<INCOMPLETE>
The Skitovich-Darmois Theorem

Or: "Why Rotated SVI Can Only Use Normally Distributed Variables"

dmerrell@cs.wisc.edu

Integrating Probabilities with SVI

Symbolic Volume Integration (SVI) of a joint distribution function $f(x_1, \ldots, x_n)$ relies on the independence of its random variables; i.e. that

$$f(\underline{x}) = f_1(x_1) \cdots f_n(x_n).$$

where each f_i is the univariate PDF of the i^{th} random variable.

When the region of integration is decomposed into axes-aligned hyper-rectangles, integration over each hyper-rectangle h reduces to an evaluation of the variables' CDFs:

$$\int_h f(\underline{x}) d\underline{x} \tag{1}$$

$$= \int_{l_1}^{u_1} \cdots \int_{l_n}^{u_n} f_1(x_1) \cdots f_n(x_n) d\underline{x} \tag{2}$$

$$= \left(\int_{l_1}^{u_1} f_1(x_1) dx_1 \right) \cdots \left(\int_{l_n}^{u_n} f_n(x_n) dx_n \right) \tag{3}$$

$$= (F_1(u_1) - F_1(l_1)) \cdots (F_n(u_n) - F_n(l_n)) \tag{4}$$

where $l_i \leq u_i$ are the bounds of the hyperrectangle in each dimension and the F_i are the cumulative density functions corresponding to the f_i.

Note that in order for an integral to be compatible for SVI (as above), it must satisfy two requirements:

1. The region of integration must be "aligned" with the axes; i.e. none of the iterated integrals have bounds dependent on the variables of integration.

2. The integrand must be separable; i.e. it must be the joint PDF of independent random variables, and hence the product of n univariate PDFs.

Suppose, however, that we permit hyper-rectangles that are not aligned with the axes. This corresponds to a substitution of variables in the integral[1]. Specifically, if an axes-aligned hyperrectangle h is rotated with respect to the axes by a matrix R, then its integral is

$$\int_{Rh} f(\underline{x})d\underline{x} \tag{5}$$

$$= \int_{h} f(R\underline{y})|\det R|d\underline{y} \tag{6}$$

where Rh denotes the region obtained by transforming every point in h by R and r_i^T are the rows of matrix R. Note that none of these integrals are compatible for SVI according to the above criteria; integral (5) generally does not satisfy criterion 1 and integral (6) generally does not satisfy criterion 2. Specifically, Rh will not generally have bounds that are independent of the integration variables, and we cannot generally rewrite the integrand $f_1(r_1^T\underline{y})\cdots f_n(r_n^T\underline{y})$ as a product of functions of the individual y_i.

While integrals over rotated hyper-rectangles do not generally satisfy the SVI-compatibility criteria, we will identify sufficient (and necessary!) conditions for SVI to handle them.

A Sufficient Condition: All Random Variables are Gaussian

Suppose all of the free variables X_i associated with our program are probabilistically assigned to gaussian distributions having means μ_i and variances σ_i. Within the program, we can always "shift and scale" each of them, obtaining $X_i' = (X_i - \mu_i)/\sigma_i$ with zero mean and unit variance. Hence, without loss of generality, we can assume all gaussian free variables X_i have zero mean and unit variance.

[1]See the appendix "Review of Integration by Substitution of Variables" for a quick refresher on this topic.

Since the X_i are independent, their joint PDF is given by the product of their individual PDFs.

$$f(\underline{x}) = \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_1^2}{2}\right)\right] \cdots \left[\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_n^2}{2}\right)\right] \tag{7}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{\underline{x}^\mathsf{T}\underline{x}}{2}\right) \tag{8}$$

If we integrate this joint PDF over a transformed hyper-rectangle Rh, we get

$$\int_{Rh} \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left(-\frac{\underline{x}^\mathsf{T}\underline{x}}{2}\right) d\underline{x} \tag{9}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \int_h \exp\left(-\frac{(R\underline{u})^\mathsf{T}(R\underline{y})}{2}\right) |\det R| d\underline{y} \tag{10}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \int_h \exp\left(-\frac{\underline{y}^\mathsf{T}(R^\mathsf{T}R)\underline{y}}{2}\right) |\det R| d\underline{y} \tag{11}$$

where a substitution of variables yields integral (10) from integral (9).

