AIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202

Centimeter

Inches

MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

1 of 1

RECEIVED

SEP 19 1994

OSTI

Analytic Perturbation Analysis of Discrete Event Dynamic Systems¹

Stanislav Uryasev
Brookhaven National Laboratory,
Bldg.130, Upton, NY, 11973

Abstract

This paper considers a new Analytic Perturbation Analysis (APA) approach for Discrete Event Dynamic Systems (DEDS) with discontinuous sample-path functions with respect to control parameters. The performance functions for DEDS usually are formulated as mathematical expectations, which can be calculated only numerically. APA is based on new analytic formulas for the gradients of expectations of indicator functions; therefore, it is called an analytic perturbation analysis. The gradient of performance function may not coincide with the expectation of a gradient of sample-path function (i.e., the interchange formula for the gradient and expectation sign may not be valid). Estimates of gradients can be obtained with one simulation run of the models.

1 Introduction

Discrete Event Dynamic Systems (DEDS) are an important class of dynamic systems with discontinuous (discrete) stochastic behavior. The dynamics of such systems include logical rules, which discontinuously (discretely) change the state of the system. For example, a customer can change a routing if a server is busy in a queuing network.

In analyzing of such systems, one often is interested in how their performance depends upon these parameters. Gradients provide useful information on this dependence. Performance functions for the DEDS usually are formulated as mathematical expectations of the discontinuous functions with respect to control parameters. Gradients of performance functions are difficult to calculate analytically. The gradient of a smooth function can be estimated with finite differences, but, for DEDS, these estimates have high variances. Low-variance stochastic

estimates of the gradients (*stochastic quasigradients*) can be obtained for a broad class of discrete event systems with one simulation run. The *infinitesimal perturbation analysis (IPA)* proposed by Ho and co-workers [8] is one such efficient method, which can be used to calculate stochastic quasigradients for DEDS (see also [4, 10, 17, 24]). An alternative approach is the *likelihood-ratio (LR)* or *score-function (SF)* method. For background on this method see [7, 11, 14, 16]. Both approaches are based on the well-known formula for the interchange of mathematical expectations and gradient signs

$$\nabla_x \mathbb{E} g(x, \omega) = \mathbb{E} \nabla_x g(x, \omega). \quad (1)$$

Generally, a disadvantage of the standard form of perturbation analysis is that, it does not work when the sample-path function is discontinuous in the relevant parameter. In that case, it may no longer be valid to change the order of taking the expectation and the derivative. Ho and Li [9], presented an approach to circumvent this difficulty by using a finite difference approximation ideas. Ermoliev and Gaivoronski [2] suggested using artificial random variables to smooth the function $g(x, \omega)$, combining concurrent approximation and optimization techniques. Gong [5] suggested smoothing the function $g(x, \omega)$ by taking the conditional expectation (Smoothed Perturbation Analysis (SPA)). SPA was developed further by Gong and Ho [6] and Brémaud and Gong [1]. However, this approach is applicable only if there are explicit formulas for these conditional expectations. Rubinstein [15] proposed first changing the variables in the expectations (the *push-out* method) and only then using formula (1). The push-out method generalized the LR/SF approach. Gaivoronski introduced the notion of *infinitesimal augmented perturbation analysis* which is developed further in a paper by Gaivoronski, Shi, and Sreenivas [3]. This approach allows gradient formulas to be derived for the *generalized semi-Markov processes* and generalizes the IPA [8]. In Pflug's forthcoming book [13], the gradient of the performance function is represented as a conditional expectation.

¹ Work performed under the auspices of the U.S. Department of Energy (USDOE). The views expressed in this paper are those of the author and do not reflect any policy or position of the USDOE.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Zazanis [23] used a compensator identity in conjunction with IPA to calculate low-variance, unbiased estimates for DEDS at the expense of additional computational requirements.

