Uncovering the Hidden Triangles of Uncertainty Analysis

Julia Peters, Gabriella Magasic

ICEAA Professional Development & Training Workshop

February 18th, 2025

Abstract

Cost estimators require a deep understanding of risk and uncertainty to produce reliable, quality cost models. Since analysts come from diverse academic fields, they may not possess the intersectional knowledge of statistics and linear algebra that manifests in uncertainty analysis. The objective of this paper is to demonstrate this intersection using the intuitive language of geometry: we reduce the abstract problem of determining cumulative uncertainty to the more tangible problem of solving a triangle. The authors leverage the Pythagorean theorem and the Law of Cosines to geometrically explain the summations of independent and linearly correlated standard deviations, respectively. By adopting this theoretical framework for uncertainty analysis, estimators can establish a foundation from which to develop an understanding of other advanced analytical methods at the frontier of cost estimation.

Contents

List of Figures

List of Tables

Uncovering the Hidden Triangles of Uncertainty Analysis

> "When you love a problem, its contours, obstacles, and resistances are all just parts of its character."
>
> – Steven Strogatz

1. Introduction

Cost estimators require a deep understanding of risk and uncertainty to produce reliable, high quality cost models. According to the Government Accountability Office best practices, all credible cost estimates must include an analysis of events that may positively or negatively influence the program [1]. A variety of numerical methods, such as Monte Carlo-based simulations, are built into automated cost estimating software, making the process of quantifying uncertainty accessible to analysts without backgrounds in higher mathematics. Although these software programs lower the barrier to entry for estimators approaching uncertainty analysis, estimators may rely on the software as a problem-solving black box without conceptualizing the underlying mechanisms. In this paper, we aim to intuitively explain the mathematical foundations of uncertainty analysis in a way that is accessible to estimators of diverse academic backgrounds.

Our strategy is to present the formulas for summing together independent and linearly correlated standard deviations and demonstrate the connections to their geometric analogs: the Pythagorean theorem and the Law of Cosines. Geometry makes for a convenient pedagogical tool because we can discuss uncertainty in terms of shapes that are easily visualized in the plane. These simplified, two-dimensional examples function as conceptual steppingstones leading up to higher-dimensional problems. A central tenet of our approach is that mathematics in 2D space generalizes nicely to the higher-dimensional cases that appear in an estimators' day-to-day work. We argue that understanding tangible, 2D examples equips the cost estimator with a conceptual foundation that makes abstract, higher-dimensional mathematics more attainable.

Of the two primary approaches to applying statistical uncertainty, top-down and bottom-up, the methodology described in this paper aligns with the bottom-up application. Both methods arrive at a total point estimate by summing together low-level cost input variables. Whereas the top-down application utilizes a probability distribution on the highest-level point estimate, the bottom-up application assigns an intentionally curated distribution to each cost input variable. Accordingly, the bottom-up approach captures the nuance of high- and low-risk elements, allowing the estimator to derive valuable risk insights that are specific to a given program [2]. However, the bottom-up approach is more computationally involved; it raises the intermediary problem of combining the child-level uncertainties with respect to their interrelated correlations. Confronting this problem brings an analyst into the territory of linear algebra, and this paper aims to intuitively explain the foundational mathematics for uncertainty propagation in the field of cost estimation.

The academic diversity in the field of cost estimation creates a demand for innovative pedagogical techniques that communicate advanced quantitative methods to professionals without backgrounds in higher mathematics. Based on 2008 data, 44% of cost estimators have an academic background in business management, and 25% have a background in finance or accounting. Only 7% of cost estimators hold a degree in mathematics or statistics [3]. In this paper, we utilize geometric intuition to construct a fast-track for acquiring the mathematical foundations of uncertainty analysis. We aim to motivate estimators to think critically about equations, deepening their skillsets as effective quantitative problem solvers.

Our objective is to frame statistical concepts in the context of geometry, where estimators can visualize what it means to quantify uncertainty. With this goal in mind, we structure this paper as follows. Section 2 introduces vectors and demonstrates their utility in statistics. A univariate distribution presents the standard deviation as a scalar quantity, whereas a combination of multiple distributions characterizes the standard deviation as a vector—a quantity with both magnitude and direction. This framework is applied in Section 3, which establishes how to sum independent standard deviations using the Pythagorean theorem. In the field of cost estimating, it is uncommon to encounter uncertainties that are best modeled as being independent from one another [2]. However, the discussion of independent standard deviations serves as a simple case from which we develop the discussion of correlated standard deviations in Section 4. Here, we demonstrate how the Law of Cosines is the geometric analog to the formula for the variance of a sum of random variables, enabling an estimator to conceptualize correlation as the amount that two uncertainty vectors are aligned with one another. In Section 5, we discuss how automated cost estimating software propagates uncertainty through a cost model, enhancing estimators' ability to translate these mathematical principles into their everyday work.

