INTERVENTION IN ORNSTEIN-UHLENBECK SDES

ALEXANDER SOKOL

Abstract. We introduce a notion of intervention for stochastic differential equations and a corresponding causal interpretation. For the case of the Ornstein-Uhlenbeck SDE, we show that the SDE resulting from a simple type of intervention again is an Ornstein-Uhlenbeck SDE. We discuss criteria for the existence of a stationary distribution for the solution to the intervened SDE. We illustrate the effect of interventions by calculating the mean and variance in the stationary distribution of an intervened process in a particularly simple case.

1. Introduction

Causal inference for continuous-time processes is a field in ongoing development. Similar to causal inference for graphical models, see [9], one of the primary objectives for causal inference for continuous-time processes is to identify the effect of an intervention given assumptions on the distribution and causal structure of the observed continuous-time process.

Several flavours of causal inference are available for continuous-time processes, see for example [3, 4, 10]. In this paper, we outline a causal interpretation of stochastic differential equations and a corresponding notion of intervention, we calculate the distribution of an intervened Ornstein-Uhlenbeck SDE, and we calculate analytical expressions for the mean and variance of the stationary distribution of the resulting process for particular examples of interventions.

2. Causal interpretation of stochastic differential equations

Consider a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)\) satisfying the usual conditions, see [11] for the definition of this and other notions related to continuous-time stochastic processes. Let \(Z\) be a \(d\)-dimensional semimartingale and assume that

2010 Mathematics Subject Classification. Primary 60G15.
Key words and phrases. Causality, Intervention, SDE, Ornstein-Uhlenbeck process, Stationary distribution.
$a : \mathbb{R}^p \to \mathcal{M}(p, d)$ is a Lipschitz mapping, where $\mathcal{M}(p, d)$ denotes the space of real $p \times d$ matrices. Consider the stochastic differential equation (SDE)

$$X^i_t = x^i_0 + \sum_{j=1}^{d} \int_0^t a_{ij}(X^j_s) \, dZ^j_s, \quad i \leq p.$$  

(2.1)

By the Lipschitz property of $a$, it holds by Theorem V.7 of [11] that there exists a pathwisely unique solution to (2.1). The following definition yields a causal interpretation of (2.1) based on simple substitution and inspired by ideas outlined in Section 4.1 of [1].

**Definition 2.1.** Consider some $m \leq p$ and $c \in \mathbb{R}$. The $(p - 1)$-dimensional intervened SDE arising from the intervention $X^m := c$ is defined to be

$$U^i_t = x^i_0 + \sum_{j=1}^{d} \int_0^t b_{ij}(U^j_s) \, dZ^j_s$$

for $i \leq p$ with $i \neq m$,

(2.2)

where $b_{ij}(y_1, \ldots, y_{m-1}, y_{m+1}, \ldots, y_p) = a_{ij}(y_1, \ldots, c, \ldots, y_p)$, and the $c$ is on the $m$th coordinate. Letting $U$ be the unique solution to the SDE and defining $Y$ by putting $Y = (U^1, \ldots, U^{m-1}, c, U^{m+1}, \ldots, U^p)$, we refer to $Y$ as the intervened process and write $(X|X^m := c)$ for $Y$.

By Theorem V.16 and Theorem V.5 of [11], the solutions to both (2.1) and (2.2) may be approximated by the Euler schemes for their respective SDEs. Making these approximations and applying Pearl’s notion of intervention in an appropriate sense, see [9], we may interpret Definition 2.1 as intervening in the system (2.1) under the assumption that the driving semimartingales $Z^1, \ldots, Z^d$ are noise processes unaffected by interventions, while the processes $X^1, \ldots, X^p$ are affected by interventions. Note that the operation of making an intervention takes a $p$-dimensional SDE as its input and yields a $(p - 1)$-dimensional SDE as its output, and this operation is crucially dependent on the coefficients in the SDE: These coefficients in a sense correspond to the directed acyclic graphs of [9]. A major benefit of causality in systems such as (2.1) as compared to the theory of [9] is the ability to capture feedback systems and interventions in such feedback systems.

