Regression Revisited (again)

Isobel Clark
Geostokos Limited, Scotland

Abstract

One of the seminal pioneering papers in reserve evaluation was published by Danie Krige in 1951. In this paper he introduced the concept of regression techniques in providing better estimates for stope grades and correcting for what later became known as the “conditional bias”. In South Africa, the development of this approach led to the phenomenon being dubbed the “regression effect” and ultimately formed the basis of Simple Kriging in Krige’s later papers.

In the late 1950s and early 1960s, Georges Matheron (1965) formulated the general theory of “regionalised variables” and included copious discussion on what he termed the “volume-variance” effect. Matheron defined mathematically the reason for and quantification of the difference in variability between estimated values and the actual unknown values.

In 1983, this author published a paper which combined these two philosophies so that the “regression effect” could be quantified before actual mining block values were available.

In 1996 and some earlier presentations, Krige revisited the regression effect in terms of the conditional bias and suggested two measures which might enable a practitioner of geostatistics to assess the “efficiency” of the kriging estimator in any particular case.

In this article, we revisit the trail from “regression effect” to “kriging efficiency” in conceptual terms and endeavour to explain exactly what is measured by these parameters and how to use (or abuse) them in practical cases.

Introduction

In recent years, major mining and geostatistical software packages have included additional parameters in block kriging analyses such as a “regression” coefficient and a measure of “kriging efficiency”. Danie Krige (1996) presented a paper at the APCOM in Wollongong discussing factors which affect these parameters during resource estimation. Viv Snowden (2001) discussed these measures in her section on classification guidelines released by the AusIMM.

Whilst “kriging efficiency” is a relatively new concept, the “regression effect” has been in use in Southern Africa for over 60 years. In this paper, the basis and development of the regression parameter will be explained in detail and illustrated with an uncomplicated case study. This illustration shows that while a regression correction might correct for the “conditional bias” it does not necessarily improve the confidence in the estimated value for an individual mining block.

Also discussed is the relevance of “kriging efficiency” in the assessment of confidence in estimated block values and a simple discriminator between indicated and inferred values.

The aim of this paper is to show that modern practice world-wide is well founded on over a half-century of established practice in Southern Africa pioneered by Krige on the gold mines.
The case study

This presentation is illustrated by a simulated example where:

- Geology is continuous and homogeneous
- Sampling is precise and accurate
- Sampling is carried out on an absolutely regular square grid
- Sample values are pretty close to Normal

The full database of sampling is as close to exhaustive as possible (less than 0.5% of the range of influence). A block size typical for mine planning and grade control is used throughout this paper, although other (larger) block sizes were studied. This block size corresponds to around 10% of the full range of influence on the semi-variogram. Around 700 blocks are available for assessment and analysis.

The exhaustive sampling is combined to provide 'actual' averages for each block in this model. Each block contains 625 sample points. These will be referred to as actual block averages.

The simplest estimation method – Ordinary Kriging – is used throughout this discussion although similar estimation results could probably be obtained using inverse distance squared.

Varying the sample density

Sampling grids were studied at grid sizes varying from centre of every block up to centre of every 6th block. For simplicity, we present results here only from grids at 3 times the block size and 6 times the block size. Expressed another way, the two sampling grids correspond to 30% and 60% of the range of influence respectively.

Quantifying the regression effect

Krige’s original discussions (1951) were based on very large data sets available from Wits-type gold reefs where chip samples were taken initially in development drives and then, as mining proceeded, on each stope face advance. The average of the development samples was used to evaluate the likely average of a stope panel of (say) 30m by 30m. As the stoping panel was mined out, face samples were used to evaluate the next advance into the stope. As a result, once the stope is mined out, a dense grid of samples is available to determine what was actually mined from that stope.

Krige found that, even allowing for the lognormal nature of the gold values, there was a discrepancy between the average stope value and the average development value. A simple scatter plot of development versus stope averages showed that the relationship between them is neither perfect nor clustered around the 45° line.

In 1972, this author was asked to look into the same question for Geevor Tin Mines Limited in Cornwall, UK. Again, the question was why development averages did not match the stope values found during mining. Again, the obvious approach was to plot development against stope averages to find out where the discrepancy arose.

In this paper we illustrate the approach with a simple example which could be considered as a single bench through an open pit with drilling on a square grid used to estimate the values within each planned mining block.
Sparse sampling – sampling grid spacing 6 blocks

Ordinary Kriging using an isotropic semi-variogram model and search radius at the full range of influence was applied to produce a block model for this illustration. The block size is realistic at 10% of the range of influence and sampling is relatively sparse at 6 times the block size. Around 700 blocks were estimated.

