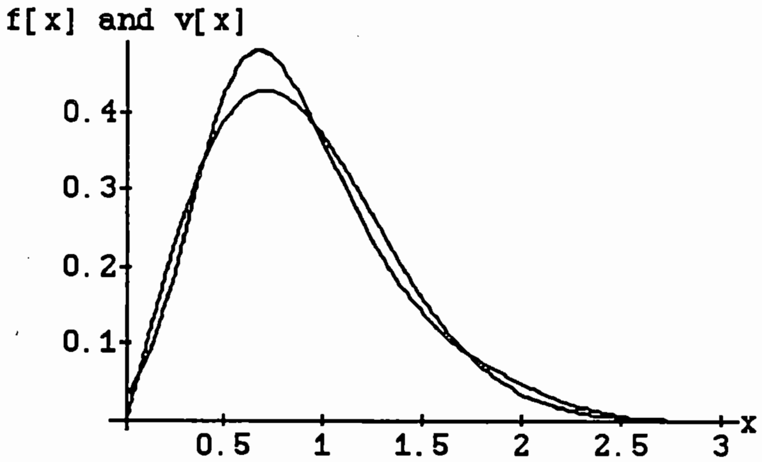

Figure 4.4

Comparison of

f[x] and v[x]

Example 2: Gamma distribution.

Gamma distribution is defined

$$P(y) = \frac{1}{\beta^{\alpha}}\left[\frac{y^{\alpha-1}}{\Gamma(\alpha)}\right]\exp\left(-\frac{y}{\beta}\right) \tag{4.81}$$

the moment generating function m(t) for a continue random variable y is defined as E[exp(ty)].

therefore

$$m(t) = E[e^{ty}] = \int_0^{\infty} \frac{\exp(ty)\, y^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} \exp\left(-\frac{y}{\beta}\right) dy$$

$$= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\left[\frac{\beta^{\alpha}}{(1-\beta t)^{\alpha}}\Gamma(\alpha)\right]$$

$$= \frac{1}{(1-\beta t)^{\alpha}} \tag{4.82}$$

using Taylor's expansion series, we can expand Equation (4.82) thus

$$= 1 + (-\alpha)\,(1)^{-\alpha-1}(-\beta t) + \frac{(-\alpha)(-\alpha-1)(1)^{-\alpha-2)}(-\beta t)^2}{2!} + \ldots$$

$$= 1 + t(\alpha\beta) + \frac{t^2[\,\alpha(\alpha+1)\beta^2]}{2!} + \frac{t^3[\,\alpha(\alpha+1)(\alpha+3)\beta^3]}{3!} + \ldots \tag{4.83}$$

therefore, in general

$$< x^k > \; = \; \alpha(\alpha+1)(\alpha+2)...(\alpha+k-1)\beta^k \qquad (4.84)$$

If we choose $\alpha = 2.5$, $\beta = 2$, then

$<x^0> \; = \; 1$

$<x^1> \; = \; 5$

$<x^2> \; = \; 35$

$<x^3> \; = \; 315$

$<x^4> \; = \; 3465$

$<x^5> \; = \; 45045$

$<x^6> \; = \; 675675$

using Equation (4.76), we can approximate the Lagrange multipliers

$\lambda_1 \; = \; -1.50001$

$\lambda_2 \; = \; 0.300003$

$\lambda_3 \; = \; -0.190478E-1$

$\lambda_4 \; = \; 0.396831E-3$

therefore, the function can be approximated as

$$V(x) \; = \; \exp(-4.1 + 1.5\,x - 0.3\,x^2 + 0.019\,x^3 - 0.000397x^4) \qquad (4.85)$$

This function is plotted at Figure 4.5.

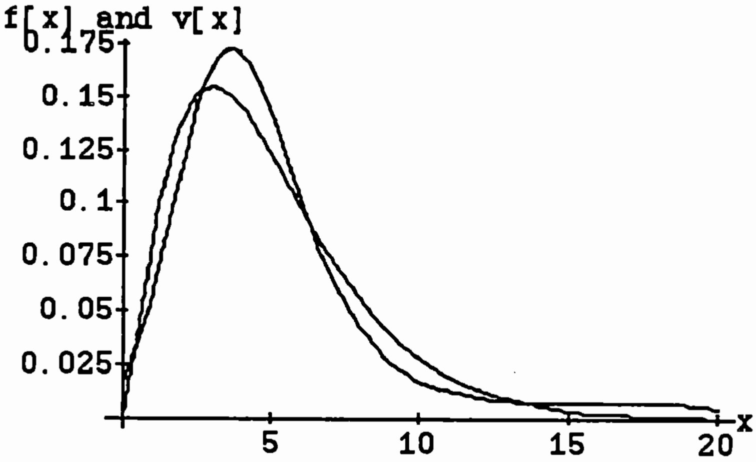

Figure 4.5

Comparison of

f[x] and v[x]

Constructing PDF from the Experimental Data

Experiment 1. This experiment is plotted at Figure 4.6.