As the solutions to (2.1) and (2.2) are defined on the same probability space, we may even consider the process $Y - X$, where $Y = (X|X^m := c)$, allowing us to calculate for example the variance of the effect of the intervention. As $Y$ and $X$ are never observed simultaneously in practice, however, we will concentrate on analyzing the differences between the laws of $Y$ and $X$ separately.

### 3. Intervention in Ornstein-Uhlenbeck SDEs

Recall that for an $\mathcal{F}_0$ measurable variable $X_0$ and for $A \in \mathbb{R}^p$, $B \in \mathcal{M}(p, p)$ and $\sigma \in \mathcal{M}(p, d)$, the Ornstein-Uhlenbeck SDE with initial value $X_0$, mean reversion
level \( A \), mean reversion speed \( B \), diffusion matrix \( \sigma \) and \( d \)-dimensional driving noise is

\[
X_t = X_0 + \int_0^t B(X_s - A) \, ds + \sigma W_t,
\]

where \( W \) is a \( d \)-dimensional \((\mathcal{F}_t)\) Brownian motion, see Section II.72 of [12]. The unique solution to this equation is

\[
X_t = \exp(tB) \left( X_0 - \int_0^t \exp(-sB)BA \, ds + \int_0^t \exp(-sB)\sigma \, dW_s \right)
\]

where the matrix exponential is defined by \( \exp(A) = \sum_{n=0}^{\infty} A^n / n! \). This is a Gaussian homogeneous Markov process with continuous sample paths. The following lemma shows that making an intervention in an Ornstein-Uhlenbeck SDE yields an SDE whose nontrivial coordinates solve another Ornstein-Uhlenbeck SDE.

**Lemma 3.1.** Consider the Ornstein-Uhlenbeck SDE (3.1) with initial value \( x_0 \). Fix \( m \leq p \) and \( c \in \mathbb{R} \), and let \( X \) be the unique solution to (3.1). Furthermore, let \( Y = (X|X^m := c) \) and let \( Y^{-m} \) be the \( p-1 \) dimensional process obtained by removing the \( m \)'th coordinate from \( Y \). Let \( \tilde{B} \) be the submatrix of \( B \) obtained by removing the \( m \)'th row and column of \( B \), and assume that \( \tilde{B} \) is invertible. Then \( Y^{-m} \) solves

\[
Y_t^{-m} = y_0 + \int_0^t \tilde{B}(Y_s^{-m} - \tilde{A}) \, ds + \tilde{\sigma} W_t,
\]

where \( y_0 \) is obtained by removing the \( m \)'th coordinate from \( x_0 \). \( \tilde{\sigma} \) is obtained by removing the \( m \)'th row of \( \sigma \) and \( \tilde{A} = \alpha - \tilde{B}^{-1} \beta \), where \( \alpha \) and \( \beta \) are obtained by removing the \( m \)'th coordinate from \( A \) and from the vector whose \( i \)'th component is \( b_{im}(c - a_m) \), respectively, where \( b_{im} \) is the entry corresponding to the \( i \)'th row and the \( m \)'th column of \( B \), and \( a_m \) is the \( m \)'th element of \( A \).

**Proof.** By Definition 2.1 we have

\[
Y_t^i = y_0 + \int_0^t b_{im}(c - a_m) + \sum_{j \neq m} b_{ij}(Y_s^j - a_j) \, ds + \sum_{j=1}^p \sigma_{ij} W_t^j
\]

for \( i \neq m \). Note that for any vector \( y \), the system of equations in \( \tilde{\sigma} \)

\[
b_{im}(c - a_m) + \sum_{j \neq m} b_{ij}(y_j - a_j) = \sum_{j \neq m} b_{ij}(y_j - \tilde{a}_j) \text{ for } i \neq m,
\]

is equivalent to the system of equations

\[
\sum_{j \neq m} b_{ij} \tilde{a}_j = \left( \sum_{j \neq m} b_{ij}a_j \right) - b_{im}(c - a_m) \text{ for } i \neq m.
\]