2. Interpreting the Standard Deviation as a Vector

This section introduces the concept of vectors and demonstrates their utility in combining the uncertainties of cost elements. As compared to a scalar quantity, which is fully specified by its magnitude, a vector quantity has both a magnitude and a direction. Whereas the standard deviation of a single variable distribution is treated as a scalar, multivariate distributions may have uncertainty along each dimension, allowing for the standard deviations to be characterized as vectors. Cost estimators encounter multivariate distributions when they apply the bottom-up methodology, combining several single-variable distributions at the child level to an estimate at the parent level.

Vector addition exemplifies how the directional differences between addends influence the sum that is obtained. We present the following physical example to illustrate a situation where vectors make for convenient problem-solving tools. Consider a box initially at rest on a frictionless surface. Two external forces, F_1 and F_2, are applied to the box, which causes the box to move. In this example, F_1 is stronger than F_2, so it is drawn longer in Figure 1, enabling us to geometrically visualize the vectors' magnitudes. To determine the box's subsequent motion, it is not enough to know only the magnitudes of F_1 and F_2; we also need to know the direction each force is acting. Figure 1 shows several ways that the forces F_1 and F_2 could point. For simplicity, all five situations have F_1 pointing horizontally to the right and allow F_2 to vary in direction. Although the lengths, or magnitudes, of F_1 and F_2 do not differ between scenarios a-e, the length of the net force changes depending on the orientation of F_2 relative to F_1.

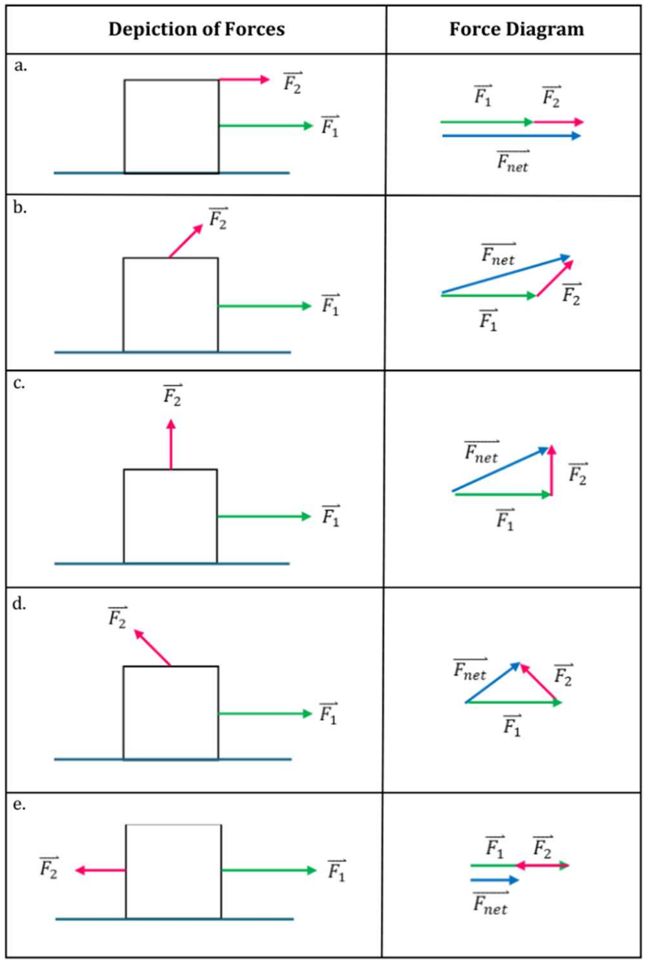


Figure 1 - Using forces to illustrate vector addition

A key takeaway from Figure 1 is that vectors are added tip-to-tail. To geometrically construct the sum $F_1 + F_2$, we draw F_1 as a horizontal arrow to the right. Then, at the arrowhead of F_1, we draw the endpoint of F_2. Shown in blue, the net force is represented as the shortest distance from the endpoint of F_1 to the arrowhead of F_2. Unless the vectors fall along the same line, F_{net} is visualized as the third leg of the triangle, connecting the endpoint of F_1 to the arrowhead of F_2. Instead of applying two forces, F_1 and F_2, separately, one could apply a singular force whose magnitude and direction is the same as F_{net}, and the box's motion would be the same. Therefore, each scenario in Figure 1 represents the equation $F_{net} = F_1 + F_2$ for a particular orientation of the "child-level" vectors.

