The Charms of the Quadratic VaR

By Olivier Croissant

1 Introduction

There is no doubt that the VaR is one of the most powerful financial concepts that has been added to the tools of financial engineers and risk managers in the last years. The only problem, and it is a big one, is that the VaR is difficult to compute. In fact, what ever the method we choose, we are led to do an approximation at a level or another, in order to get a working schema. Two methods received most of the attention. One of them, recommended with the Risk Metrics Dataset but not compulsory, is the linear approximation of the PV. It has the advantage to be fast . The other method, situated at the other end of the scale of speed, is the Monte Carlo method. It has the advantage to be accurate if we wait long enough, for the variance of the computation to go below our required precision. For large portfolios, it can take hours or even days on ordinary computers. In some cases , it is possible to reduce the computation time dramatically by using sophisticated random generators. But not all the time. As we see, the enormous disadvantage of this method is its slowness. This drawback prevents us from using it for intraday calculation for large portfolios, which is not the case with the linear method. Unfortunately, in order to make an operational system with risk limits set with VaR , it is important to provide desks managers and ultimately traders with VaR simulation tools that allow them to check the influence of the potential deals on their limits and also to optimize their hedges. Monte Carlo computations made in a few seconds are still a dream, Thus traders are left with linear approximation methods. The limit of these methods is their inabilities to handle options and more generally, situations where the first derivatives of the PV don't represent all the risk associated with the position. So, this prevents us from making simulation for deals like OTC options or exotic swaps .

Facing the need for quick VaR calculations for positions having a strong option flavor, it is very tempting to use the second order approximation of the PV. One of the reasons for hoping to have a satisfactory risk calculation , is that traders currently use gamma maps to hedge positions and to asses risk. The goal is therefore to guide their decisions with more synthetic measures of risk. If they can do it with gammas ,deltas and an intuitive knowledge of the market volatility, it should be possible to build a probabilistic framework that grounds their intuitions.

1.1 The quadratic approximation

We represent the PV of a portfolio of market positions or of a single deal as a function of market variables that we will consider to be fundamental and to be the determinants the state of the market. This is the set of market variables for which we are able to build a complete covariance matrix C. We consider therefore the second order approximation of the PV:

$$PV[x_1, x_2, ..., x_n] \approx PV_0 + \sum_{i=1}^{n} \Delta_i (x_i - x_{0,i}) + \frac{1}{2} \sum_{i,j=1}^{n} \Gamma_{i,j} (x_i - x_{0,i})(x_j - x_{0,j}) \qquad \text{(EQ 1)}$$

At this point, we need to do three remarks .

- First, the PV function which should be used is not the PV of today but the expectation of the PV in one business day if we are looking at a one business day VaR calculation. This is important for options.

-Second , If we want a reasonable risk description, we need to structure the PV function as a function of all variables that are not constant in the PV calculation. That means that beyond the classical interest rate curve , it is necessary to address the dependency of the PV on the volatility of rates, on the major-minor spreads and on volatility smile of the volatility dimensions. The reason of this remark lies in the fact that a lot of exotic positions are hedged in classical risks and are very sensitive toward risks such as spreads and smile effects.

-Third, traders are used to looking at diagonal gammas only (the second derivative with respect to a variable) .However most of the time crossed gammas are important to consider and sometimes they carry all the information about risk .

Even in the case of a simple option, if the expiration date is 18 month and the fondamental variables for which we have entries in the covariance matrix have only 1 years and two years data, it is necessary to map the second order sensitivities. In regular cases, we have the intuition that the one year and two years point should carry a comparable weight. This decomposition of the gamma on several points implies crossed gamma.(See The Encarta)

Many other examples in which off diagonal elements are important can be found : diff swaps, yield curve swaps, CMS and CMT swaps, bond options, ...

1.2 Mapping process and implicit interpolation function

We need to describe the first and second derivatives of a position with vectors and matrices that have the dimensions of the available covariance matrix. For example, we know how to decompose the first senstivity with respect to 18months rate into a one year and a two year points. This is provided by techniques like the one described in the Risk Metrics technical manual (See Ref. [Longerstaey1]). But we also need to do it also with the gamma matrix . In order to deduce how to do it we need to introduce an interesting concept: the notion of implicit interpolation function.