Experiment 2. This experiment is plotted at Figure 4.7.

Experiment 3. This experiment is plotted at Figure 4.8.

Experiment 4. This experiment is plotted at Figure 4.9.

Experiment 5. This experiment is plotted at Figure 4.10.

All these experimental data are obtained from the lab at the University of Colorado at Denver. They are unknown probability density functions.

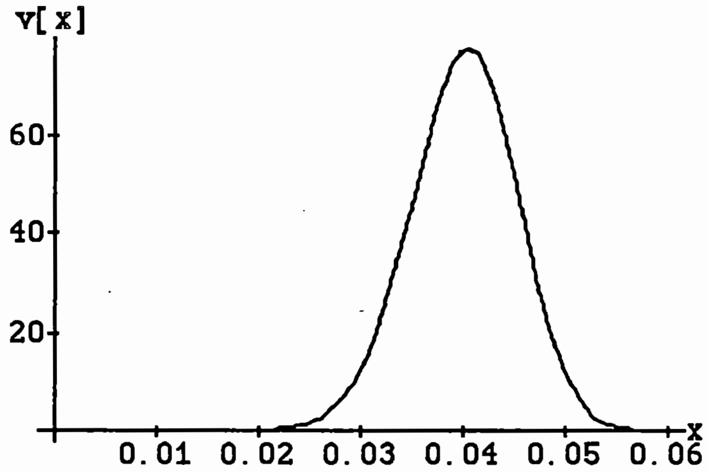

Figure 4.6

v[x] is an unknown

probability density function

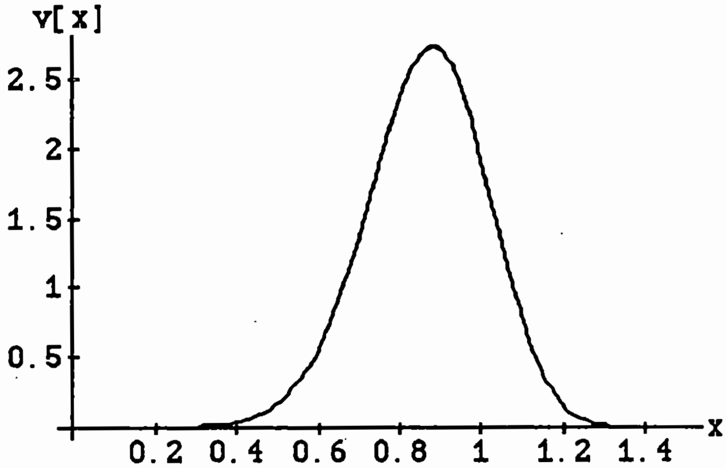

Figure 4.7

v[x] Is an unknown

probability density function

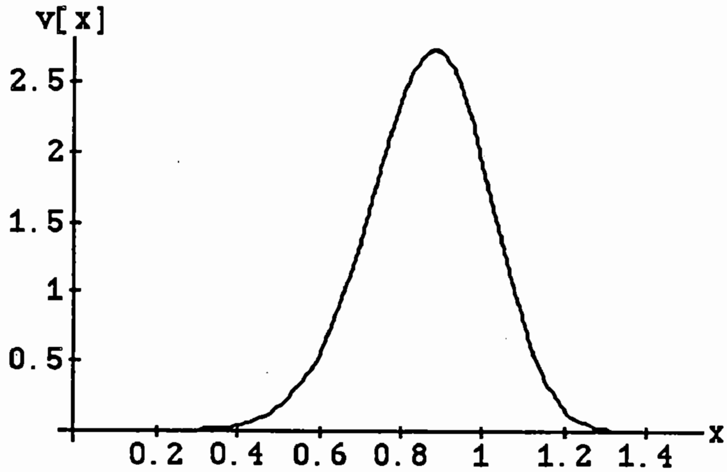

Figure 4.8

v[x] is an unknown probability density function

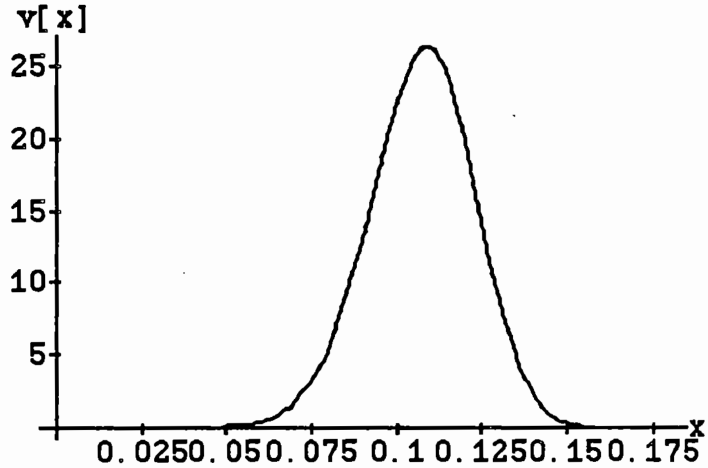

Figure 4.9

v[x] is an unknown

probability density function

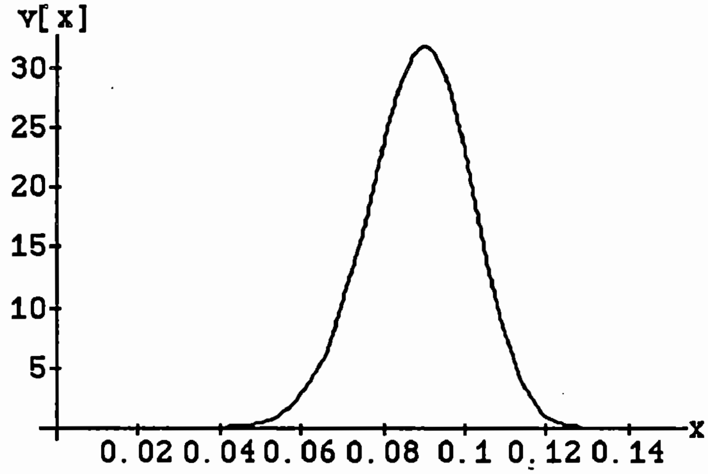

Figure 4.10

v[x] is an unknown

probability density function

CHAPTER 5

DISCUSSION AND CONCLUSION

In this paper, we have carefully examineD the classical moment problems, using maximum entropy. An explicit procedure for approximating the Lagrange multipliers has been developed foe infinite and finite domains for the classical problems. This method is based on developing a differential equation for the Fourier Transform of the probability distribution. This new method allows the physical interpolation of the various Lagrange multipliers in term of the cumulants and moments of the distribution. The algorithm is good to solving certain types of non-linear differential equations. The difference between this algorithm and the old classical algorithm is this new method provides a significant result which eliminates the tedious solution of systems of non-linear equations by solving two linear differential systems of equations. Also, the new algorithm does not has an assumption as in the maximum likelihood problem that the underlying probability distribution is Gaussian distribution, which is not always true in the real world.

The MAXENT algorithm allows us to truncate the infinite sum in the expansion of the Fourier Transform the probability density function. In fact, this algorithm only use the first few moments to approximate the Lagrange multipliers, but give me good results.