This paper proposes an alternative approach, which is called the *Analytic Perturbation Analysis* (APA), using analytic formulas for the derivatives of expectations of indicator functions [19, 20, 21]. An earlier paper [22] establishes the relation of APA to an infinitesimal augmented perturbation analysis [3], and to the push-out method [15] showing that gradients calculated with these approaches can be obtained as special cases of APA. Since these two approaches cover IPA and LR/SF methods, APA can be viewed as a framework unifying different methods and covering a large area of applications.

1.1 Explanation of the Approach

Let $(\mathbf{P}, \mathcal{F}, \Omega)$ be a probability space, and $G(x) = \mathbb{E}g(x, \omega)$ be an expectation of function $g(x, \omega)$ depending upon the control variables $x \in \mathbb{R}^n$ and a random element $\omega \in \Omega$. For convenience, we often omit argument ω and use bold face font for random variables (e.g., $q = q(\omega)$). The function $g(x, \omega)$ can be discontinuous w.r.t. x and ω . We suppose that:

1. the function $g(x, \omega)$ is piece-wise differentiable w.r.t. x and ω ;
2. the set Ω can be split into subsets $\mu^q(x) \in \mathcal{F}$, $q = 1, 2, \dots$

$$\Omega = \bigcup_{q=1}^{\infty} \mu^q(x), \quad (2)$$

and the function $g(x, \omega)$ is differentiable w.r.t. ω on each subset $\mu^q(x)$, $q = 1, 2, \dots$;

3. for any $q \neq j$

$$\mathbf{P}(\mu^q(x) \cap \mu^j(x)) = 0; \quad (3)$$

4. each subset $\mu^q(x)$, $q = 1, 2, \dots$ can be presented by the system of inequalities

$$\mu^q(x) = \{\omega \in \Omega : f^q(x, \omega) \leq 0\}$$

$$\stackrel{\text{def}}{=} \{\omega \in \Omega : f_l^q(x, \omega) \leq 0, 1 \leq l \leq k^q\},$$

where $f_l^q : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$, $1 \leq l \leq k^q$.

$I_{\{f^q(x, \omega) \leq 0\}}$ denotes an indicator function which corresponds to the set $\mu^q(x)$

$$I_{\{f^q(x, \omega) \leq 0\}} = \begin{cases} 1, & \text{if } f^q(x, \omega) \leq 0; \\ 0, & \text{otherwise.} \end{cases}$$

With these definitions, the performance function $G(x) = \mathbb{E}g(x, \omega)$ can be presented as the sum

$$\begin{aligned} G(x) &= \mathbb{E}g(x, \omega) = \mathbb{E}\left[\sum_{q=1}^{\infty} I_{\{f^q(x, \omega) \leq 0\}} g(x, \omega)\right] \\ &= \sum_{q=1}^{\infty} \mathbb{E}[I_{\{f^q(x, \omega) \leq 0\}} g(x, \omega)]. \end{aligned} \quad (4)$$

The function

$$\phi_q(x) \stackrel{\text{def}}{=} \mathbb{E}[I_{\{f^q(x, \omega) \leq 0\}} g(x, \omega)]$$

is an integral of the function $g(x, \omega)$ over the set $\mu^q(x)$. This function is differentiable under general conditions (see, Appendix). The gradient $\nabla_x \phi_q(x)$ can be presented as an integral of another function $a^q(x, \omega)$ over the same set $\mu^q(x)$ plus an additional function $\psi_q(x)$, which is a surface integral, or equivalently, as a mathematical expectation of the product $a^q(x, \omega) I_{\{f^q(x, \omega) \leq 0\}}$, plus $\psi_q(x)$, i.e.,

$$\nabla_x \phi_q(x) = \mathbb{E}[I_{\{f^q(x, \omega) \leq 0\}} a^q(x, \omega)] + \psi^q(x).$$

