Novel Response Surface Methodologies with Design of Experiment for Source Localization in Unknown Spatial-Temporal Fields

Zhenyi Liu, Philip Smith, Trevor Park, A. Alexandre Trindade, and Qing Hui

Abstract—The spread of contaminant into the environment due to a high volume of leakage from point sources is a very real threat in the modern world. Since the contaminant can often be dangerous to human operators, multiple robotic agents equipped with suitable sensors are playing an increasingly important role in the so-called contaminant detection problem than before. This paper proposes a contaminant detection methodology which can not only locate the (possibly multiple) source of the contaminant, but also estimate the intensity in an environment evolving over both space and time. The method can deal with measurement noise and does not need any gradient information.

I. INTRODUCTION

The spread of contaminant in the environment due to high volume of leakage from point sources is a very real threat in the modern world [1]. It is also important to be able to map the intensity and distribution of the contaminant quickly, in order to effectively move people and property out of harm’s way, warn others to stay away, and allow for clean-up operations to be conducted. Often the contaminant can be dangerous to human operators, or they may cause visual impairment (e.g., from smoke or fumes) which prevents human agents from carrying out these tasks. An alternative is to use multiple robotic agents equipped with suitable sensors to detect the intensity of contaminant at a point. In this paper, we propose new algorithms which can not only locate the source(s) of contaminant, but also estimate the intensity of the contaminant while simultaneously accounting for temporal evolution of the spread and source(s). The methods can deal with measurement noise and do not need any gradient information.

There are many papers in the literature dealing with threat or contaminant detection utilizing multiple robots like [1]–[5], to cite but a few examples. In the contaminant detection problem, apart from localizing the sources, another important task is to determine the distribution of the intensity of the contaminant so that the subsequent aid or cleaning can be conducted efficiently based on this information. In the contaminant detection problem, most of the papers shown above in the literature only focus on locating the sources of contaminant. Few consider dealing with the two problems simultaneously. In one of the previous papers, the sources of contaminant and the boundary of the affected region can be found. However, the method needs a large number of robots and does not include any noise in the measurement. In addition, if the contaminant is radioactive waste on the sea, as the waste moves with waves, the position and the boundary of the contaminant also move. Besides, to our knowledge, even though some of the papers consider the situation where the wind may change the concentration of the contaminant, none of the above referenced works considers the case where the sources of contamination are time-varying. This is quite common if the sources of contaminant are in the moving water such as rivers and the sea.

In this paper, we propose methods to conduct threat detection and optimal allocation of robotic agents using key ideas from statistical design of experiments. The main contributions and features of the methodology are as follows:

1. It is able to pinpoint not only the source(s) of contamination, but also the extent of contamination in a given geographic region.
2. It can deal with multiple sources.
3. The positions of the contaminant sources can evolve with time.
4. Contrary to many current competing methods, our approach does not use any gradient information.
5. No formation of robots is needed, and the measurements recorded by the robots are assumed to be noisy.
6. Fewer robots are needed than many state-of-the-art techniques.

After introducing basic statistical concepts about nonlinear regression and optimal designs, we will briefly formulate the problem. Next, we propose our methods and algorithms, and illustrate the performance via some simulation results. In the end, convergence analysis of estimates is discussed.

II. PRELIMINARIES ON STATISTICS

In this section we introduce some basic notions of response surface methodology (RSM), design of experiments (DoE), which are the fundamental statistical tools used in our method.

A. Response Surface Methods

RSM is a technique for modeling how a response variable of interest is affected by several control variables [6]. In our problem, the response surface is the surface formed by the
concentration of the contaminant in a region and the variables that influence the surface are the coordinates in the region.

In response surface methodology, if the data points (observations of explanatory variables and response variables) are given, then the surface can be easily fitted by regression methods. However, in our problem, we need robots to move in the contaminated region and pick the data points. Therefore, a main task for us is to propose a method to pick the data points that can “best” fit the response surface.

B. Optimal Experimental Designs

Optimal designs indicate where the response surface should be sampled (which design points) in order to minimize the variability of the estimates resulting from fitting a model with data collected at the chosen points [6]. Standard designs such as Central Composite Design, Box-Behnken, and their variants, are widely used because they are quite general and flexible.