APPENDIX

```
! This program is used to plotted Figure 4.1.
! Program language: Mathematica.
a = 0
b =1
c = 0
d = 0
g[x_] = Exp[-a*x - b*x^2 - c*x^3 - d*x^4]
e = 0
v[x_] = Exp[-e - a*x - b*x^2 - c*x^3 - d*x^4]
Plot[ v[x], {x, -3, 3}, AxesLabel -> {"x", "v[x]"}]
f[x_] =  Exp[-x^2]
Plot[ { v[x], f[x] }, {x, -3, 3}, AxesLabel -> {"x", "f[x] and v[x]"} ]
```

```
! This program is used to plotted Figure 4.2.
! Program language: Mathematica.
a = 1
b = 0
c = 2.9*10^-8
d = 3.7*10^-9
g[x_] = Exp[-a*x - b*x^2 - c*x^3 - d*x^4]
init = NIntegrate[ g[x], { x, 0, 10} ]
e = Log[ init ]
v[x_] = Exp[-e - a*x - b*x^2 - c*x^3 - d*x^4]
Plot[ v[x], {x, 0, 3}, AxesLabel -> {"x", "v[x]"}]
f[x_] =  Exp[ -x ]
Plot[ { v[x], f[x] }, {x, 0, 3}, AxesLabel -> {"x", "f[x] and v[x]"} ]
```

```
! This program is used to plotted Figure 4.3.
! Program language: Mathematica.
a = -3
b = 1.5
c = - 0.22222
d = 0.0104167
g[x_] = Exp[-a*x - b*x^2 - c*x^3 - d*x^4]
init = NIntegrate[ g[x], { x, 0, 10} ]
e = Log[ init ]
v[x_] = Exp[-e - a*x - b*x^2 - c*x^3 - d*x^4]
Plot[ v[x], {x, 0, 5}, AxesLabel -> {"x", "v[x]"}]
f[x_] =  x*Exp[ -x ]
Plot[ { v[x], f[x] }, {x, 0, 3}, AxesLabel -> {"x", "f[x] and v[x]"} ]
```

```
! This program is used to plotted Figure 4.4.
! Program language: Mathematica.
a = -10.2
b = 12.415
c = - 5.653
d = 0.969
g[x_] = Exp[-a*x - b*x^2 - c*x^3 - d*x^4]
init = NIntegrate[ g[x], { x, 0, 10} ]
e = Log[ init ]
v[x_] = Exp[-e - a*x - b*x^2 - c*x^3 - d*x^4]
Plot[ v[x], {x, 0, 5}, AxesLabel -> {"x", "v[x]"}]
f[x_] =  x*Exp[ -x ]
Plot[ { v[x], f[x] }, {x, 0, 3}, AxesLabel -> {"x", "f[x] and v[x]"} ]
```

```
! This program is used to plotted Figure 4.5.
! Program language. Mathematica.
a = -1.5
b = 0.3
c = - 2/105
d = 1/2520
g[x_] = Exp[-a*x - b*x^2 - c*x^3 - d*x^4]
init = NIntegrate[ g[x], { x, 0, 10} ]
e = Log[ init ]
v[x_] = Exp[-e - a*x - b*x^2 - c*x^3 - d*x^4]
Plot[ v[x], {x, 0, 5}, AxesLabel -> {"x", "v[x]"}]
f[x_] =  x*Exp[ -x ]
Plot[ { v[x], f[x] }, {x, 0, 3}, AxesLabel -> {"x", "f[x] and v[x]"} ]
```

```
! program: MAXENT.BAS
! programming language: Vax Basic
print chr$(30); chr(92); chr(50);chr(74);
print chr$(27); chr(92); chr(74);
main:
y$ = " y"
        until ( y$ <> " y " and y$ <> " y " )
                input" input (size) n1 "; n1%
                input" input ( size) n "; n%
                input" iteration time 'd' "; d
                length = 2 * n1% - 1
                length = 2 * n% - 1
                dim  m(n1%,n1%), inv_m(n1%,n1%), b(n1%,1),
                        lamda(n1%,1), g(length,1)
                dim  ma(n%,n%), inv_ma(n1%,n1%), c(n1%,1),
                        x(n1%,1), h(length,1)
                print" input matrix element g[i] size is n1"
                mat input g
                when error in
                        for d1 = 1 to d
                        mat b = zer
                        for l = 2 to n1%
                                b(l,1) = (l-1)*g(l-1,1)
                        next l
                        k = 0
                        for i = 1 to n%
                                for j = 1 to n1%