Let see how it works :

We know that any fix cash flow C_{18m} occurring in 18 month can be computed as the derivative of the PV with respect to the zero coupon value associated with the 18 month maturity :

$$C_{18m} = \frac{\partial}{\partial B_{18m}} PV$$

So we can rewrite this relationship as:

$$dPV = C_{18m} dB_{18m}$$

Splitting the cash flow arriving in 18 months into cash flows occurring in 1 year and 2 years is equivalent to decompose dPV differently :

$$dPV = C_{1y} dB_{1y} + C_{2y} dB_{2y}$$

So the mapping process is equivalent to a decomposition of the differential:

$$dB_{18m} = \frac{C_{1y}}{C_{18m}} dB_{1y} + \frac{C_{2y}}{C_{18m}} dB_{2y}$$

Now, there is a bijective relationship between the zero coupon price and the (continuous) zero coupon rate $B_t = e^{-rt}$. It implies a relationship between the differentials:

$$dB_t = -tB_t dr_t$$

and putting evrything together, we deduce the relationship between the differentials:

$$:dr_{18m} = \frac{C_{1y} \cdot 1y \cdot B_{1y}}{C_{18m} \cdot 18m \cdot B_{18m}} dr_{1y} + \frac{C_{2y} \cdot 2y \cdot B_{2y}}{C_{18m} \cdot 18m \cdot B_{18m}} dr_{2y}$$

This relationship reminds us the one we get when we differentiate an interpolation function as $r_{18m} = f(r_{1y}, r_{2y})$: which is ;

$$dr_{18m} = \frac{\partial f}{\partial r_{1y}} dr_{1y} + \frac{\partial f}{\partial r_{2y}} dr_{2y}$$

So we can ask when this relationship between differential implies the existence of such an interpolation function. The answer comes from a powerful mathematical theorem called the Poincare lemma. The only condition that has to be checked is that

$$\frac{\partial}{\partial r_{2y}}\left[\frac{C_{1y} \cdot 1y \cdot B_{1y}}{C_{18m} \cdot 18m \cdot B_{18m}}\right] = \frac{\partial}{\partial r_{1y}}\left[\frac{C_{2y} \cdot 2y \cdot B_{2y}}{C_{18m} \cdot 18m \cdot B_{18m}}\right]$$

Here comes the miracle . If we uses the usual mapping process, it works ! (The one described in the Risk Metrics manual for exemple) .

Thus the implicit interpolation function exists and we will say that the mapping derives from it.

Let’s assume that the implicit mapping derives from the following interpolation function :

$$r_{18m} = \frac{r_{1y} + r_{2y}}{2}$$

Then any function of the 18m rate should be considered to be a function of the 1 year and the 2 year rate. Let’s compare the taylor development :

$$PV \approx PV_0 + \delta \cdot (r_{18m} - r_{0,\,18m}) + \frac{1}{2}\gamma \cdot (r_{18m} - r_{0,\,18m})^2$$

Thanks to the implicit interpolation function, this implies that :

$$PV \approx PV_0 + \delta \cdot \left(\frac{(r_{1y} - r_{0,\,1y}) + (r_{1y} - r_{0,\,1y})}{2}\right) + \frac{1}{2}\gamma \cdot \left(\frac{(r_{1y} - r_{0,\,1y}) + (r_{2y} - r_{0,\,2y})}{2}\right)^2$$

This formula gives birth not only to the mapping for the delta that we already knew , but also to the mapping for the gammas and it also shows us how the non diagonal gammas are created.

Particularly, we deduce the new (mapped) delta vector and the new (mapped) gamma matrix :

$$PV \approx PV_0 + \Delta \cdot \begin{pmatrix} dr_{1y} \\ dr_{2y} \end{pmatrix} + \frac{1}{2} {}^t\begin{pmatrix} dr_{1y} \\ dr_{2y} \end{pmatrix} \Gamma \begin{pmatrix} dr_{1y} \\ dr_{2y} \end{pmatrix} \qquad \Delta = \delta \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \qquad \Gamma = \gamma \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

This method gives us a prescription to map any derivative to the available dimensions even if the implicit interpolation function is not globally defined .

1.3 Parametric and non parametric approximation of the distribution

Having computed the delta vector and the gamma matrix of the future PV of the portfolio, the problem is well posed. We need to compute the distribution of the right member of the equation (EQ1). We will get the VaR associated with a level of confidence of x% with the confidence intervals associated with the level of x%.