Since we have assumed \( \tilde{B} \) to be invertible, this system of equations has the unique solution \( \tilde{A} = \tilde{B}^{-1}(\tilde{B}\alpha - \beta) = \alpha - \tilde{B}^{-1}\beta \). For \( i \neq m \), we therefore obtain that \( Y_t^i = y_0 + \int_0^t \sum_{j \neq m} b_{ij}(Y_s^j - \tilde{a}_j) \, ds + \sum_{j=1}^p \sigma_{ij} W_t^j \), proving the result. \( \square \)
Recall that a principal submatrix of a matrix is a submatrix with the same rows and columns removed. In words, Lemma 3.1 states that if a particular principal submatrix $\tilde{B}$ of the mean reversion speed is invertible, then making the intervention $X^{m} := c$ in an Ornstein-Uhlenbeck SDE results in a new Ornstein-Uhlenbeck SDE with mean reversion speed $\tilde{B}$ and modified mean reversion level involving the inverse of $\tilde{B}$. Now assume that an Ornstein-Uhlenbeck SDE is given such that the solution has a stationary initial distribution. A natural question to ask is what interventions will yield intervened processes where stationary initial distributions also exist. In the following, we consider this question.

Recall that a square matrix is called stable if its eigenvalues have negative real parts and semistable if its eigenvalues have nonpositive real parts, see [2]. Theorem 4.1 of [13] yields necessary and sufficient criteria for the existence of a stationary probability measure for the solution of (3.1). One criterion is expressed in terms of the controllability subspace of of the matrix pair $(B, \sigma)$, which is the span of the columns in the matrices $\sigma, B\sigma, \ldots, B^{p-1}\sigma$. In the case where $\sigma$ has full column span, meaning that the columns of $\sigma$ span all of $\mathbb{R}^p$, the controllability subspace is all of $\mathbb{R}^p$, and Theorem 4.1 of [13] shows that the existence of a stationary probability measure is equivalent to $B$ being stable. The case where $\sigma$ is not required to have full column span is more involved.

In the following, we will restrict our attention to Ornstein-Uhlenbeck processes with $\sigma$ having full column span. By Theorem 4.1 of [13], it then holds that there exists a stationary distribution if and only if $B$ is stable. Furthermore, applying Theorem 2.4 and Theorem 2.12 of [8], it holds in the affirmative case that the stationary distribution is the normal distribution with mean $\mu$ and variance $\Gamma$ solving $B\mu = BA$ and $\sigma\sigma^t + B\Gamma + \Gamma B^t = 0$. Note that as $B$ is stable, zero is not an eigenvalue of $B$, thus $B$ is invertible and $\mu = A$. Also, stability of $B$ yields that $\Gamma = \int_0^{\infty} e^{sB}\sigma\sigma^te^{sB^t}ds$. For the $(p-1)$-dimensional Ornstein-Uhlenbeck process resulting from an intervention according to Lemma 3.1 the diffusion matrix $\tilde{\sigma}$ is obtained by removing the $m$’th row of $\sigma$. As the columns of $\sigma$ span $\mathbb{R}^p$, the columns of $\tilde{\sigma}$ span $\mathbb{R}^{p-1}$. Therefore, it also holds for the intervened process that there exists a stationary distribution if and only if the mean reversion speed is stable.

We conclude that for diffusion matrices with full column span, the existence of stationary distributions for both the original and the intervened SDE is determined solely by stability of the mean reversion speed matrix $B$ and the corresponding principal submatrices.

Consider a stable matrix $B$. It then holds that if all principal submatrices of $B$ are stable, all interventions will preserve stability of the system. We are thus lead to the question of when a principal submatrix of a matrix is stable. That stability does not in general lead to stability of principal submatrices may be seen from the following example. Define $B$ by putting

$$B = \begin{bmatrix} 1 & 7 \\ -1 & -3 \end{bmatrix}.$$
The matrix $B$ has eigenvalues $-1 \pm i \sqrt{3}$ and is thus stable, while the principal submatrix obtained by removing the second row and second column trivially has the single eigenvalue 1 and thus is not stable, in fact not even semistable. Conversely, $-B$ has eigenvalues $1 \pm i \sqrt{3}$ and thus is neither stable nor semistable, while the principal submatrix obtained by removing the second row and second column of $-B$ is stable.