                                        m(i,j) = j * g(j+k,l)
                                next j
                                k = k + 1
                        next i
        call print_matrix ("matrix m", m(,) )
        call print_matrix("matrix b", b(,) )
                mat inv_m = inv(m)
        call print_matrix (" inversed matrix inv_m", inv_m(,) )
                call calculate_solution_x ( lamda(,), inv_m(,), b(,))
        call print matrix ("lamda ", lamda(,) )
        mat c = zer
        c(1,1) = -g(5,1) * lamda(1,1) + 4 * g(4,1)
        c(2,1) = 5 * g(5,1)
        mat h = zer
        h(1,1) = -6
        for l = 2 to 5
```

```
                h(l,1) = lamda( l-1, 1)
        next l
        mat ma = zer
        k = 3
        for i = 1 to n%
                b0 = 0
                for j = 1 to n%
                        if ( ( b0+k) >= 1) then
                                ma(i,j) = h(l,1)
                                h(l,1) = h(l,1) - 1
                        else
                                ma(i,j) = ( b0+k-1) * h(b0+k, 1)
                        end if
                end if
                b0 = b0 + 1
                next j
                k = k - 1
        next i
        call print_matrix ( "matrix ma", ma(,) )
        call print_matrix ( "matrix c ", c (,) )
        call print_matrix ( "matrix h ", h (,) )
                mat inv_ma = inv( ma )
        call print_matrix (" inversed matrix ma", inv_ma(,) )
                call calculate_solution_x ( x(,), inv_ma(,), c(,) )
        call print_matrix ( "x", x(,) )
        for l = lengh-1 to lengh
                g(l,1) = x(l-5,1)
        next l
        call print_matrix ("solution matrix g", g(,) )
next d1
call print_matrix ("solution matrix x", x(,) )
use
print " unresolvable matrix, try again"
end when
input" continue.; y$
next
end

sub initial_lamda ( hfloat m(,), b(,) )
        mat b = zer
        for l = 2 to n1%
                b(l,1) = ( l-1 ) * g( j+k, 1)
        next l
```

```
        k =  0
        for i = 1 to n1%
              for j = 1 to n1%
                    m( i, j ) = j * g(j+k,1)
              next j
              k = k + 1
        next i
end sub

sub initial_ma ( ma(,), c(,), h(,), lamda(,) )
        mat c = zer
c(1,1) = -g(5,1) * lamda(1,1)  + 4 * g(4,1)
c(2,1) = 5 * g(5,1)
mat h = zer
h(1,1) = -6
for l = 2 to 5
        h(l,1) = lamda(l-1,1)
next l
mat ma = zer
k = 3
for i = 1 to n%
        b0 = 0
        for j = 1 to n%
              if ( k+b0) >= 1) then
                    if ( ( b0+k-1) = 0 ) then
                          h(l,1) = h(l,1) -1
                          mat( i,j) = h( l,1)
                    else
                          a(i,j) = (b0+k-1) * h( b0+k, 1)
                    end if
              end if
              b0 = b0 + 1
        next j
        k = k - 1
next i
end sub

sub calculate_solution x ( hfloat x(,), inv_a(,), b(,) )
        mat x = inv_a * b
end sub

sub print_matrix ( a$, hfloat out_matrix(,) )
        print a$
```

```
    mat print out_matrix;
end sub
```