At first sight the problem is simple . We have to estimate the distribution of a second order polynomial with n variables , assuming that those n variables are jointly normal with a covariance matrix C that we assume given. Let’s define the second order polynomial by the vector Δ and the matrix Γ .Several particular cases are easily handled:

- if $\Gamma = 0$ this case is reduced to the linear case.

- Γ has only k non zero elements that are all located on the diagonal and that are all equal . Also all the off diagonal elements of C with 1 index belonging to the non zero elements of Γ are zero. Then , the searched distribution is the distribution of the sum of a chi2 with k degrees of freedom with an independent normal variable. It can be easily computed with a one dimensional numerical integration.

- for all other cases, in which the number of dimension is greater than 7-8, no general algorithms giving a quick answer are known.

1.3.1 An Alternative to Var Calculation

The difficulty of computing the quadratic VaR can be circumvented by accepting a different definition for VaR. Instead of computing a confidence interval associated with a level of, let say 99 % , we can compute the minimum of the P&L on the domain where the density of probability is greater than 0.5 %. This method results in a computation of the maximum loss and the state of the market that generate this maximum loss (See Ref. [Wilson1]) . But because the definition of this VaR is different from the classical one, it is impossible to compare risks obtained with both VaR Calculations. Nevertheless this method gives an idea of the worst case and can be used for that purpose. The main drawback of this method is that it is necessary to diagonalize the product ΓC , and any algorithm that uses diagonalisation is very sensitive to the accuracy problems associated with the determination of the correlations.

1.3.2 Cumulants and the exact distribution method

The only practical way to reach the distribution, is to determine the cumulants of the distribution. The cumulants are defined as the derivatives in 0 of the logarithm of the characteristic function. Mathematically ,

$$c_n = \frac{d^n}{dt^n} Log[E_x[e^{i(t \cdot PV(x))}]]\Big|_{t=0}$$

Despite their odd definition those quantities are still very natural

- c_1 is the expectation of PV(x) . We are not interested in the expectation since by definition, we will have a constant PV_0 computed in such a way that this expectation will be the actual expected future market value of the portfolio.

- c_2 is the variance of PV(x) , i. e. the square of the standard deviation , the main ingredient used in linear methods.

-c_3 is the skewness of the distribution . It measures the dissymmetry between the risk associated with buying an option and the risk associated with selling this option.

-c_4 is the kurtosis of the distribution . It measures the balance between very rare events and the most likely ones. c_4 introduces a sort of leptokurticity in the distribution.

-c_5, c_6, ... increase our knowledge of the distribution. In Regular cases, the knowledge of all cumulants is equivalent to the knowledge of the distribution. However, it is worth noticing that the higher order is a cumulant, the less information it brings about the distribution

In the normal case, the only non zero cumulants are the first one and the second one. This explains why the cumulants of order greater than 3 can be used to measure how far from a normal curve, the distribution is. These cumulants also have the very interesting property of summing over independent variables. This property explains the origin of their names.

The good news in the context of quadratic VaR is that these quantities are easy to compute, and that their values are robust (for the low order cumulants) . Fo the first orders, their values are :

$$c_2 = {}^t\Delta C\Delta + \frac{1}{2}Tr[(\Gamma C)^2]$$
$$c_3 = 3{}^t\Delta C\Gamma C\Delta + Tr[(\Gamma C)^3]$$
$$c_4 = 12{}^t\Delta C(\Gamma C)^2\Delta + 3Tr[(\Gamma C)^4]$$
....

And it even exists a general formula . In these formula we used the symbol Tr[M] to designate the trace of the matrix M. The trace is obtained with the sum of all diagonal elements of the matrix. This trace is independent of the basis in which it is computed and it has many interesting properties that can be used to speed up the computation.

Computing the distribution from the knowledge of the cumulant is a very classical problem in probability. It is called the inversion problem. In general, it is a difficult task. Some algorithms exist (See Ref. [Imhof1] and Ref [Rice1]). But they need a few dozen cumulants to give us a computation of the 99% confidence interval. By making reasonable assumptions about the shape of the distribution, we can shortcut this need . This is the implicit idea behind the parametrisation of the distribution.