There are classes of matrices satisfying that all principal submatrices are stable. For example, by the inclusion principle for symmetric matrices, see Theorem 4.3.15 of [7], it follows that a principal submatrix of any symmetric stable matrix again is stable. In general, though, it is difficult to ensure that all principal submatrices are stable. However, there are criteria ensuring that all principal submatrices are semistable. For example, Lemma 2.4 of [5] shows that if $B$ is stable and sign symmetric, then all principal submatrices of $B$ are semistable. Here, sign symmetry is a somewhat involved matrix criterion, it does however hold that any stable symmetric matrix also is sign symmetric. Furthermore, by Theorem 1 of [2], either of the follow three properties are also sufficient for having all principal submatrices being semistable:

1. $A - D$ is stable for all nonnegative diagonal $D$.
2. $DA$ is stable for all positive diagonal $D$.
3. There is positive diagonal $D$ such that $AD + DA^t$ is negative definite.

4. An example of a particular intervention

Consider now a three-dimensional Ornstein-Uhlenbeck process $X$ with $\sigma$ being the identity matrix of order three and upper diagonal mean reversion speed matrix $B$, and assume that the diagonal elements of $B$ all are negative. As the diagonal elements of $B$ in this case also are the eigenvalues, $B$ is then stable, and all principal submatrices are stable as well. The interpretation of having $B$ upper diagonal is that the levels of both $X^1$, $X^2$ and $X^3$ influence the average change in $X^1$, while only the levels of $X^2$ and $X^3$ influence the average change in $X^2$ and only $X^3$ influences the average change in $X^3$. Figure 4.1 illustrates this, as well as the changes to the dependence structure obtained by making interventions $X^2 := c$ or $X^3 := c$.

We will investigate the details of what happens to the system when making the intervention $X^2 := c$ or $X^3 := c$. To this end, we calculate the mean and variance in the stationary distribution for the nontrivial coordinates in each of the intervened processes. Consider first the case of the intervention $X^2 := c$. Let $\mu$ and $\Gamma$ denote the mean and variance in the stationary distribution after intervention. Applying Lemma 3.1, the SDE resulting from making this intervention is a two-dimensional
Figure 4.1. Graphical illustrations of the dependence structures of \((X_1, X_2, X_3)\) (left), of the dependence when making the intervention \(X_2 := c\) (middle) and of the dependence when making the intervention \(X_3 := c\) (right).

Ornstein-Uhlenbeck SDE with mean reversion speed and mean reversion level

\[
\begin{bmatrix}
  b_{11} & b_{13} \\
  0 & b_{33}
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
  a_1 \\
  a_3
\end{bmatrix}
- \begin{bmatrix}
  b_{11} & b_{13} \\
  0 & b_{33}
\end{bmatrix}^{-1}
\begin{bmatrix}
  b_{12}(c - a_2) \\
  0
\end{bmatrix}.
\]

As we have

\[
\begin{bmatrix}
  b_{11} & b_{13} \\
  0 & b_{33}
\end{bmatrix}^{-1}
= \begin{bmatrix}
  \frac{1}{b_{11}} & -\frac{b_{13}}{b_{11}b_{33}} \\
  0 & \frac{1}{b_{33}}
\end{bmatrix},
\]

this immediately yields that

\[
\mu = \begin{bmatrix}
  a_1 - \frac{b_{13}}{b_{11}}(c - a_2)
\end{bmatrix}.
\]

As for the variance, recall that we have the representation

\[
\begin{aligned}
\Gamma &= \int_0^\infty \exp \left( s \begin{bmatrix}
  b_{11} & b_{13} \\
  0 & b_{33}
\end{bmatrix} \right) \exp \left( s \begin{bmatrix}
  b_{11} & 0 \\
  b_{13} & b_{33}
\end{bmatrix} \right) \, ds.
\end{aligned}
\]