At this point, we provide the additional stipulation that R be a *scaled rotation*; that is, R cannot be simply *any* invertible matrix, but must be of the form $R = c{\cdot}R'$, where $c > 0$ is a scalar constant and R' is a rotation matrix. This stipulation is consistent with our original goal—the construction of an SVI method that permits rotated hyperrectangles—and has several useful implications:

- Since the determinant of a rotation matrix is 1, it follows that $\det R = \det(c \cdot R') = c^n$;
- Since the inverse of a rotation matrix is its transpose, we have $R^\mathsf{T}R = (c \cdot R')^\mathsf{T}(c \cdot R') = c^2$;
- Permitting a scaled rotation allows us to avoid matrices containing irrational entries; this is important for SVI, as it relies on an SMT solver that does not generally allow irrational numbers.

Hence, continuing from integral (11) we have

$$= \left(\frac{1}{\sqrt{2\pi}} \right)^n |c^n| \cdot \int_h \exp\left(-\frac{\underline{y}^T \underline{y}}{2c^{-2}} \right) d\underline{y} \tag{12}$$

$$= \left(\frac{1}{\sqrt{2\pi c^{-2}}} \right)^n \cdot \int_h \exp\left(-\frac{\underline{y}^T \underline{y}}{2c^{-2}} \right) d\underline{y} \tag{13}$$

which is an integral satisfying both of the SVI compatibility criteria; its bounds are constant, and its integrand is in fact a product of n PDFs—in this case, the product of n gaussian PDFs with $\mu = 0$ and $\sigma^2 = c^{-2}$.

The point of this section was to illustrate that the following are *sufficient* conditions for rotated SVI to work:

- That the rotation affect spaces spanned only by gaussian-distributed random variables. Without loss of generality, these must have zero mean and unit variance.
- That the rotation is, in some strict sense, a rotation. Specifically, we require them to be scaled rotations.

In the upcoming sections, we will present the Skitovich-Darmois Theorem and show that it provides a *necessary* condition for rotated SVI to work.

The Skitovich-Darmois Theorem

This is the classical Skitovich-Darmois Theorem:

Thm: Given independent random variables $X_1, \ldots, X_n$, and nonzero α_i, β_i $i \in \{1, \ldots, n\}$ such that

$$\sum_{i=1}^{n} \alpha_i X_i \tag{14}$$

$$\text{and } \sum_{i=1}^{n} \beta_i X_i \tag{15}$$

are independent of each other, it follows that the X_i must be normally distributed.

In other words: if linear combinations of independent variables are, themselves, independent, then the original variables must be normally distributed.

This can be taken as (yet another) definition of the normal distribution. I would describe this as "spooky".

In the next section, we will use this theorem to provide a necessary condition for rotated SVI to work.

A Necessary Condition: All Random Variables are Gaussian

Suppose we have independent random variables $Z_1, \dots, Z_n$, making no assumption about their individual distributions. Let their joint PDF be denoted $f(\underline{z}) = f_1(z_1) \cdots f_n(z_n)$ and let R be a scaled rotation matrix. We define the random variables $Y_1, \dots, Y_n$ by the mapping

$$\underline{Y} = R(\underline{Z})$$

It follows (from substitution of variables) that $Y_1, \dots, Y_n$ have the following joint PDF:

$$g(\underline{y}) = f(R^{-1}(\underline{y}))|\det R| \tag{16}$$

In short, if some "new" random variables are defined to be a scaled rotation of "old" random variables, then their joint PDF is of the form found in (16).

We now consider the converse. Given a function of the form found in (16), is it guaranteed to be the joint PDF of some random variables $Y_1, \dots, Y_n$? It is a joint PDF of n random variables if and only if it is nonnegative over $\mathbb{R}^n$ and its integral over $\mathbb{R}^n$ evaluates to 1. Since f is itself a PDF and is multiplied by an absolute value, we know that (16) is nonnegative everywhere. When we integrate (16) over any region Γ in $\mathbb{R}^n$, we

We have now established that the integrand of (6) is, in fact, the joint PDF of random variables Y_i given by $\underline{Y} = R^{-1}\underline{X}$. So each Y_i is a linear combination of the "original" random variables $Z_1, \dots X_n$. In order for integral (6) to be SVI-compatible, it is necessary for the $Y_1, \dots, Y_n$ to be stochastically independent. We cannot generally assume that the linear combinations have coefficients equal to zero; hence,