1.3.3 Parametrization of the distribution

In linear methods, the knowledge of the variance is transformed into monetary terms. Because options introduce a dissymmetry between the buying case and the selling case, it is expected that the knowledge of skewness and kurtosis of the portfolio bring the correction to VaR computations that will make these computations usable for the option cases. We are looking for a natural way to transform these risk measures (because it is what cumulants are) into monetary terms. The answer is a parametrization of the distribution.

The linear method and the generalized linear method (in which we use c_2 instead of only the part of c_2 with no Γ for the linear case) both parametrize the distribution with a normal curve. In other words, a family of curves is determined by only one parameter : the variance. In the quadratic method, because we can have at least 2 other pieces of information we need to use a more general family . Consequently, the different quadratic methods will be characterized by the following precise point : “what family of distribution is it used to parameterize the distribution ”.

We will present here two such families. Because the determination of the parameters of the family will be based on cumulants fitting, and because we want to use the 3d and 4th cumulants, these family will be parametrized by four parameters. The additional parameter is a localization parameter that will be used to be sure that we deal with a zero expectation distribution, because we are only interested to modelize the variation of the PV value around its expectation.

1.3.4 The Johnson Family of distribution

In 1949, N. L. Johnson came up with three families of distribution that have attractive features. First they are all derived from the normal distribution by a transformation that insures that to find a confidence interval, we just need to compute it in the normal cases, by for example looking into a table. Then by applying the inverse transformation we get the confidence interval. The simplicity of this method was an invaluable advantage at a time where computers where inexistent. Another advantage, is that a member of the family is available for every set of variance skewness and kurtosis. The family is described by the following equations

$$z = \gamma + \delta \log[y] \quad \text{(EQ 2)}$$

$$z = \gamma + \delta \log\left[\frac{y}{1-y}\right] \quad \text{(EQ 3)}$$

$$z = \gamma + \delta Sinh^{-1}[y] \quad \text{(EQ 4)}$$

where z is assumed to be a standard normal variable and y is a linear function of the PV.

Equation (EQ 2) is the lognormal family of distribution called S_L , Actually, it is a 3 parameters family and it defines the boundary between the family associated with the equation (EQ 3) also called the bounded family S_B , and the unbounded family S_U described by the equation (EQ 4) .

To study the way this family fit the searched distribution , we introduce the notion of modality of a density associated with a distribution. Basically a unimodal density is a density that has only one local maximum. By extension, we call mutimodal a distribution that has several maximums of density. These notions are important because when we deal with non normal densities we always have interest to study the density starting from the maximum density point. Very powerful methods such as the saddle point methods are then applicable (See Ref. [Jensen1]) . The normal distribution is unimodal but when we deal with hedged options , the multimodality with infinite density spikes appear. To understand that, consider a simple case where we have sold a call on an equity stock and we hedge it for a value of the stock that we anticipate and that is different from the present value of the stock. Let begin by assuming that we are in a simplified world where the risk is entirely described by the probabilistic normal distribution of the underlying stock price . Then around its present value, the stock price shows a maximum of density. Therefore in general the density of the position, which is not hedged in the current values of the markets, will reflects this maximum by showing also a maximum of density. Now if we look at the value of the stock for which the option is hedged, then by definition of a hedge when we vary slightly the value of the underlying stock, the value of the position doesn’t vary. This means that the density associated with the associated value of the position is very high. In fact in our simplistic model we can show that it is infinite. In monodimensional case, hedging situations are always associated with an infinite density spike. In reality as soon there is an other source of randomness, like interest rates or stochastic volatility, then we have to make a convolution of this infinite density spike with a normal noise, and the spike is smoothed and become finite. Nevertheless, this spike can be situated far enough from the maximum associated with the present value of the market to be well differentiated and this situation gives rise to a bimodal distribution.

These typical behavior of portfolios containing options show us a problem when we want to parametrize their distribution with a member a the Johnson family. If the searched distribution present a bimodality, then it is necessary to find also a bimodal distribution to match. Fortunately, S_B sometimes presents a bimodality, this happens when $\delta < \frac{1}{\sqrt{2}}$ and $|\gamma| < \delta^{-1}\sqrt{1-2\delta^2} - 2\delta Tanh^{-1}\sqrt{1-2\delta^2}$

but their is no reason that this situation coincides with the cases where the searched distribution is bimodal.

The deep reason is that the fact that the searched distribution is the one of a quadratic function has not been used to select the family of fitting distribution, which is not the case of the next one we present.