This paper demonstrates how the formulas for summing together standard deviations can be visualized and understood in terms of vector addition. Figure 2 contains two bell curve distributions, where each one corresponds to a child element of a cost model. Since distribution A has a wider spread, its standard deviation is greater, so its vector is drawn longer than its counterpart from distribution B. Analogous to the force vectors in Figure 1, the magnitude of the cumulative uncertainty, σ_c, in Figure 3 depends on the directional orientation of the child-level standard deviation vectors.

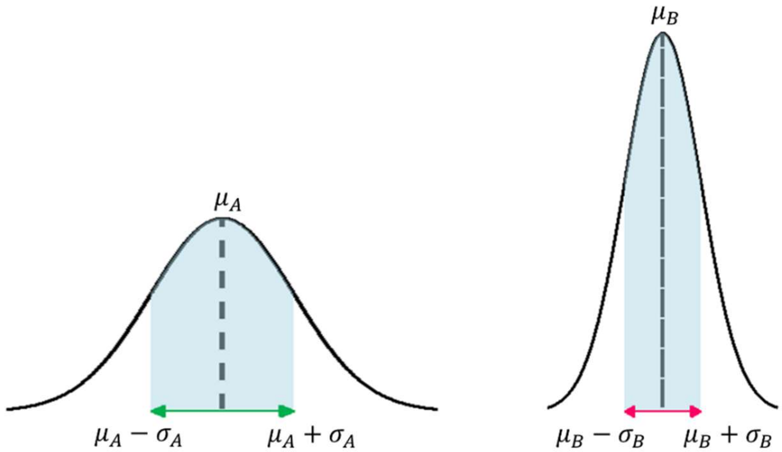


Figure 2 - Probability densities of two child-level cost elements

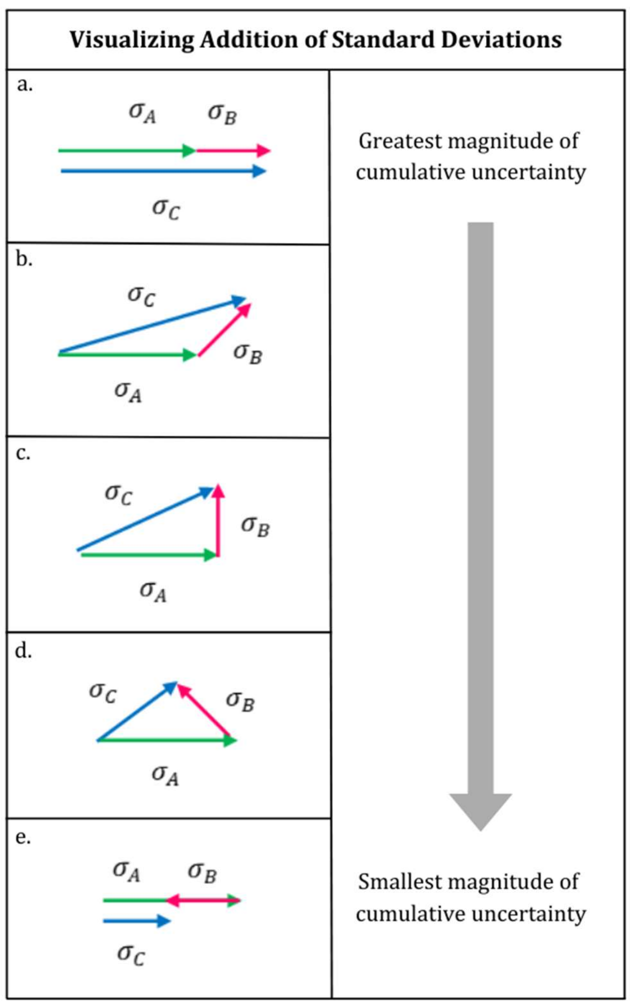


Figure 3 - Various ways to combine the standard deviations in Figure 2

In the physical analogy, it is straightforward to conceptualize pushing or pulling on a box at various angles, validating the use of vectors to approach the problem. However, the source of directional differences in uncertainty analysis is more subtle. Here, the angle between the standard deviation vectors is encoded in the correlation coefficient between the two variables. On a conceptual level, the angle between two vectors and the correlation coefficient between two variables serve a similar purpose: both values quantify how aligned one object is with respect to another. In the following sections, we develop this conceptual foundation, providing cost estimators with geometric explanations of the formulas for summing together independent and correlated uncertainties.

3. Independent Uncertainties and the Pythagorean Theorem

Although independent random variables are relatively rare in cost estimation, we present the special case of independence to establish a foundation for the more common—and complicated—correlated case. This section leverages an example problem to motivate the use of the formula for summing together independent standard deviations. Utilizing a two-dimensional application, we demonstrate how two uncertainties can be visualized as the legs of a right triangle, where the hypotenuse is the cumulative uncertainty. Accordingly, we verify through geometric reasoning why independent standard deviations are added in quadrature.