A graph was produced with the kriged estimate along the horizontal axis and the actual block average as the vertical axis. This is shown as Figure 1. The “perfect estimator” is shown as a dashed line with a slope of 1 on this graph. It can be seen clearly that the points on the graph do not lie around this 45° line. The estimated values have a much smaller spread or “dispersion” around thecentre of the graph than do the actual values.

In classical statistics, a “best fit” line can be fitted through the points to find the slope of the line which “best” fits the points. In Least Squares regression (LS) the “best” line is that which minimises the difference between the true average block value and that value which would be estimated using the regression line. This slope is calculated by:

\[
\text{covariance between estimated and actual value} / \text{variance of estimated values}
\]

and the intercept on the line is determined by making sure that the line goes through the average value of all the points for each variable. This classic line is shown in Figure 1 as the solid line.

If these regression factors are applied to the estimates, that is a new estimate is produced which would be:

\[
\text{intercept} + \text{slope} \times \text{kriged estimate}
\]

this should ‘correct’ for the regression effect and produce estimates which lie around the 45° line. At least, this is the stated theory. Figure 2 shows a graph of the kriged estimate corrected for the regression effect using the Least Squares approach. There seems to be little difference between this and Figure 1.

Other types of regression

There are other types of regression lines which can be used to ‘correct’ the estimated values. One of these is the “reduced major axis” (RMA) form, which calculates the slope as:

\[
\text{Standard deviation of actual value} / \text{Standard deviation of estimate}
\]

(cf. Till, 1974). This slope simply rescales the estimates to have the same dispersion as the true values. This is a highly simplified version of an “affine correction”. Figure 1 shows the RMA regression as a dotted line and Figure 3 shows a plot of the corrected estimates using RMA versus the actual block averages. This scatter lies more pleasingly around the 45° line.
As an alternate illustration, the grade/payability curves were calculated for each option:

- Actual block average
- Kriged block average
- Kriged estimate corrected for least squares regression
- Kriged estimate corrected for reduced major axis

These calculations are summarised in Figure 4. It can be seen from this graph that – in this case – severely smoothed estimates can be more realistically corrected by an RMA line than using Least Squares.

It should, perhaps, be noted that the confidence on individual corrected estimates will not be affected by the correction if (and only if) the actual average value over the study area is exactly known. The intercept on the line requires the knowledge of the actual block values. In this illustration the true average value is known for every block and for the whole area. This should be borne in mind for the further sections of this paper.

**Deriving the regression parameter without historical mining**

In Krige’s early work and in the illustration here, the database contains enough information to evaluate the “actual” values which are being estimated. Wits-type gold has been mined since the 19th century and copious amounts of sampling were available for a study such as this. In this author’s early studies, around 50 stopes had already been mined before the correction factors were developed. The advent of Matheron’s Theory of Regionalised Variables (1965) it became apparent that the regression parameters could be determined before half the deposit had been mined out.

Matheron showed that the semi-variogram model – under certain assumptions – was mathematically equivalent to calculating the covariance model. That is, the covariance between two values a certain distance apart can be found from:

\[
\text{Total sill on the semi-variogram model – semi-variogram at that distance}
\]

This relationship can also be used to derive dispersion variances for blocks, viz:

\[
\text{Total sill on the semi-variogram model – average semi-variogram within the block}
\]

In fact, the covariance between any two sets of entities – samples or blocks – can be derived from the semi-variogram model for the samples. The development of the mathematics for the regression factors can be found in Clark (1983).

Krige (1996) uses the following notation:

- \(BV\) represents the dispersion variance of block values within the deposit or study area;
- \(KV\) represents the kriging variance obtained during the estimation of the block

In more traditional geostatistical notation:

- \(BV = \text{total sill on the semi-variogram} - \bar{\gamma}(A,A)\)
- \(KV = \sum w_i \gamma(S_i,A) + \lambda - \bar{\gamma}(A,A)\)
where the total sill on the semi-variogram is taken to be the best estimate for the dispersion variance of single samples within the study area. \( \gamma(A,A) \) represents the average semi-variogram within the block, also known as the “within block variance”. \( \gamma(S_i,A) \) is the semi-variogram between each sample and the block being estimated and \( w_i \) represents the weight given to that sample.