1.3.5 A more natural family of distribution

We start from a quadratic form with a set of initial variables that are jointly normal . We know that we can reduce the problem to a basis where the covariance matrix is the identical matrix and the second order matrix of the quadratic form is diagonal (See for example the Ref [Prasolov1]) . In this basis we have to estimate the distribution of the sum of independent second order polynomial of a normal variable.:

$$z = \underbrace{\sum_{i=1}^{p} a_i(x_i+b_i)^2}_{P} - \underbrace{\sum_{i=p+1}^{p+q} a_i(x_i+b_i)^2}_{N} \qquad (a_i > 0)\forall i$$

We have separated the positive quadratics from the negative ones.Let 's assume for now that N=0 which means that we have to compute the distribution of a sum of positive quadratic polynomials . Two extreme cases exist :

- No a_i is more important than another , and we have to sum a large number of them. We are then in the conditions of central limit theorems and we can consider safely that the sum will be close to a normal variable.

-One of the a_i is more important than the others. The others a_j realize together a kind of normal noise. This case will be well represented by a law with the form : $z = aX^2 + bX + cY$ where X and Y are two independent standard normal variables. This general form also parametrizes the preceeding case.

Hopefully, that in general any intermediate cases will be well represented by a member of this family that is able to represent the two extrem cases.

If we come back to the general case with a non zero part P , then we can consider that the general case is represented by a difference between two independent variables having the statistical behavior described previously . So to be fully general, we should use a representation of the form $z = aX^2 + bX + cY^2 + dY$. Among the caracteristics of this distribution family, it is noticeable that :

-The cumulants are easy to compute and the fitting using a solving algorithm is fast.

- The computation of the distribution is equivalent the convolution of two functions using simple operators and being computed by a monodimensional numerical integral. The determination of the confidence intervals is therefore of the same complexity than the computation of a numerical integral and is fast.

Remark: The described family has four essential parameters . Therefore,we need to fit the variance, the skewness, the kurtosis and an additional fifth cumulant to determine the parameters . If we want to use only three data to fit it ; the variance, the skewness and the kurtosis, it is possible to do so by reducing the preceding family to two three parameters subfamily with a boundary as followed :

$$z = aX^2 + bX + cY \quad \text{(EQ 5)}$$

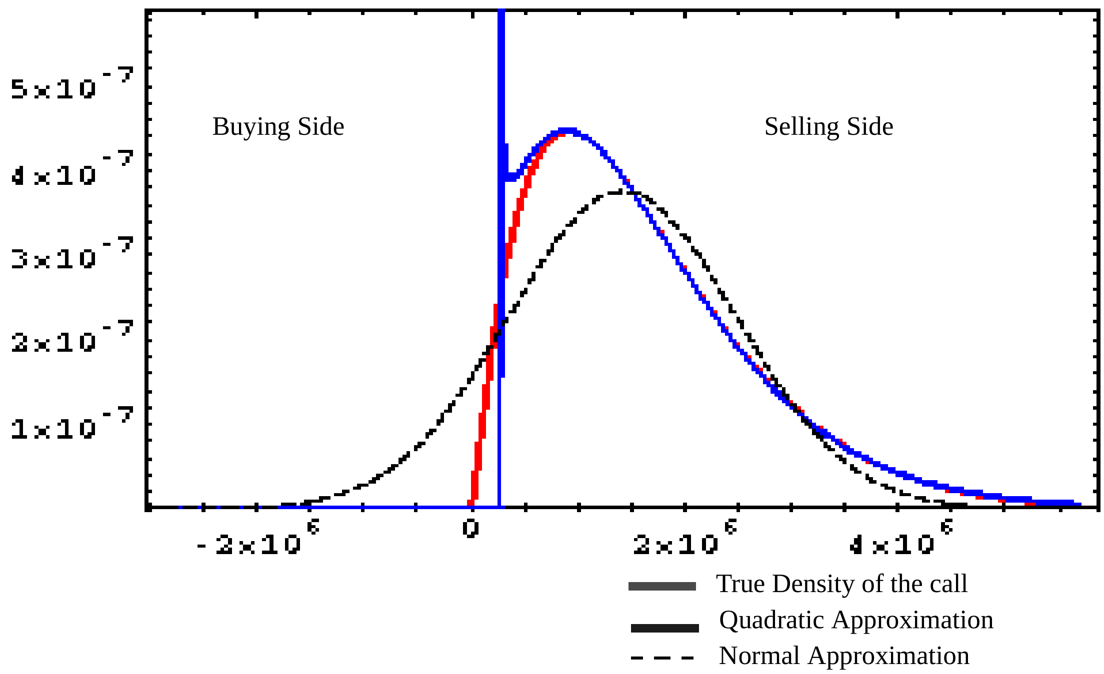
$$z = aX^2 + bX + cY^2 \quad \text{(EQ 6)}$$