3.1. Problem Statement

The following problem is adapted from Paul Garvey's *Probability Methods for Cost Uncertainty Analysis* [4]. Suppose the total cost of a system is given by Cost = $X_1 + X_2$. Let X_1 denote the cost of the system's prime mission product. The distribution of X_1 is normal with a mean of 30 \$M and a standard deviation of 12 \$M. Let X_2 denote the cost of system testing. The distribution of X_2 is normal with a mean of 20 \$M and a standard deviation of 5 \$M. Suppose X_1 and X_2 are independent random variables. Determine the expected value of the total project cost and the total project standard deviation.

3.2. Solution

Since the central tendencies are scalar quantities, the calculation of E(Cost) is straightforward: E(Cost) = $E(X_1) + E(X_2)$ = 30 \$M + 20 \$M = 50 \$M. To sum together the standard deviations, we apply

$$\sigma_{total} = \sqrt{\sum_{i=1}^{n} \sigma_i^2} \qquad \text{(Equation 1)}$$

Where *n* is the number of child-level standard deviations. Here, all the values in $\{\sigma_i\}$ are independent from one another. Setting *n* = 2 and plugging in the given quantities, we have

$$\sigma_{total} = \sqrt{\sigma_1^2 + \sigma_2^2} = \sqrt{(12\,\$M)^2 + (5\,\$M)^2} = 13\,\$M$$

So the expected value for the total program cost is 50 \$M with a standard deviation of 13 \$M.

3.3. Geometric Intuition

Evidently, we can map the two-dimensional application of the formula for summing together standard deviations onto the Pythagorean theorem, as both formulas employ addition in quadrature. The Pythagorean theorem is

$$a^2 + b^2 = c^2 \qquad \text{(Equation 2)}$$

Where a, b are the leg lengths of a right triangle, and c is the length of the hypotenuse. In the example above, the right triangle representing the summation of uncertainties has leg lengths a = 5 and b = 12 with a hypotenuse of c = 13.

On a high level, our approach to this problem involves (1) determining the space of possible outcomes for cumulative uncertainty and (2) employing statistical reasoning to enhance our understanding of this space. For a two-dimensional problem, the space of possible outcomes can be visualized as in Figure 4. Here, σ_1 is shown in green, σ_2 in red, and σ_{total} in blue. As a mnemonic device, let's consider σ_1 and σ_2 as the components of a fair spinner. Here, σ_1 is fixed in space, and σ_2 is free to rotate such that its arrowhead traces out the circle as shown. Since the two uncertainties are independent, σ_2 can take any landing position on the circumference with equal probability. The magnitudes of σ_1 and σ_2 are constant, but the magnitude of their sum, σ_{total}, depends on the landing position of σ_2. Thus, the rotation of σ_2 about σ_1 in the plane specifies the length of every possible σ_{total}.

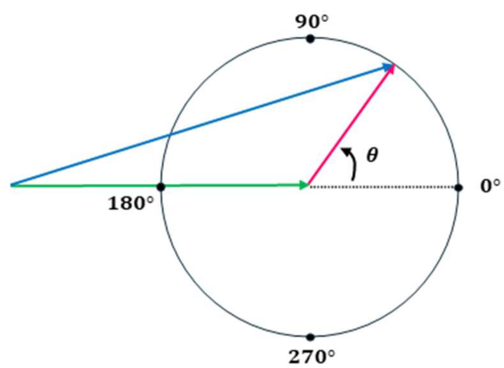


Figure 4 - Rotating one standard deviation vector about the other

However, we can further constrain this space such that our diagram utilizes only the upper semicircle. We are tracking the values that are accessible for the magnitude of σ_{total}. If σ_2 lands below the horizontal, we can reflect it over the horizontal axis of the circle, and the magnitude of σ_{total} is preserved through this reflection. Therefore, the set of endpoints around the entire circumference double-counts each accessible magnitude of σ_{total}. To account for this, we define the bounds of θ, the angle that specifies how much σ_2 has deviated from 0°, as $0° \leq \theta \leq 180°$. Visually, we picture our spinner as a half-spinner, allowing σ_2 to take any value in the upper semicircle with equal probability.

Moreover, the choice to limit the domain of θ so that σ_2 sweeps out a semicircle is consistent with the upper and lower bounds of the magnitude of σ_{total}. The maximum attainable value of σ_{total} occurs when θ = 0°. In other words, when there is zero angular separation between the child-level vectors, they superimpose along the same line. The minimum attainable value of σ_{total} occurs when θ = 180°. In this case, σ_1 and σ_2 fall along the same line, but they cancel one another out. Our objective is to determine a value for θ in between these two extremes at 0° and 180°.