Thus,

$$\begin{aligned} \nabla_x G(x) &= \nabla_x \sum_{q=1}^{\infty} \phi_q(x) = \sum_{q=1}^{\infty} \nabla_x \phi_q(x) \\ &= \sum_{q=1}^{\infty} \mathbb{E}[I_{\{f^q(x, \omega) \leq 0\}} a^q(x, \omega)] + \sum_{q=1}^{\infty} \psi^q(x) \\ &= \mathbb{E}a^q(x, \omega) + \sum_{q=1}^{\infty} \psi^q(x), \end{aligned} \quad (5)$$

where,

$$q = \min\{\nu : \omega \in \mu^\nu(x)\}.$$

The random vector $a^q(x, \omega)$ usually can be obtained with one Monte-Carlo simulation run of the model, analogous to the random value $g(x, \omega)$. Generally,

$$a^q(x, \omega) \neq \nabla_x g(x, \omega), \quad \sum_{q=1}^{\infty} \psi^q(x) \neq 0, \quad (6)$$

therefore, in many cases, formula (1) is not valid.

The idea to split the set Ω into subsets $\mu^q(x)$, $q = 1, 2, \dots$ is not new and was used in different forms in papers [1, 2, 3, 5, 13, 15, 23] and others. In this paper, we combine this idea with new analytic formulas for the differentiating the expectations of indicator functions.

Figure 1: System Flow Chart.

2 Example: A simple Multiclass Queuing Network

The following example was considered by Gong and Ho [5]. Here, we consider a case with arbitrary smooth distributions and probability performance function.

The system consists of two servers and two classes of customers (Figure 1). Each class consists of only one customer and has a different routing. The class 1 customer goes back to server 1 immediately after leaving the server. The class 2 customer goes from server 1 to server 2, then back to server 1. Service times on server 1 for the class 1 customer are independent random variables and have the same density function ρ . We assume that the density function ρ and its derivatives are available. We denote by $L(\omega)$ the interdeparture time of class 2 customer from server 1. The performance we are concerned with here is the cumulative distribution function of interdeparture time $L(\omega)$, i.e.,

$$G(x) = \mathbb{P}[L(\omega) \leq x] = \mathbb{E}[I_{\{L(\omega) \leq x\}}].$$

As Gong and Ho discussed [5], interchange formula (1) is not applicable to these kind of problems.

Suppose that each simulation run begins and ends with the successive arrivals of class 2 customer at server 2; this is a regenerative cycle. Suppose, in this regenerative cycle, a class 1 customer was served q times by server 1. Denote by ζ_i , $i = 1, \dots, q$ service times of customer 1 by server 1 and by η_2 the service time of customer 2 by server 2, and by η_1 the service time of customer 2 by server 1 (Figure 2).

One regenerative cycle is defined by the positive random variables η_2 , η_1 and ζ_i , $i = 1, \dots, q$, where the random variable q can be an arbitrary large positive integer number. The set Ω of random elements can be split into subsets $\mu^q(x)$, $q = 1, 2, \dots$

$$\mu^q(x) = \{\omega \in \Omega : \sum_{i=1}^{q-1} \zeta_i \leq \eta_2 \leq \sum_{i=1}^q \zeta_i, \sum_{i=1}^q \zeta_i + \eta_1 \leq x\}.$$

Figure 2: One Regenerative Cycle.

The first and second inequality which define the set $\mu^q(x)$ means that customer 2 arrived at server 1 when customer 1 had service number q . The third inequality follows from the fact that interdeparture time

$$L(\omega) = \sum_{i=1}^q \zeta_i + \eta_1$$

is less than x . The performance function $G(x)$ can be presented as

$$G(x) = \sum_{q=1}^{\infty} \mathbb{E}[I_{\{\mu^q(x)\}}].$$

Let $\mathbb{E}_q$ be conditional expectation with respect to σ -algebra generated by the random variables η_2 , η_1 and ζ_i , $i = 1, \dots, q-1$, then

$$\mathbb{E}[I_{\{\mu^q(x)\}}] = \mathbb{E}\mathbb{E}_q[I_{\{\mu^q(x)\}}].$$

The function $\mathbb{E}_q[I_{\{\mu^q(x)\}}]$ is smooth with respect to x ; therefore, under general conditions

$$\begin{aligned} \nabla_x G(x) &= \nabla_x \sum_{q=1}^{\infty} \mathbb{E}[I_{\{\mu^q(x)\}}] = \nabla_x \sum_{q=1}^{\infty} \mathbb{E}[\mathbb{E}_q[I_{\{\mu^q(x)\}}]] \\ &= \sum_{q=1}^{\infty} \mathbb{E}[\nabla_x \mathbb{E}_q[I_{\{\mu^q(x)\}}]] \end{aligned} \quad (7)$$