For each block in the model, the slope of the relevant LS regression line can be calculated as follows:

\[
\text{Regression slope} = \frac{(BV - KV + \lambda)}{(BV - KV + 2\lambda)}
\]

where \( \lambda \) represents the lagrangian multiplier produced in the solution of the Ordinary Kriging equations when using the semi-variogram form. If the OK equations are solved using covariances instead of semi-variograms, the sign should be reversed on the lagrangian multiplier. In Snowden’s (2001) notation \( \mu = -\lambda \).

Note, this formula is not actually quoted in Krige (1996). In more traditional geostatistical notation, this formula (cf. Clark 1983) would be written:

\[
\left\{ \text{Total sill} - \sum w_i \gamma(S_i,A) \right\} \div \left\{ \text{Total sill} - \sum w_i \gamma(S_i,A) + \lambda \right\}
\]

To calculate the RMA slope for each block, the formula becomes:

\[
\left\{ \text{Total sill} - \bar{\gamma}(A,A) \right\} \div \left\{ \text{Total sill} - \sum w_i \gamma(S_i,A) + \lambda \right\}
\]

Note also that the formula in Snowden (2001) uses the absolute value of \( \mu \). This would be correct if \( \mu \) is negative but not if \( \mu \) is positive (i.e. \( \lambda \) is negative). It is possible for the \( \lambda \) value to be negative if the sampling layout is very dense or (at least) excessive to needs.

All of the parameters needed to generate the regression slope are available during the Ordinary Kriging process, so that the calculation of the regression slope demands very little extra computation time during block estimation. In this way, the regression slope appropriate to each individual block estimate can be evaluated and included in the output for the block kriging exercise. In Figures 2, 3 and 4 the individual regression parameters were used to provide the “corrected” estimate in each case.

**Kriging efficiency**

The major parameter proposed in Krige (1996) and documented in Snowden (2001) is the “kriging efficiency”. This is a comparative measure for confidence in the individual block estimate. In Krige’s (1996) notation, this parameter is calculated as:

\[
\text{Kriging efficiency} = \frac{(BV - KV)}{(BV)}
\]

And is usually expressed as a percentage rather than the proportion which would be given by the formula. Figure 5 shows the relationship of the two regression parameters (discussed previously) and the kriging efficiency to the kriging variance for each block, using the sparse data grid.

It should be noted that kriging efficiency can take a negative value. As discussed in Krige (1996) the situation at which \( KV = BV \) is when the kriging estimate provides the same level
of reliability as simply using the global average value as the block estimate. This author coined the term \textit{ygiagam}\textsuperscript{1} for this phenomenon and uses this criterion to determine when resources or reserves have to be classified as inferred. As a general rule, any block with a negative kriging efficiency should never be included in a measured resource category.

\textbf{Denser sampling grid}

As a (possibly) more realistic exercise, a sampling grid at 3 block spacing was also studied. Figure 6 shows the kriged estimates versus the actual block values, the 45° line and the overall regression slope for the almost 700 blocks. Figure 7 shows the block estimates corrected by the Least Squares regression coefficient for each block individually. It can be seen from this graph that the LS regression correction works well for this data spacing in this illustration.

\textit{Effect on the grade/payability curve}

To assess the impact of data spacing on the estimated block values, further graphs were produced. Figure 8 shows the comparison between the two restricted data sets at 3 block spacing (3bk) and 6 block spacing (6bk) and the exhaustive potential data set. It is clear from this graph that a limited data set is unlikely to reach the high values which will be encountered during mining (or the lows). According to this example, the limited data sets give a similar general behaviour as regards value (pay grade) but seriously underestimate the likely payable tonnage.

Figure 9 shows a comparison between the kriged block models based on the two sample spacings and the actual block values. The sparse data set is seriously over-estimating the payability for low grades and under-estimating the payability for higher cutoffs – as would be expected by the smoothing shown in previous sections. The grade estimated by kriging from the 6bk sparse grid seriously underestimates the recovered grade at any cutoff compared to the actual block values.

In contrast, the closer spaced sampling (3 times block size) almost exactly mirrors the pay grades in the actual blocks for every cutoff. However, the 3bk sampling seems to over-estimate payability at low cutoffs.

A similar graph using the RMA corrected block grades for the 6bk sampling and the LS corrected block grades from the 3bk sampling would show almost identical grade/payability curves in this particular case.

\textbf{Mis-classification of ore and waste}

The studies reported in this paper, in Clark (1983) and in Krige (1951, 1996) all illustrate how consideration of the regression effect – or conditional bias – can improve estimates for a mining block or stoping model. It is apparent, however, that emphasis on the potentially marginal improvements achieved by regression correction has masked a far more important consideration in mine planning based on estimated block models. Whatever the regression coefficient or the kriging efficiency, there still remains the fact that values allocated to potential mining blocks are still only \textit{estimates}.