The rule being to try to use the first one to fit the cumulants. If it doesn’t work, use the second one. This simplification speed up the computation and most of the time is sensible.

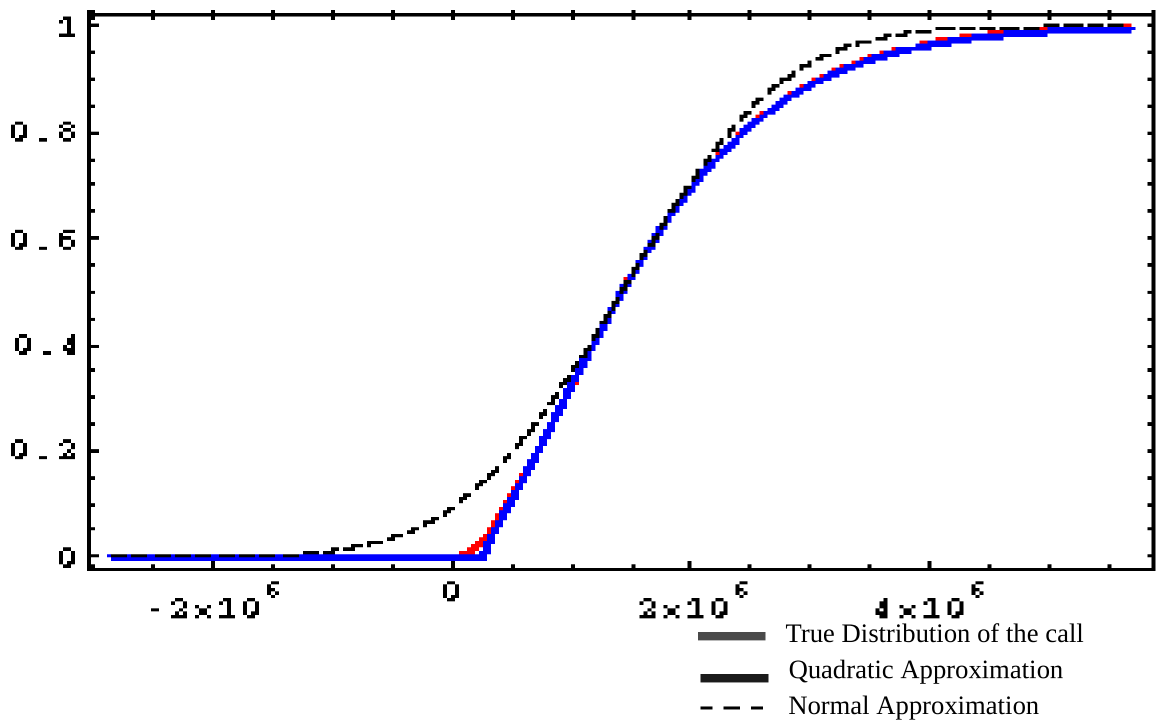
2 Performance of the quadratic approximation

To understand how the quadratic approximation performs compoared to the linear approximation and the real curves, let’s look at a simple at the money call . This case is perfectly adapted to test because this type of option is an important building block of derivative portfolios and we can draw easily the density and the distribution of its value the next buisness day.