The function $\mathbb{E}_q[I_{\{\mu^q(x)\}}]$ can be presented as an integral

$$\begin{aligned} \mathbb{E}_q[I_{\{\mu^q(x)\}}] &= \int_{\substack{\eta_2 \leq \sum_{i=1}^{q-1} \zeta_i + \zeta_q \\ \sum_{i=1}^{q-1} \zeta_i + \zeta_q + \eta_1 \leq x}} \rho(\zeta_q) d\zeta_q \\ &= \int_{f^q(x, \zeta_q) \leq 0} \rho(\zeta_q) d\zeta_q \stackrel{\text{def}}{=} F(x), \end{aligned} \quad (8)$$

where,

$$f^q(x, \zeta_q) = \left(\frac{-\sum_{i=1}^{q-1} \zeta_i + \eta_2 - \zeta_q}{\frac{\sum_{i=1}^{q-1} \zeta_i - \eta_2 + \zeta_q}{x - \eta_1 - \eta_2} - 1} \right).$$

Let us use formula (16) of Appendix to calculate $\nabla_x F(x)$. The change of variables

$$z = \gamma^{-1}(x, \zeta_q) \stackrel{\text{def}}{=} \frac{-\sum_{i=1}^{q-1} \zeta_i + \eta_2 - \zeta_q}{x - \eta_1 - \eta_2}$$

eliminates the variable x from the function $f^q(x, \zeta_q)$ and

$$\zeta_q = \gamma(x, z) = -z(x - \eta_1 - \eta_2) - \sum_{i=1}^{q-1} \zeta_i + \eta_2.$$

The gradient of the function $F(x)$ is presented with the function (see (18))

$$\begin{aligned} H(x, y) &= \nabla_x \gamma(x, z)|_{z=\gamma^{-1}(x, y)} = -z|_{z=\gamma^{-1}(x, y)} \\ &= \frac{\sum_{i=1}^{q-1} \zeta_i - \eta_2 + \zeta_q}{x - \eta_1 - \eta_2}. \end{aligned}$$

With formula (16)

$$\begin{aligned} \nabla_x F(x) &= \int_{f^q(x, \zeta_q) \leq 0} \text{div}_{\zeta_q} (\rho(\zeta_q) H(x, \zeta_q)) d\zeta_q \\ &= \int_{f^q(x, \zeta_q) \leq 0} \frac{\partial}{\partial \zeta_q} \left(\rho(\zeta_q) \frac{\sum_{i=1}^{q-1} \zeta_i - \eta_2 + \zeta_q}{x - \eta_1 - \eta_2} \right) d\zeta_q \\ &= \frac{1}{x - \eta_1 - \eta_2} \int_{f^q(x, \zeta_q) \leq 0} \left[1 + \left(\sum_{i=1}^{q-1} \zeta_i - \eta_2 + \zeta_q \right) \frac{\partial}{\partial \zeta_q} \ln \rho(\zeta_q) \right] \rho(\zeta_q) d\zeta_q \\ &= \mathbb{E}_q [a^q(x, \omega) I_{\{f^q(x, \zeta_q) \leq 0\}}] = \mathbb{E}_q [a^q(x, \omega) I_{\{\mu^q(x)\}}] \quad (9) \end{aligned}$$

where,

$$\begin{aligned} a^q(x, \omega) &= \frac{1}{x - \eta_1 - \eta_2} \\ &\times \left[1 + \left(\sum_{i=1}^{q-1} \zeta_i - \eta_2 + \zeta_q \right) \frac{\partial}{\partial \zeta_q} \ln \rho(\zeta_q) \right]. \end{aligned}$$

Finally, combining (7), (8), and (9)

$$\nabla_x G(x) = \sum_{q=1}^{\infty} \mathbb{E} [\mathbb{E}_q [a^q(x, \omega) I_{\{\mu^q(x)\}}]]$$

$$= \sum_{q=1}^{\infty} \mathbb{E} [a^q(x, \omega) I_{\{\mu^q(x)\}}] = \mathbb{E} a^q(x, \omega). \quad (10)$$

Therefore, $a^q(x, \omega)$ is an unbiased estimate of the gradient $\nabla_x G(x)$.