Having defined the space of possible outcomes, we now utilize statistical reasoning to determine a sensible value between the upper and lower bounds, intuitively validating Equation 1. Suppose we spin σ_2 a large number of times and record the θ that we obtain from each trial. Since every landing position is equally probable, we argue by symmetry that the expected outcome of θ is given by the arithmetic mean of the upper and lower bounds. The mean of 0° and 180° is 90°, which means, geometrically, that the child-level uncertainties represent the legs of a right triangle. The cumulative uncertainty is given by the hypotenuse of this triangle, so we use the Pythagorean theorem to solve for this length.

Table 1 summarizes the meaning of the rotations of σ_2 about σ_1 that we have discussed so far. The special case of independence only requires that we know the extreme values of θ and the average value between them. This is because modeling the child-level uncertainties as independent enables us to make the assumption that every landing position for a random spin of σ_2 is equally probable.

Table 1- Summary of outcomes for θ and σ_{total} for independent σ_1 and σ_2

Value of θ	**Description of child level uncertainties**	**Description of magnitude of cumulative uncertainty**
0°	Aligned, superimposing	Upper bound
90°	Perpendicular	Expected value
180°	Anti-aligned, canceling	Lower bound

4. Correlated Uncertainties and the Law of Cosines

This section modifies the previous example problem so that the cost variables have a positive linear correlation, motivating the use of the formula for the variance of a sum of correlated random variables. To frame this problem in the context of geometry, we revisit the spinner analogy. Since the cost variables are not perfectly independent, we cannot assume that every landing position for σ_2 is equally probable, so we use the correlation coefficient to specify the expected landing position for the correlated case. In doing so, we construct a non-right triangle for which we implement the Law of Cosines to solve for the missing side, σ_{total}.

4.1. Problem Statement

Suppose the total cost of a system is given by Cost = $X_1 + X_2$. Let X_1 denote the cost of the system's prime mission product. The distribution of X_1 is normal with a mean of 30 $M and a standard deviation of 12 $M. Let X_2 denote the cost of system testing. The distribution of X_2 is normal with a mean of 20 $M and a standard deviation of 5 $M. Suppose X_1 and X_2 are positively correlated with correlation coefficient $\rho = 0.5$. Determine the expected value of the total project cost and the total project standard deviation.

4.2. Solution

As before, the calculation of E(Cost) is straightforward: E(Cost) = E(X_1) + E(X_2) = 30 $M + 20 $M = 50 $M. To sum together the standard deviations, we apply Equation 3, the formula for the variance of a sum of random variables:

$$Var(Cost) = \sum_{i=1}^{n} \sigma_i^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \rho_{X_i,X_j}\sigma_i\sigma_j \qquad \text{(Equation 3)}$$

Equation 3 supports the summation of an arbitrary number, *n*, of correlated standard deviations. Setting *n* = 2 and plugging in the given quantities, we have

$$Var(Cost) = \sigma_1^2 + \sigma_2^2 + 2\rho_{X_1,X_2}\sigma_1\sigma_2$$

$$Var(Cost) = (12\ \$M)^2 + (5\ \$M)^2 + 2(0.5)(12\ \$M)(5\ \$M) = 229\ \$M^2$$

$$\sigma_{total} = \sqrt{Var(Cost)} = \sqrt{229\ \$M^2} \approx 15.13\ \$M$$

So the expected value for the total program cost is 50 $M with a standard deviation of approximately 15.13 $M.

Notably, the expected value for the mean is 50 $M regardless of whether the variables are independent or correlated; however, the total uncertainty increases from 13 $M in the independent case to ~15.13 $M in the positively correlated case. We use the Law of Cosines to provide the geometric intuition behind Equation 3 and explain the change in σ_{total} from the independent case to the correlated one.

4.3. Geometric Intuition

The correlated case maps nicely onto the spinner analogy established in Section 3. Instead of assuming that the spinner is fair, suppose the spinner is biased towards a particular landing position other than 90°. For this example, we rig the spinner so that the expected value of θ is 60°. This is because taking the cosine of 60° returns the correlation coefficient of 0.5. The corresponding expected value of σ_{total}, shown in Figure 5, is represented as the length of side c. To solve for the missing side, σ_{total}, we implement the Law of Cosines[1]:

$$c^2 = a^2 + b^2 + 2ab\cos\theta \qquad \text{(Equation 4)}$$

Where a, b, c are side lengths, and θ is the supplement of the angle opposite side c. We choose the convention for θ to be the rotation of σ_1 counterclockwise from 0°. The 0° angular position of σ_1 relative to σ_2 represents perfect alignment, maximizing the sum σ_{total}. It follows that a small deviation from perfect alignment is represented by a small, positive value of θ, so it is convenient to define θ as the supplement of the interior angle opposite the side σ_{total}.

[1] See Appendix A for a discussion of the chosen sign convention.