However, occasionally an experimenter encounters a situation where standard response surface designs may not be appropriate. (1) If the region of interest is not a regular region, standard designs may not be the best choice. (2) Usually an experimenter selects a first- or second-order response surface model, realizing that this is only an approximation to the true underlying mechanism. In this case, special knowledge or insight about the process being studied may suggest a nonstandard model [6]. In our problem, the robots may encounter some obstacles which might prevent us from placing them anywhere we want. Besides, the model we use may not be the standard first- or second-order response surface model. Thus, the optimal designs are able to tell us how to move our robots.

By an optimal design, we mean a design that is “best” with respect to some criterion. The usual approach is to specify the optimality criterion, and then choose the design points from a set of candidate points that the experimenter would consider using.

There are several popular design optimality criteria and in this paper we just choose one of them. Suppose we have the nonlinear model as follows:

\[
y_i = f(x_i, \gamma) + \varepsilon_i = f(x_i; \gamma) + \varepsilon_i, i = 1, \ldots, N
\]

where \(N\) is the number of observations, \(y_i\) is the \(i\)-th observation of the response variable, \(x_i\) is the \(i\)-th observation of the explanatory variable, \(\varepsilon_i\) is noise, \(\varepsilon_1, \ldots, \varepsilon_N \sim \text{i.i.d.}(0, \sigma^2)\), where i.i.d. denotes independent and identically distributed. \(\gamma\) is the vector of unknown parameters. Then we have the matrix \(F\):

\[
F = \begin{bmatrix}
  f_{i1} & \cdots & f_{im} \\
  \vdots & \ddots & \vdots \\
  f_{n1} & \cdots & f_{nm}
\end{bmatrix}
\]

where the \(f_{ij} = \frac{\partial f(x, \gamma)}{\partial \gamma_j} \big|_{\gamma = \gamma_0}\) for a given \(\gamma_0\).

A design is said to be D-optimal if \(|\text{det}(F^T F)|^{-1}\) is minimized, where \(|A|\) denotes the determinant of square matrix \(A\). It turns out that a D-optimal design minimizes the volume of the joint confidence region on the vector of regression coefficients [7].

III. Problem Formulation

In this section, we formulate our problem into three different cases which we call respectively fixed-source case, moving-source case, and fixed-source case with autocorrelated errors. The fixed-source case means that the positions of sources of contaminant are time-invariant. Similarly, the moving-source case means that the positions of sources of contaminant are time-varying. Since our robots take data over time, it is highly possible that the errors are not independent and identically distributed (i.i.d.) but autocorrelated. Therefore, we also need to consider the fixed-source case with autocorrelated errors. The moving-source case with autocorrelated errors is much more complicated and hence, we put it into the future work. The fixed-source case with autocorrelated errors is similar to the fixed-source case except that the errors are not independent and identically distributed (i.i.d.) but autocorrelated. Before we proceed to introduce the methods, we make several assumptions. First, we assume that the real distribution function of contaminant is a multiple-hump continuous function whose form is unknown and the sources of contaminant stay at the peak of the humps. This is reasonable because for most common contaminant like chemical, biological, and even radiation, the concentration is larger nearer the source than further away from it. It is true especially when the contaminant is of radiative property like nuclear waste. We use a sum of Radial Basis Functions (RBFs) as an approximating function to approximate the real form of the distribution function because sums of Radial Basis Functions are typically used to approximate some given function. The second assumption is that the number of Radial Basis Functions we use in the approximating function is always greater than the number of the sources of contaminant. This assumption is also reasonable because as technology advances the computational cost will decrease with time so that we may choose as many RBFs as we can to approximate the real function.