\textsuperscript{1} your guess is as good as mine
For any particular cutoff value, there will be blocks which are estimated as payable which will actually be waste. There will be blocks which are estimated to be below cutoff which will actually be payable. This problem is also discussed in detail in Krige’s early work (1951) and many later papers. Figure 10 illustrates the problem of applying cutoff values to the estimated block values using our example with the 3bk sampling, where the regression effect is minimal.

The majority of estimated block values are classified correctly as payable or waste. However, a significant number of blocks are mis-classified. The practical implication of this is that blocks which are actually payable on average will be sent to the waste pile (or not mined) whilst waste material will be mined and delivered to the plant as payable. However much mathematics is applied to this situation, the result will be the same tonnage mined for a lower overall recovered value.

It should be emphasised that this is not a result of using a particular estimation technique but is a fact of production. The only way to eliminate this effect is to instigate a grade control sampling plan which will narrow the scatter on the graph enough to achieve an absolute minimum of blocks in the two punitive quadrants. Or, in simpler language, to achieve a good enough grade control programme to be fully confident in the block values during production.

As a final illustration, a single cutoff was applied to the block estimates using the 3bk sampling for Ordinary Kriging. Those blocks estimated as above cutoff were separated from those classified as “waste”. Grade/payability curves were produced from these two sets of blocks and are shown in Figure 11. In this case, the payability curves are scaled to reflect how many blocks are classified as payable as opposed to waste. This example illustrates the proportion of “payable” blocks which are actually non-payable and the values in the blocks which are classified as “waste”.

Conclusions

Most of the above illustrations could equally well have been produced using theoretical methods. It is not necessary to have a vast database of previously mined areas to produce the regression factors or an assessment of the likely mis-classification errors which will be incurred during production. This paper has used a case study where the database is exhaustive because it has been created from a conditional simulation based on a real life case study.

It should be borne in mind that regression corrections, kriging efficiency and mis-classification assessments depend heavily on the sample values following a Normal(-ish) behaviour. Statistics such as variance and covariance have little real meaning when applied to skewed sample data. The example presented shows that least squares regression works well when the regression slope is less than (or not much greater than) 1. If a high degree of smoothing is present, an alternative approach such as RMA regression or affine corrections should be utilised rather than least squares. Alternatively, simulation studies may be valuable when data is too sparse to achieve realistic results.

It cannot be emphasised strongly enough that a regression correction does not improve confidence in individual estimated block values. There will always be uncertainty in the true value of the block until it has been mined (and maybe afterwards). Mis-classification of payable and non-payable blocks will inevitably lead to reconciliation problems during production – unless allowances are built into the mine plan for those recovery factors.
One puzzling factor is that many software packages now supply regression factors but do not seem to use them in adjusting the block estimates.

Kriging efficiency is simply a standardised form of the kriging variance. An advantage over simply considering the kriging variance is that it does provide an immediate indication of the guideline as to where to stop trying to provide a local estimate for an individual block.

The main aim of this paper has been to show that approaches developed by Danie Krige 60 years ago are still vital in the production of mineral resource and reserve models and in ongoing mine planning.

References


Figures

Figure 1: block values estimated from samples at grid spacing 6 times block size: estimated versus actual value with regression lines

Figure 2: block values estimated from samples at grid spacing 6 times block size: kriged estimated value corrected for least squares regression versus actual value
Figure 3: block values estimated from samples at grid spacing 6 times block size: kriged estimated value corrected for reduced major axis regression versus actual value

Figure 4: block values estimated from samples at grid spacing 6 times block size: grade/payability curves calculated on actual and estimated block values
Figure 5: block values estimated from samples at grid spacing 6 times block size: regression coefficients and kriging efficiency plotted against kriging variance.

Figure 6: block values estimated from samples at grid spacing 3 times block size: estimated versus actual value with regression line.
Figure 7: block values estimated from samples at grid spacing 3 times block size: kriged estimated value corrected for least squares regression versus actual value
Figure 8: samples at various grid spacing: grade/payability curves calculated on sample values

Figure 9: block values estimated from samples at various grid spacing: grade/payability curves calculated on actual and estimated block values
Figure 10: block values estimated from samples at 3 block grid spacing: applying the cutoff value to the estimates instead of the actual values.

Figure 11: block values estimated from samples at 3 block grid spacing: grade/payability curves of actual block values mined as ‘above cutoff’.