We used formula (16) to obtain the last estimate. Since the gradient of function (8) can be presented also with formulas (13) and (15), then estimate (10) can be considered as one representative

References

- [1] Brémaud, P. and W.B. Gong, Derivatives of Likelihood Ratios and "Smoothed Perturbation Analysis for Routing Problem." *ACM Transactions on Modelling and Computer Simulation*, 3, No. 2, 134-161, 1993.
- [2] Ermoliev, Yu. and A. Gaivoronski, *On Optimization of Discontinuous Systems*. Working Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria, WP-91-49, 1991.
- [3] Gaivoronski, A.A., Shi, L.Y. and R.S. Sreenivas, "Augmented Infinitesimal Perturbation Analysis: An Alternative Explanation." *Discrete Event Dynamic Systems: Theory and Applications*, 2, pp. 121-138, 1992.
- [4] Glasserman P., *Gradient Estimation Via Perturbation Analysis*. Kluwer Academic Publishers, Boston-Dordrecht-London, 1991.
- [5] Gong W.B., *Smoothed Perturbation Analysis of Discrete Event Dynamic Systems*. Ph.D. thesis, Div. of Applied Sciences, Harvard Univ., Cambridge, Mass, 1987.
- [6] Gong W.B. and Y.C. Ho, "Smoothed (Conditional) Perturbation Analysis of Discrete Event Dynamic Systems". *IEEE - Transactions on Automatic Control* 32, pp. 858-866, 1987.
- [7] Glynn P.W., "Likelihood ratio gradient estimation for stochastic systems". *Communications of the ACM*, 33, pp. 75-84, 1990.
- [8] Ho Y.C., Eyler M.A. and T.T. Chien, "A Gradient Technique for General Buffer Storage Design in a Serial Production Line." *Intern. J. Production Res.* 17, pp. 557-580, 1979.
- [9] Ho Y.C. and S. Li, Extensions of Infinitesimal Perturbation Analysis. *IEEE - Transactions on Automatic Control*, 33, pp. 427-438, 1988.

- [10] Ho Y.C. and X.R. Cao, *Perturbation Analysis of Discrete Event Dynamic Systems*. Kluwer, Boston, 1991.
- [11] L'Ecuyer P.L., "A unified version of the IPA, SF, and LR gradient estimation techniques." *Management Science*, 36, pp. 1364-1383, 1990.
- [12] Marti K., *Approximations and Derivatives of Probability Functions in Structural Reliability and Design*, Proc. of 6-th Intern. Conf. on Structural Reliability and Design, Innsbruck, Austria, 1295-1299, 1993.
- [13] Pflug G.Ch., *Simulation and Optimization - the Interface*, Kluwer Academic Publishers, (in press).
- [14] Reiman, M.I. and A. Weiss, "Sensitivity Analysis for Simulations via Likelihood Ratios." *Operations Research*, 37, pp. 830-844, 1989.
- [15] Rubinstein R., "Sensitivity Analysis of Discrete Event Systems by the Push Out Method." *Annals of Operations Research*, 39, 1992.
- [16] Rubinstein R. and A. Shapiro, *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization via the Score Function Method*, Wiley, Chichester, 1993.
- [17] Suri R., "Perturbation Analysis: the State of the Art and Research Issues Explained via the GT/G/1 Queue." *Proceedings of the IEEE*, 77, pp. 114-137, 1989.
- [18] Uryasev, S.P., "A Differentiation Formula for Integrals over Sets Given by Inclusion." *Numerical Functional Analysis and Optimization*. 10(7 & 8), pp. 827-841, 1989.
- [19] Uryasev, S.P., "Differentiation Formula for Integrals over Sets Given by Inequalities." Working Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria, WP-91-38, 1991.
- [20] Uryasev, S.P., "Derivatives of Probability Functions and Integrals over Sets Given by Inequalities." *J. Computational and Applied Mathematics*, Vol. 56., 1994.
- [21] Uryasev, S.P., *Derivatives of Probability Functions and some Applications*. Proceedings of VI International Conference on Stochastic Programming. Udine, Italy, September, 1992 (in press).
- [22] Uryasev, S.P., "Analytic Perturbation Analysis for DEDS with Discontinuous Sample-Path Functions." *Stochastic Models*, (in press).