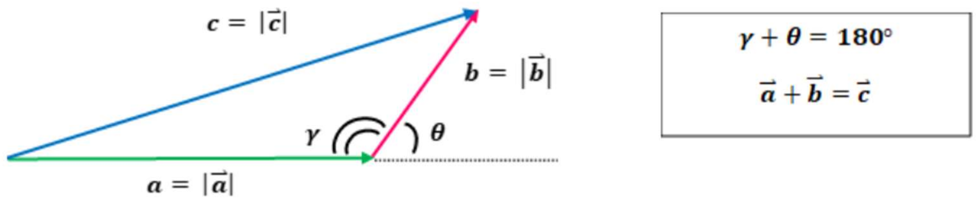


Figure 5 - Chosen convention for the Law of Cosines

This definition of θ is also consistent with the derivation of the Law of Cosines from basic definitions in the algebra of vectors. We present the following derivation to emphasize how Equations 3 and 4 come directly from vector addition. The vectors shown in Figure 5 represent the equation

$$\vec{c} = \vec{a} + \vec{b} \qquad \text{(Equation 5)}$$

We introduce the vector dot product to derive the Law of Cosines. For vectors $\vec{u}, \vec{v} \in R^n$, the dot product is defined as

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos \theta_{\vec{u},\vec{v}} \qquad \text{(Equation 6)}$$

Where $|\vec{u}|, |\vec{v}|$ denote the magnitudes of $\vec{u}, \vec{v}$ and θ is the angle between the vectors. In the two dimensional case, we square both sides of Equation 5.

$$\vec{c} \cdot \vec{c} = (\vec{a} + \vec{b}) \cdot (\vec{a} + \vec{b}) \qquad \text{(Equation 7)}$$

Carrying out each dot product gives

$$|\vec{c}||\vec{c}| \cos\left(\theta_{\vec{c},\vec{c}}\right) = |\vec{a}||\vec{a}| \cos(\theta_{\vec{a},\vec{a}}) + |\vec{b}||\vec{b}| \cos(\theta_{\vec{b},\vec{b}}) + 2|\vec{a}||\vec{b}| \cos \theta_{\vec{a},\vec{b}} \qquad \text{(Equation 8)}$$

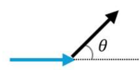
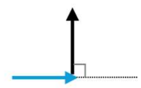
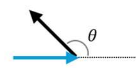
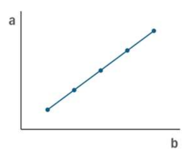
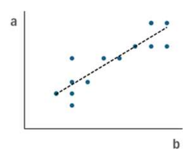
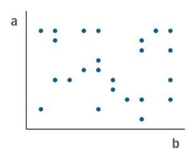
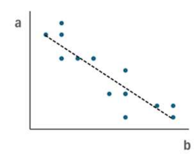
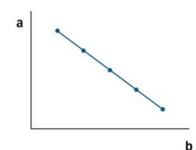
Since the angle of separation between a given vector and itself is zero, $\cos(\theta_{\vec{a},\vec{a}})$, $\cos(\theta_{\vec{b},\vec{b}})$, and $\cos\left(\theta_{\vec{c},\vec{c}}\right)$ all evaluate to 1. Simplifying, we have

$$|\vec{c}|^2 = |\vec{a}|^2 + |\vec{b}|^2 + 2|\vec{a}||\vec{b}| \cos \theta_{\vec{a},\vec{b}} \qquad \text{(Equation 9)}$$

As before, $\theta_{\vec{a},\vec{b}}$ is the supplement of the interior angle opposite side c. Accordingly, we have shown that the formula for the variance of a sum of random variables is an application of the Law of Cosines, which follows from vector addition.

Having shown that σ_1, σ_2, and σ_{total} can be represented as the sides of a triangle, we now justify how Pearson's correlation coefficient, ρ, is equal to the cosine of the angle θ. The angle θ characterizes the alignment between the two uncertainty vectors. Its domain spans from 0° to 180°, allowing the rotating vector to specify the endpoint of every attainable vector sum. An angular separation of 0° means that there is zero directional deviation between the child-level uncertainties, so their sum is maximized. By contrast, 180° of angular separation means that the uncertainty vectors directly oppose one another, so their sum is minimized. Much like θ, the correlation coefficient, ρ, quantifies the strength of the association between two variables. Its maximum value, 1, represents perfect positive correlation, and its minimum value, -1, represents perfect negative correlation. In order to map θ onto ρ, we simply take the cosine of the angle. Read top-to-bottom, Table 2 exemplifies how the conversion of θ into ρ builds a bridge from the geometric representation of uncertainty vectors to the familiar use of ρ to characterize the strength of a linear relationship. Through studying this table, cost estimators can gain a deeper understanding of why accounting for correlations between variables impacts the overall uncertainty on a sum of child elements.