A. Fixed-source Case

According to the previous assumptions, an approximating function or a model based on RBFs is proposed here to approximate the real form of the distribution function of contaminant:

\[
f(x; \gamma) = a_0 + \sum_{k=1}^{h} \frac{a_k}{1 + \|L_k(x - c_k)\|^2}
\]

where \(h\) is the number of RBFs we use in the model, \(x = [x_1, x_2]^T \in \mathbb{R}^2\) is a 2-dimensional Cartesian coordinate which is a spatial location, \(a_0 \in \mathbb{R} = (-\infty, \infty), a_k \in \mathbb{R}^+ = (0, \infty)\), \(k = 1, \ldots, h\) are unknown parameters, \(c_k = [c_{k1}, c_{k2}]^T\) is the center point of the hump of \(k\)-th RBF in the function, and \(L_k = \begin{bmatrix} l_{k11} & 0 \\ l_{k21} & l_{k22} \end{bmatrix}\) is the parameter matrix in the hump of \(k\)-th RBF which determines the shape of the hump. Let \(d(x)\) be the unknown real distribution function and with data
\[ y_i = d(x_i) + \varepsilon_i, \quad \varepsilon_1, \ldots, \varepsilon_N \sim \text{i.i.d.}(0, \sigma^2) \quad (N \text{ is the number of observations}) \quad \text{and} \quad x_i, \quad \text{our goal is that by choosing} \]

\[ \text{The procedure here begins with initialization of the vector} \quad \begin{bmatrix} \gamma_0 \end{bmatrix} \]

\[ \text{At the} \ t \text{-th iteration, for a given} \quad \gamma^{(t-1)}, \text{collect data at} \]

\[ q \text{ points about a circle of a given radius centered at the robot. Among the} \]

\[ q \text{ points thus obtained, identify the D-optimal one (minimizes} \quad |F^T F|^{-1}) \quad \text{and record this direction} \]

\[ \text{for subsequent movement at the next step. Before going to} \]

\[ \text{phase one again,} \quad \gamma^{(0)} \text{is updated. A summary is provided in} \]

\[ \text{Algorithm 1 Detection and Location of Contaminant in the} \]

\[ \text{Fixed-source Case} \]

\[ \text{Move the robots to get the initial data set} \quad D^{(0)} \]

\[ \text{With the data and the initial guess of the parameter vector} \quad \gamma^{(0)}, \text{obtain} \]

\[ \gamma^{(1)} = \text{lsqcurvefit}(D^{(1)}, \gamma^{(0)}) \]

\[ \text{for} \quad t = 2, \ldots, v \quad \text{do} \]

\[ \text{Collect} \quad q \text{ points on the circle around the robots with some radius.} \]

\[ \text{With data set} \quad D^{(t-1)} \quad \text{and each point around the circle, calculate matrix} \]

\[ F = \{ f_i \} = \left\{ \frac{\partial f(x; \gamma)}{\partial \gamma} \right\}_{\gamma = \gamma^{(t-1)}} \]

\[ \text{Find the point on the circle that can minimizes} \quad |F^T F|^{-1}. \]

\[ \text{Make it the direction of the next step.} \]

\[ \text{Move all the robots toward the respective directions we} \]

\[ \text{obtain above and at the same time take data and put all} \]

\[ \text{the data just taken and all previous data set} \quad D^{(t-1)} \quad \text{into} \]

\[ \text{the new data set} \quad D^{(t)}. \]

\[ \text{Estimate the parameter vector:} \quad \gamma^{(t)} = \text{lsqcurvefit}(D^{(t)}, \gamma^{(t-1)}) \]

\[ \text{end for} \]

\[ \text{Output the parameters estimated.} \]

In the algorithm, \text{lsqcurvefit} is the MATLAB® function to obtain ordinary least squares (OLS) estimation. \( D^{(t)} \)

\[ \text{is a data matrix containing observations of explanatory and response variables up to the} \quad t \text{-th iteration.} \]

\[ \text{B. Moving-source Case} \]

\[ \text{In this case there are two differences from the fixed-source case. One of them is that the data taken from earlier time steps should be weighted less than those taken from later time steps. So we down-weight data from earlier time steps by} \phi \in (0, 1). \text{As a result, we estimate} \gamma \text{with weighted least squares estimation} \]

\[ \text{WLS}(\gamma) = [y - f(X; \gamma)]^T \Lambda [y - f(X; \gamma)] \]

\[ \text{where} \quad y = [y_1, \ldots, y_N]^T, \quad f(X; \gamma) = [f(x_1; \gamma), \ldots, f(x_N; \gamma)]^T \]

\[ \text{and} \]