BIBLIOGRAPHY

[1] C. Shannon, "A mathematical theory of communication." Tech Rep. Bell Syst. Technical Journal,1948.

[2] E. Jaynes, "Information theory and statistical mechanics." Phys. Rev. Vol. 106, P. 620, 1957.

[3] L. Mead and N. Papanicolaon, "Maximum entropy in the problem of moments."Journal of Mathematical Physics, Vol. 25, P 2204,1984.

[4] J. Baker-Jarvis, J. Alameddine, and M. Racine, "Solving differential equations by a maximum entropy-minimum norm method with a application to fokker-plank equations." J. Math Phys,Vol.30, PP.1459-1463, 1989.

[5] A. K. Majumdar, "Uniqueness of statistics derived from moments of irradiance fluctuations in atmospheric optical propagation." Optical Communications, Vol. 50, PP. 1-7,1984.

[6] J. Baker-Jarvis and R. Geyor, "Application of the method of maximum entropy in stochastic processes and Inverse problems." (To be published).

[7] M. G. Kendall and A. Stuart, "The advanced theory of statistics." New York: Hafner publishing company, Vol.1, PP. 55-104,1969.

[8] W. Menderhall, R. L. Scheaffer, and D. D. Wockerly, "Mathematical statistics with application." Boston: Duxbury Press, 1978.

[9] R.G. Frehlich and J. H. Churnside, "Probability density function for estimates of the moments of laser scintillation." SPIE Proc., Infrared and Millimeter wave Propagation Engineering, 1988.

[10] S. A. Brown-Van Hoozer, "Statistical laser irradiance fluctuations through atmospheric turbulence based on higher-order moments." Master Thesis, University of Colorado at Denver, 1990.

[11] A. K. Majumdar, "Higher-order statistics of laser-irradiance fluctuations due to turbulence." Journal of the Optical Society of America, Vol. I, P.1067. Nov. 1984.