In order to calculate this integral, first consider the case \(b_{11} = b_{33}\). By Theorem 4.11 of [6], we in this case obtain

\[
\begin{bmatrix}
  b_{11} & b_{13} \\
  0 & b_{33}
\end{bmatrix}^{-1}
= \begin{bmatrix}
  1 & \frac{b_{12}}{b_{11}} \\
  0 & \frac{1}{b_{33}}
\end{bmatrix},
\]

and similarly for the transpose. Applying that \(\int_0^\infty x^\alpha e^{\beta x} \, dx = \Gamma(\alpha + 1)/(-\beta)^{\alpha + 1}\) for all \(\alpha > -1\) and \(\beta < 0\), we conclude

\[
\begin{aligned}
\Gamma &= \int_0^\infty e^{2sb_{11}} \begin{bmatrix}
  1 & sb_{13} \\
  0 & 1
\end{bmatrix} \begin{bmatrix}
  1 & 0 \\
  sb_{13} & 1
\end{bmatrix} \, ds \\
&= \int_0^\infty e^{2sb_{11}} \begin{bmatrix}
  1 & s^2b_{13} \\
  0 & sb_{13}
\end{bmatrix} \begin{bmatrix}
  1 & 0 \\
  sb_{13} & 1
\end{bmatrix} \, ds = \begin{bmatrix}
  -\frac{1}{2b_{11}} & -\frac{b_{13}}{4b_{11}^2} \\
  \frac{b_{12}}{4b_{11}^2} & -\frac{1}{2b_{11}}
\end{bmatrix}.
\end{aligned}
\]

In the case \(b_{11} \neq b_{33}\), we put \(\zeta = b_{13}/(b_{11} - b_{33})\) and Theorem 4.11 of [6] yields

\[
\begin{aligned}
\exp \left( s \begin{bmatrix}
  b_{11} & b_{13} \\
  0 & b_{33}
\end{bmatrix} \right) &= \begin{bmatrix}
  e^{sb_{11}} & \zeta(e^{sb_{11}} - e^{sb_{33}}) \\
  0 & e^{sb_{33}}
\end{bmatrix}.
\end{aligned}
\]
and we then obtain
\[
\exp \left( s \begin{bmatrix} b_{11} & b_{13} \\ 0 & b_{33} \end{bmatrix} \right) \exp \left( s \begin{bmatrix} b_{11} & 0 \\ b_{13} & b_{33} \end{bmatrix} \right) = \begin{bmatrix} e^{sb_{11}} \zeta(e^{sb_{11}} - e^{sb_{33}}) & e^{sb_{11}} \\ e^{sb_{33}} & e^{sb_{11}} \end{bmatrix} \begin{bmatrix} e^{sb_{11}} & 0 \\ \zeta(e^{sb_{11}} - e^{sb_{33}}) & e^{sb_{33}} \end{bmatrix}^{-1} \begin{bmatrix} 1 + \zeta^2 e^{2sb_{11}} - 2\zeta^2 e^{(b_{11} + b_{33})} + \zeta^2 e^{2sb_{33}} & \zeta e^{(b_{11} + b_{33})} - \zeta e^{2sb_{33}} \\ \zeta e^{(b_{11} + b_{33})} - \zeta e^{2sb_{33}} & e^{2sb_{33}} \end{bmatrix}^{-1},
\]
(4.8) implying that
\[
\Gamma = \begin{bmatrix} -\frac{(1+\zeta^2)}{2b_{11}} + \frac{2\zeta^2}{b_{11}+b_{33}} - \frac{\zeta^2}{b_{11}+b_{33}} + \frac{\zeta}{b_{33}} - \frac{1}{b_{33}} \\ -\frac{1}{b_{33}} - \zeta^2 \frac{\zeta(b_{11}+b_{33})}{2b_{33}(b_{11}+b_{33})} - \frac{1}{2b_{33}} \end{bmatrix}.
\]
(4.9)

Note in particular that (4.9) also yields the correct result in the case $b_{11} = b_{33}$.

Next, considering the intervention $X^3 := c$, we let $\nu$ and $\Sigma$ denote the mean and variance in the stationary distribution of the nontrivial coordinates after intervention. By Lemma 3.1, the result of making this intervention is an Ornstein-Uhlenbeck SDE with mean reversion speed and mean reversion level
\[
\nu = \begin{bmatrix} a_1 - \frac{b_{13}}{b_{11}} - \frac{b_{13}b_{23}}{b_{11}b_{22}} (c - a_3) \\ a_2 - \frac{b_{13}b_{23}}{b_{22}} (c - a_3) \end{bmatrix},
\]
(4.10)
yielding by calculations similar to the previous case that
\[
\Sigma = \begin{bmatrix} -\frac{1}{2b_{11}} - \frac{b_{13}^2}{b_{11}b_{22}} & \frac{b_{13}}{b_{22}}(c - a_3) \\ -\frac{b_{13}b_{23}}{b_{22}(b_{11} + b_{22})} & \frac{b_{13}b_{23}}{b_{22}(b_{11} + b_{22})} \end{bmatrix}.
\]
(4.12)

We have now calculated the mean and variance in the stationary distribution for both intervened processes. We next take a moment to interpret our results.