Density of a Call at the money, 2 weeks from expiration on a market at 20% volatility per year

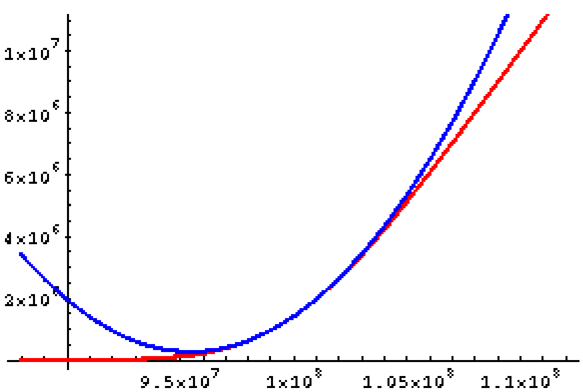


Distribution of a Call at the money, 2 weeks from expiration on a market at 20% volatility per year



We need to do several remarks:

P&L in function of the variation of the underlying



•-Looking at the density ans the distribution, it is obvious that the quadratic approximation performs better on the selling side than on the buying side. There is a topological reason . On selling side, the true P&L curve and the quadratic approximation have a branch that go to infinity, but on the buying side,the true P&L branch tends towards 0 asymtotically when the branch of the quadratic approximation goes to infinty .

•-Despite the large oscillation of the density around the 0 of the premium, the influence of this swing on the distribution is small.

•The Small difference in the zeros of the distributions comes from the fact that the zero derivative point of the quadratic approximation doesn't occur at the zero point of the real P&L curve. The zero property is a global property , whereas the quadratic approximation is a local property.

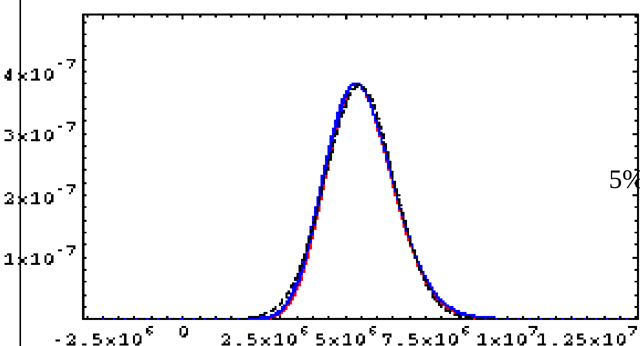
-

Expected Value : 1.417524 Standard deviation : 1.06258 (in $ millions)	Risks	level	Real Risk	Quadratic	relat. err.	Linear	relat. err.	Improv./Lin.
	Buyer	5%	0.988	1.056	-3.29%	1.747	60.1%	94.7%
		1%	1.2721	1.1415	-11.4%	2.472	94.3%	89.1%
	Seller	5%	2.223	2.274	2.25%	1.747	-21.4%	89.2%
		1%	3.401	3.472	4.79%	2.472	-27.3	81.6%

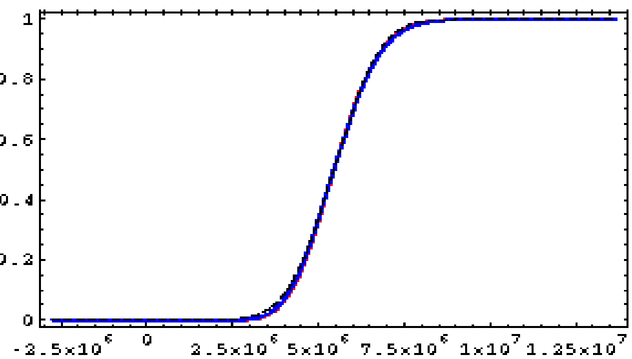
The Column Improvement/Linear measures how well the quadratic approximation performs in replicating the real value of the risk compared to the linear risk. It is computed as $\zeta = 1 - |r-q|/|r-l|$. Where r (respectivly q and l) is the real risk (respectively the quadratic risk and the linear risk) . It works for all non linear instruments.A positive value shows an improvement on the linear approximation, a value of 100% means an exact value for the risk and a negative value is associated with a worst value than with the linear method.

2.1 Other options

Call at the money, 6 months from expiration on a market at 20% volatility per year