- [23] Zazanis, M.A., "Derivative Estimation via Stochastic Intensities: Event Averages in Queuing Systems." *Letters of Operations Research*, (in press).
- [24] Zazanis, M.A., "Perturbation Analysis of the GI/GI/1 Gueue." *QUESTA*, (in press).

Appendix. The Analytical Derivatives of the Integrals over Sets Given by Inequalities

Let the function

$$F(x) = \int_{f(x,y) \leq 0} p(x,y) dy \quad (11)$$

be defined on the Euclidean space $\mathbb{R}^n$, where $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^k$ and $p : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ are some functions. The inequality $f(x,y) \leq 0$ should be treated as a system of inequalities

$$f_i(x,y) \leq 0, \quad i = 1, \dots, k.$$

For example, let

$$\begin{aligned} F(x) &= \mathbb{E}[I_{\{f(x,\zeta) \leq 0\}} g(x,\zeta)] \\ &= \int_{f(x,y) \leq 0} g(x,y) \rho(x,y) dy = \int_{f(x,y) \leq 0} p(x,y) dy, \end{aligned} \quad (12)$$

where random vector ζ in $\mathbb{R}^m$ has a probability density $\rho(x,y)$ that depends on a parameter $x \in \mathbb{R}^n$.

Further, we presented a general formula [20, 21] for the differentiation of integral (11); a partial case of the formula was also published earlier [18]. A gradient of the integral is presented as a sum of integrals taken over a volume and over a surface.

Let us introduce the following shorthand notations

$$f(x,y) = \begin{pmatrix} f_1(x,y) \\ \vdots \\ f_k(x,y) \end{pmatrix}, \quad f_{11}(x,y) = \begin{pmatrix} f_1(x,y) \\ \vdots \\ f_1(x,y) \end{pmatrix},$$

$$\nabla_y f(x,y) = \begin{pmatrix} \frac{\partial f_1(x,y)}{\partial y_1}, & \dots, & \frac{\partial f_k(x,y)}{\partial y_1} \\ \vdots & & \vdots \\ \frac{\partial f_1(x,y)}{\partial y_m}, & \dots, & \frac{\partial f_k(x,y)}{\partial y_m} \end{pmatrix}.$$

Divergence for the matrix H is denoted by

$$\operatorname{div}_y H = \begin{pmatrix} \sum_{i=1}^m \frac{\partial h_{1i}}{\partial y_i} \\ \vdots \\ \sum_{i=1}^m \frac{\partial h_{mi}}{\partial y_i} \end{pmatrix}, \quad H = \begin{pmatrix} h_{11} & \dots & h_{1m} \\ \vdots & & \vdots \\ h_{n1} & \dots & h_{nm} \end{pmatrix}$$