In the original system, all of $X^1$, $X^2$ and $X^3$ negatively influenced themselves, and in addition to this, $X^2$ influenced $X^1$ and $X^3$ influenced $X^1$ both directly and through its influence on $X^2$. Based on this, we would expect that making the intervention $X^2 := c$, the steady state of $X^3$ would not be changed, while the steady state of $X^1$ would change, depending on the level of influence $b_{13}$ of $X^2$ on $X^1$. This is what we see in (4.3). When making the intervention $X^3 := c$, however, we obtain a change in the steady state of $X^1$ based both on the direct influence of $X^3$ on $X^1$, depending on $b_{13}$, but also on the indirect influence of $X^3$
on $X^1$ through $X^2$, depending also on $b_{23}$ and $b_{12}$. Furthermore, the steady state of $X^2$ also changes. These results show themselves in (4.11).

As for the steady state variance, the changes resulting from interventions are in both cases of the same type, yielding moderately complicated analytical expressions, both independent of $c$. This implies that while we in most cases will be able to obtain any steady state mean for, say, $X^1$, by picking $c$ suitably, the steady state variance can be influenced only by the type of intervention made, that is, on which parts of the system the interventions are made. Furthermore, by considering explicit formulas for the steady state variance in the original system, it may be seen that for example positive covariances may turn negative and vice versa when making interventions.

Acknowledgements. The development of the notion of intervention for SDEs is joint work with my thesis advisor, Niels Richard Hansen, whom I also thank for valuable discussions and advice.

References

1. Odd O. Aalen, Kjetil Røysland, Jon Michael Gran, and Bruno Ledergerber, Causality, mediation and time: a dynamic viewpoint, J. Roy. Statist. Soc. Ser. A 175 (2012), no. 4, 831–861.
2. G. W. Cross, Three types of matrix stability, Linear Algebra and Appl. 20 (1978), no. 3, 253–263.
3. Jean-Pierre Florens and Denis Fougere, Noncausality in continuous time, Econometrica 64 (1996), no. 5, 1195–1212.
4. Anne Gégout-Petit and Daniel Commenges, A general definition of influence between stochastic processes, Lifetime Data Anal. 16 (2010), no. 1, 33–44.
5. Daniel Hershkowitz and Nathan Keller, Positivity of principal minors, sign symmetry and stability, Linear Algebra Appl. 364 (2003), 105–124.
6. Nicholas J. Higham, Functions of matrices, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008, Theory and computation.
7. Roger A. Horn and Charles R. Johnson, Matrix analysis, Cambridge University Press, Cambridge, 1985.
8. M. Jacobsen, A brief account of the theory of homogeneous gaussian diffusions in finite dimensions, Frontiers in Pure and Applied Probability 1, TVP Science Publishers, 1993, pp. 86–94.
9. Judea Pearl, Causality, second ed., Cambridge University Press, Cambridge, 2009, Models, reasoning, and inference.
10. Ljiljana Petrić and Sladjana Dimitrijević, Invariance of statistical causality under convergence, Statist. Probab. Lett. 81 (2011), no. 9, 1445–1448.
11. Philip E. Protter, Stochastic integration and differential equations, Stochastic Modelling and Applied Probability, vol. 21, Springer-Verlag, Berlin, 2005, Second edition. Version 2.1, Corrected third printing.
12. L. C. G. Rogers and David Williams, Diffusions, Markov processes, and martingales. Vol. 1, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Foundations, Reprint of the second (1994) edition.
13. Moshe Zakai and Jakov Snyders, Stationary probability measures for linear differential equations driven by white noise, J. Differential Equations 8 (1970), 27–33.
Alexander Sokol: Institute of Mathematics, University of Copenhagen, 2100 Copenhagen, Denmark

E-mail address: alexander@math.ku.dk

URL: http://www.math.ku.dk/~alexander