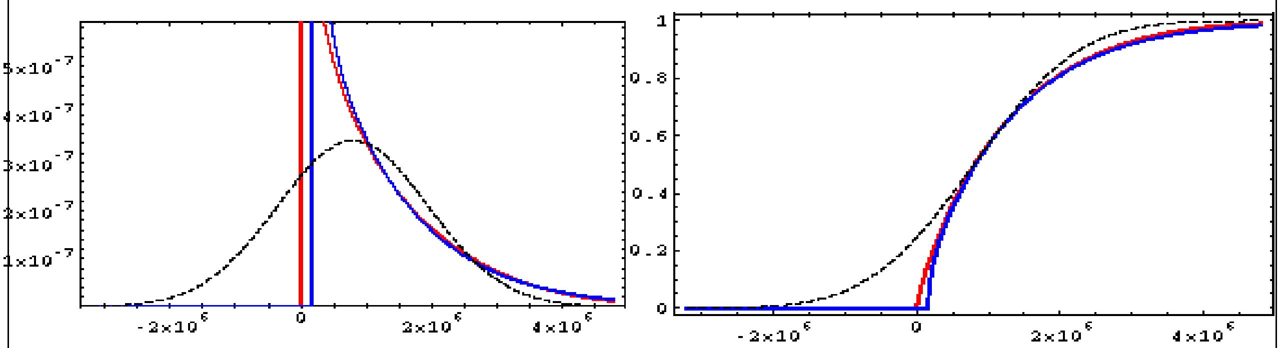


5%



Expected Value :5.4455 Standard deviation : 1.0588 (in $ millions)	Risks	level	Real Risk	Quadratic	relat. err.	Linear	relat. err.	Improv./Lin.
	Buyer	5%	1.5427	1.5828	2.53%	1.7417	12.9%	79.8%
		1%	2.0922	2.1484	2.62%	2.4632	17.7%	84.8%
	Seller	5%	2.279	2.279	0.00%	1.7417	-92.4%	100%
		1%	2.279	2.279	0.00%	2.4632	-89.2%	100%

Call at the money, 6 days from expiration on a market at 20% volatility per year



Expected Value : 0.7974
Standard deviation : 1.1551

(in $ millions)

Risks	level	Real Risk	Quadratic	relat. err.	Linear	relat. err.	Improv./Lin.
Buyer	5%	0.7590	0.7675	1.11%	1.8999	150%	99.3%
	1%	0.9379	0.8287	-13.2%	2.6871	186%	93.8%
Seller	5%	2.5358	2.7376	7.37%	1.8999	-25.1%	68.3%
	1%	3.8609	4.5038	14.3%	2.6871	-30.4%	45.2%

3 Uses of quadratic method

3.1 Risk assessment for option portfolios

The quadratic approximation for the value at risk gives usable results when the options are not to sharp. When there is no highly predominant option positions, the expected acurracy of the var calculation can be better than 5%. The risk diversification effect allied with the convergence toward normality of the central limit theorem insure us that when we deal with more complex cases, the quadratic risk calculation should be good. Of course , when the portfolio contains a predominant option, or when most of the options contained the portfolio are more dangerous than regular options (digital options), the quadratic approximation sees its validity put in danger.

3.2 Hedge optimization for option based trades

Most of the times, hedge ratios computed with the minimisation of the linear risk gives very good hedge ratios.More and more the hedging of highly non lineair complex positions will use liquid options to address the convexity risk. In this case, the only usable short method is the quadratic method

4 Conclusion

Having in a certain domaine of validity a good accuracy, the quadratic approximation finds its place between the linear approximation and the MonteCarlo simulation., It should be appreciated in intradays situation where its speed of calculation and the possibility to use it to compute efficient hedge is determinant.

5 References

[Bott1] R. Bott , L.W. Tu : Differential Forms in Algebraic Topology, 1982 Springer Verlag GTM 82

[Imhof1] J. P. Imhof : Computing the distribution of quadratic forms in normal variables , Biometrika, (48) 1961, 419-426

[Johnson1] N. L. Johnson : Systems of frequency curves generated by methods of translations, Biometrika, (36) , 1949 , p 149-176

[Longerstaey1] : J. Longerstaey : Risk Metrics Technical Manual , J. P. Morgan

[Rice1] S. O. Rice : Distribution of quadratic forms in normal random variables-Evaluation by numerical integration, SIAM J. Sci. Stat. Comput. (1) #4 , December 1980 , p 438-448

[Slifker1] Slifker and Samuel S. Shapiro : The Johnson system , selection and parameter estimation Technometrics (22) #2 , May 1980 , p239-246

[Wilson1] Thomas Wilson : Plugging the gap , Risk , October 1994

[Jensen1] J. L. Jensen : Saddle point Approximation, (1995) Oxford Science Publication

[Prasolov1] V. V. Prasolov : problems and theorems in linear algebra , AMS , (1994) Translation of mathematical monographs Vol 134