We define

$$\mu(x) = \{y \in \mathbb{R}^m : f(x, y) \leq 0\}$$

$$\stackrel{\text{def}}{=} \{y \in \mathbb{R}^m : f_l(x, y) \leq 0, 1 \leq l \leq k\}$$

and $\partial\mu(x)$ to be the surface of the set $\mu(x)$. Let us denote by $\partial_i\mu(x)$ a part of the surface which corresponds to the function $f_i(x, y)$

$$\partial_i\mu(x) = \mu(x) \cap \{y \in \mathbb{R}^m : f_i(x, y) = 0\}.$$

If we split the set $K \stackrel{\text{def}}{=} \{1, \dots, k\}$ into two subsets K_1 and K_2 , without loss of generality we can consider

$$K_1 = \{1, \dots, l\} \text{ and } K_2 = \{l+1, \dots, k\}.$$

There is a freedom in the choice of the sets K_1 and K_2 and representation of the gradient of function (11). First, we consider the case when the subsets K_1 and K_2 are not empty. In this case, the derivative of integral (11) is given by the formula

$$\begin{aligned} \nabla_x F(x) = & \int_{\mu(x)} [\nabla_x p(x, y) + \operatorname{div}_y (p(x, y) H_l(x, y))] dy \\ & - \sum_{i=l+1}^k \int_{\partial_i\mu(x)} \frac{p(x, y)}{\|\nabla_y f_i(x, y)\|} [\nabla_x f_i(x, y) \\ & + H_l(x, y) \nabla_y f_i(x, y)] dS, \end{aligned} \quad (13)$$

where the matrix function $H_l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^{n \times m}$ satisfies the equation

$$H_l(x, y) \nabla_y f_{1l}(x, y) + \nabla_x f_{1l}(x, y) = 0. \quad (14)$$

The last equation can have a lot of solutions and we can choose an arbitrary solution, differentiable with respect to the variable y .

Further, let us present the derivative of function (11) for the case with the empty set K_1 . Then, matrix function H_l is absent and

$$\nabla_x F(x) = \int_{\mu(x)} \nabla_x p(x, y) dy$$

$$- \sum_{i=1}^k \int_{\partial_i\mu(x)} \frac{p(x, y)}{\|\nabla_y f_i(x, y)\|} \nabla_x f_i(x, y) dS. \quad (15)$$

Finally, let us consider a formula for the derivative of function (11) for the case with the empty set K_2 . The integral over the surface is absent and the derivative is represented as an integral over the volume

$$\nabla_x F(x) = \int_{\mu(x)} [\nabla_x p(x, y) + \operatorname{div}_y (p(x, y) H(x, y))] dy, \quad (16)$$

where a matrix function $H : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^{n \times m}$ satisfies the equation

$$H(x, y) \nabla_y f(x, y) + \nabla_x f(x, y) = 0. \quad (17)$$

In many cases there is a simple way to solve equations (14) and (17) using change of variables. Suppose that there is a change of variables

$$y = \gamma(x, z)$$

which eliminates vector x from the function $f(x, y)$, i.e., function $f(x, \gamma(x, z))$ does not depend upon variable x . Denote by $\gamma^{-1}(x, y)$ the inverse function, defined by the equation

$$\gamma^{-1}(x, \gamma(x, z)) = z.$$

In this case

$$H(x, y) = \nabla_x \gamma(x, z)|_{z=\gamma^{-1}(x, y)}. \quad (18)$$

Formula (15) with matrix (18) covers the derivative formulas which can be obtained with change of variables in the integration set [12]. Indeed, gradient of the function $\gamma(x, y(x, z))$ with respect to x equals zero, therefore

$$\begin{aligned} 0 = \nabla_x f(x, \gamma(x, z)) = & \nabla_x \gamma(x, z) \nabla_y f(x, y)|_{y=\gamma(x, z)} \\ & + \nabla_x f(x, y)|_{y=\gamma(x, z)}, \end{aligned}$$

and function $\nabla_x \gamma(x, z)|_{z=\gamma^{-1}(x, y)}$ is a solution of equation (17).

Formulas (16), (15), and (13), should be treated, actually, as a family of formulas. The following components are not uniquely defined in these formulas

- two subsets K_1 and K_2 ,
- matrix $H(x, y)$ and matrix $H_l(x, y)$,
- different vector functions $f(x, y)$ may represent the same integration set $\mu(x)$.

**DATE
FILMED**

10/21/94

END

5.