\[ \Lambda = \begin{bmatrix} \phi^{N-1} I_{nr} & 0 & \cdots & 0 \\ 0 & \phi^{N-2} I_{nr} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_{nr} \end{bmatrix} \]

where \( \phi \in (0, 1), \text{nr} \) is the number of robots. \( I_{nr} \) is nr \times nr identity matrix. Thus, the weighted least squares estimation allows us to estimate \( \gamma \) to slowly “forget” earlier information that may no longer be relevant. And the other difference from the fixed-source case is that the information matrix is
no longer $F^T F$ but $F^T Λ F$. The algorithm in this case is the same with the fixed-source case except $γ(t)$ is obtained by the weighted least squares estimation and $F^T F$ is replaced by $F^T Λ F$.

C. Fixed-source Case With Autocorrelated Errors

Given the model

$$ y = f(X; γ) + ε, ε \sim (0, \sigma^2_{\varepsilon}) $$

where $f(X; γ) = \{f(x_1; γ), \ldots, f(x_N; γ)\}^T$ and $f(x_i; γ) = a_0 + \sum_{k=1}^{3} a_k 1 + \|L_k(x - c_k)\|^2$, $i = 1, \ldots, N$. Since the errors are autocorrelated, we should estimate $γ$ with generalized least squares criterion

$$ GLS(γ) = [y - f(X; γ)]^T Γ^{-1} [y - f(X; γ)] $$

Also, the information matrix becomes $F^T Γ^{-1} F$. One problem is that we do not know the form of covariance matrix $Γ$. First, in this case, we assume that our robots take data one after the other so that the data observed can be regarded as time series. One of the methods to specify the unknown covariance matrix $Γ$ is to entertain an ARMA($p, q$) model for $\{ε_t\}$ and use conditional least squares to estimate parameters of ARMA($p, q$) process [8].

$$ ε_t - φ_1 ε_{t-1} - \cdots - φ_q ε_{t-q} = z_t + θ_1 z_{t-1} + \cdots + θ_q z_{t-q} $$

where $\{z_t\} \sim i.i.d.(0, σ^2_z)$. And we minimize a conditional least squares criterion to estimate $φ = [φ_1, \ldots, φ_q]^T$ and $θ = [θ_1, \ldots, θ_q]^T$. The algorithm is given in Algorithm 2.

Algorithm 2 Detection and Location of Contaminant in the Fixed-source Case With Autocorrelated Errors

Move the robots to get the initial data set $D^{(1)}$.
With the data and the initial guess of the parameter vector $γ^{(0)}$, obtain OLS estimates $γ^{(1)}$.
Form the residuals $ε^{(1)} = y - f(X; γ^{(1)})$ and determine an appropriate ARMA for this process (via Akaike Information Criterion (AIC)).

For $t = 2, \ldots, v$ do
Collect $q$ points on the circle around the robots with some radius.
With data set $D^{(t-1)}$ and each point around the circle, calculate matrix $F = \{f_i\} = \{\frac{∂(x; γ)}{∂γ} | γ = γ^{(t-1)}\}$. Find the point on the circle that can minimizes $\|F^T F\|^{-1}$. Make it the direction of the next step.
Move all the robots toward the respective directions we obtain above and at the same time take data and put all the data just taken and all previous data set $D^{(t-1)}$ into the new data set $D^{(t)}$.
Obtain GLS estimates $γ^{(t)}$ with $Γ = Γ^{(t-1)}$.
Form the residuals $ε^{(t)} = y - f(X; γ^{(t)})$ and update $Γ^{(t-1)} → Γ^{(t)}$ by minimizing $Q(φ, θ; γ^{(t)})$ w.r.t. $φ, θ$.

end for
Output the parameters estimated.

V. SIMULATION RESULTS

This section presents some simulation results in the three cases.

A. Fixed-source Case

In this part, for simplicity, we choose a real distribution function of contaminant whose form is similar to our approximating function. It has the form $d(x) = b_0 + \sum_{k=1}^{2} b_k 1 + \|L_k(x - c_k)\|^2$, which means that there are two sources of contaminant and the $i$-th observation of the concentration of the contaminant we measure is $y_i = d(x_i) + ε_i$ where $ε_1, \ldots, ε_N \sim i.i.d.(0, σ^2)$. In real world, it is very difficult to know the exact number of sources of contaminant so it is highly possible that we do not choose the “right” model.