Table 2 - Explicitly connecting uncertainty vectors to the correlation between variables

	Case A	Case B	Case C	Case D	Case E
Uncertainty Vectors & Angle of Rotation					
$\boldsymbol{\theta}$	$0°$	$0° < \theta < 90°$	$90°$	$90° < \theta < 180°$	$180°$
$\boldsymbol{\rho = \cos\theta}$	1	$1 > \rho > 0$	0	$0 > \rho > -1$	-1
Relationship between Variables	Perfect Positive Correlation	Positive Correlation	Independent, No Correlation	Negative Correlation	Perfect Negative Correlation
Plot of Variables					

As demonstrated in the example problems, accounting for a positive correlation between two cost elements—holding all else constant—causes the cumulative uncertainty to increase. Looking at Table 2, modifying the correlation between two cost elements from independent to positively correlated changes the orientation of the child-level standard deviation vectors from Case C to Case B, increasing the magnitude of the cumulative uncertainty vector. Qualitatively, this means that a cost overrun in element X_1 is likely to be coupled with an overrun in X_2, and an underrun in X_1 is likely to be coupled with an underrun in X_2. Therefore, incorporating a positive correlation between the cost elements increases the cumulative uncertainty.

Cost estimators can analyze these formulas in limiting or special cases to evaluate their validity. We can check that the formulas for independent and correlated cases are consistent with one another by applying the Law of Cosines to the special case that $\theta = 90°$. Since the cosine of 90° is zero, the $2ab\cos\theta$ term in Equation 4 vanishes, so the Law of Cosines reduces to the Pythagorean theorem for $\theta = 90°$. Analogously, when $\rho = 0$, all of Equation 3 vanishes except for the summation of standard deviations in quadrature. Therefore, for $\rho = 0$, the formula for the variance of a sum of correlated random variables reduces to Equation 1, the formula for the independent case. Limiting case analysis is a widely recognized tool in other quantitative disciplines because it motivates problem solvers to think critically about equations [5]. Cost estimators can implement limiting case analysis to enhance their understanding of the mathematics of uncertainty analysis.

5. Applications in Automated Cost Estimator Software

When considering a case of dependence, a correlation will exist between two or more variables. This correlation will also be applied to the standard deviation and impact risk output results. While Automated Cost Estimator (ACE) assists in the computation of the risk results and takes correlation into account, it is imperative to understand the application of mathematics being executed.

In this instance, there are two types of correlation that may exist, Functional and Defined. Functional correlation is implicitly built into the ACE model based on the interaction between the variables. For example, if Tire Cost = 4*Rubber Cost, there is a direct positive association of the cost, and it implies Tires and Rubber are positively correlated. Therefore, if Rubber Cost were to increase in variability (have a higher spread/standard deviation), Tire cost would implicitly increase in risk as well. When Monte Carlo risk simulations are run, these relationships are considered without needing to define a "Group" or "Group Strength."

On the other hand, Defined correlation requires input from the analyst. Before an analyst begins to define correlation, ACE advises running the Ri$k Correlation Report, most likely to prevent double counting of risk through added correlation on top of the Functional correlation that already exists. In ACE the correlation between two or more variable is captured by using the "Group" and "Group Strength" columns. The "Group" column denotes which variables are interrelated, and the "Group Strength" column defines their level of correlation or strength of their relationship. The most common input from an analyst uses a heuristic approach, or a subjective assessment. There is typically an understanding of the relationship between two variables, whether positively or negatively correlated, and an approximation of the strength of that relationship. The analyst can utilize the ACE or the U.S. Air Force Cost Risk and Uncertainty Analysis Handbook guidance for input into the group strength column for two dependent variables (see Appendix A, Table 3 and Table 4). Additionally, the USAF Handbook provides insight to applying group strength given more than two variables it states, "In general, all elements measured correlation should be at a minimum plus or minus 0.50 (5 elements), 0.25 (10 elements) or 0.10 (20+ elements) correlation if no information to the contrary is available" [6].

Once grouping and group strength is designated, the Monte Carlo simulation can be run. In the simulation, the Group Strength Algorithm uses the dominant row, series of Latin Hypercube draws, and positional ranking to maintain the original distribution of the task, capture the correlation between the dependent and dominant tasks, and prevent unwanted correlations of dependent tasks within the same group. For additional information, reference the Group Strength Algorithm and Determining Group Draws during Risk Analysis sections of the ACE Help Guide [7].

6. Conclusion

Uncertainty analysis is essential in the field of cost estimation, and estimators can benefit from developing an understanding of the equations that govern how we quantify uncertainty. Although the formulas for determining cumulative uncertainty are somewhat abstract on the surface, we present them in terms of 2D geometric examples to cultivate an understanding for estimators of diverse academic backgrounds.

Child-level standard deviations are vectors that are summed together with respect to their alignment, which can be characterized by the correlation coefficient or the vectors' angular separation. Independent uncertainties have a correlation coefficient of zero, corresponding to an angular separation of 90° between the standard deviation vectors. Therefore, the child-level uncertainties are added in quadrature, which aligns with the Pythagorean theorem. Correlated uncertainties have a nonzero correlation coefficient, so we use a non-right triangle to visualize their summation. The missing side of this triangle can be obtained using the Law of Cosines, and we show that the formula for the variance of a sum of correlated standard deviations follows from vector addition. Connecting our discussion of correlation to its applications in ACE, we show how estimators can implement their understanding of uncertainty analysis to projects in their daily work.

Applications of these formulas on an arbitrary number of child-level cost elements may be impossible to visualize; however, the mathematics used in two dimensions generalizes to higher dimensional cases, and understanding 2D examples makes these abstract problems more attainable. Overall, we hope that this geometric approach to statistics lowers the barrier to entry for estimators striving to understand uncertainty analysis and higher mathematics for cost estimation.

Appendix A

In this paper, the chosen sign convention for the Law of Cosines is different from the one that is typically taught in high school geometry. The version of the Law of Cosines that may be more familiar is

$$c^2 = a^2 + b^2 - 2ab\cos\gamma \qquad \text{(Equation 10)}$$

Where a, b, c are the side lengths of a triangle, and γ is the interior angle opposite side c. In the context of a geometry class, it is convenient to define γ as an interior angle so that the student need not look for a supplementary angle *outside* the triangle to solve for a missing side.

Equation 10 derives from vector subtraction. We can visualize vector subtraction by making a slight modification to Figure 5. Simply flip the direction of $\vec{a}$ so that the vector points to the left, causing $\vec{a}$ and $\vec{b}$ to share the same endpoint. In this case, $\vec{c}$ is defined as the difference $\vec{a} - \vec{b}$, yet none of the side lengths or angle measures have changed. Accordingly, we can derive the vector version of Equation 10 as follows:

$$\vec{c}\cdot\vec{c} = (\vec{a} - \vec{b})\cdot(\vec{a} - \vec{b}) \qquad \text{(Equation 11)}$$

$$|\vec{c}||\vec{c}|\cos(\theta_{\vec{c},\vec{c}}) = |\vec{a}||\vec{a}|\cos(\theta_{\vec{a},\vec{a}}) + |\vec{b}||\vec{b}|\cos(\theta_{\vec{b},\vec{b}}) - 2|\vec{a}||\vec{b}|\cos\gamma \qquad \text{(Equation 12)}$$

$$|\vec{c}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2|\vec{a}||\vec{b}|\cos\gamma \qquad \text{(Equation 13)}$$

Where γ is the interior angle opposite side $\vec{c}$.

When making geometric sense of the formula for the variance of a sum of correlated random variables, it is convenient to adopt the sign convention shown in Equation 4. This way, the derivation from vector addition aligns with the addition of standard deviations, and the angle θ represents deviation from perfect positive alignment.

Appendix B

Table 3 - ACE Help Guide Correlation Guidance

Strength	Positive	Negative
None	0.0	0.0
Weak	0.3	-0.3
Medium	0.5	-0.5
Strong	0.9	-0.9
Perfect	1.0	-1.0

Table 4 - USAF Cost, Risk, and Uncertainty Analysis Handbook Correlation Guidance

Strength	Positive	Negative
None	0.00	0.00
Weak	0.25	-0.25
Medium	0.50	-0.50
Strong	0.90	-0.90
Perfect	1.00	-1.00

References

[1] U.S. Government Accountability Office, GAO-20-195G Cost Estimating and Assessment Guide, Washington, DC: U.S. Government Accountability Office, p. 140, 2020.

[2] S. Koellner, "Spread Too Thin: Managing Coefficient of Variation in Monte-Carlo Based Cost Models," ICEAA 2023 Professional Development & Training Workshop, p. 29, 2023.

[3] G. Vernez, H. G. Massey, “The Acquisition Cost-Estimating Workforce,” RAND Corporation, p. 9, 2009.

[4] P. R. Garvey, *Probability Methods for Cost Uncertainty Analysis*, Bedford, MA: Marcel Dekker, Inc., p. 182, 2000.

[5] G. White, T. Sikorski, J. Landay, M. Ahmed, “Limiting Case Analysis in an Electricity and Magnetism Course,” American Physical Society, 2023.

[6] Tecolote Research, Inc., U. S. Air Force Cost Risk and Uncertainty Analysis Handbook, p. 23, 2007.

[7] Tecolote Research, Inc., ACE 8.2 Help Guide, 2024.