Here we use an approximating function with three RBFs $f(x; γ) = a_0 + \sum_{k=3}^{3} a_k 1 + \|L_k(x - c_k)\|^2$ as shown in model (3). We use 6 robots to estimate the parameters. Let $v$ be iteration times and make $v = 50$. The step size in Phase 1 is 0.04 and the number of steps in each Phase 1 is 20. The contour plot in Figure 1 is generated with the real parameters. Since $σ^2$ is not used in estimating other parameters, it is not estimated. Figure 1 shows the trajectories of the 6 robots. Figure 2 is the coefficient of determination ($R^2$) as a function of iteration. The larger $R^2$ is, the better results we have. From the estimated parameters and $R^2$, we can see that the estimation is satisfactory. In some situations, the structure of the real distribution function may be more complicated and we cannot obtain any information about the estimation results by just checking the estimated parameters. In these situations, $R^2$ is an effective way to judge the estimation results.

B. Moving-source Case

In this part, still suppose the contaminant has two sources and the distribution function of the contaminant is of the form $d(x, t) = b_0 + \sum_{k=1}^{2} b_k 1 + \|L_k(x - h_k(t))\|^2$, which means there are two sources of contaminant and the $i$-th observation of the concentration of the contaminant we measure is $y_i = d(x_i, t) + ε_i$ where $ε_1, \ldots, ε_N \sim i.i.d.(0, σ^2)$. Here we still use an approximating function with three RBFs $f(x; γ) = a_0 + \sum_{k=3}^{3} a_k 1 + \|L_k(x - c_k)\|^2$ as shown in model (3). We also use 6 robots to estimate the parameters and make $v =$
30. Figure 3-6 are the snapshots of real and estimated contour of the sources of contaminant at 0, 10, 20, 30-th iteration respectively. The red circles are the real starting positions of sources of contaminant; the red upward-pointing triangles are the real final positions of sources of contaminant. In Figure 3-6, it can be shown that as the real sources of contaminant move from the starting positions to the final positions, the positions of estimated sources of contaminant can catch up with the real ones. However, one disadvantage is that the shape of the contour is not as accurate as we can have in the time-invariant case.

C. Fixed-source Case With Autocorrelated Errors

In this part, we choose the same real distribution function of contaminant as in the fixed-source case. It has the form 
\[ d(x) = b_0 + \sum_{i=1}^{2} \frac{b_i}{1 + ||x - \xi_i||^2} \]
which means that there are two sources of contaminant and the i-th observation of the concentration of the contaminant we measure is \( y_i = d(x_i) + \varepsilon_i \) and we use ARMA(1,1) process to simulate the autocorrelated errors and \( \{\varepsilon_i\} \sim \text{i.i.d.}(0, \sigma^2) \). We also use 6 robots to estimate the parameters. We pick AR(1), AR(2), MA(1), MA(2), ARMA(1,1), ARMA(2,1), ARMA(1,2), ARMA(2,2) as candidate models. The AIC value suggests to choose AR(2) as the model for estimation. Since our goal is to estimate the parameters of the approximating function as accurately as possible and ARMA process is only one way to fit the errors which belong to some unknown process, it is not essential that we do not pick the "right" ARMA process. Figure 7 shows the coefficient of determination \( R^2 \). The result also suggests the estimation is satisfactory.

VI. Convergence Analysis

In this section, we discuss the convergence analysis about the estimated parameters in the fixed-source case with i.i.d. errors and with autocorrelated errors. Since in the moving-source case the sources of contaminant move with time, the estimated parameters will not converge.

We first deal with the fixed-source case with i.i.d. errors. Consider our nonlinear model 
\[ y_i = f(x_i; \gamma^*) + \varepsilon_i, \quad (i = 1, 2, \ldots, N) \]
where \( \gamma^* \) is the true value of vector parameter \( \gamma \). We give the following assumptions first:

A(1) The \( \varepsilon_i \) are i.i.d. with mean zero and variance \( \sigma^2 \).
A(2) For each i, \( f_i(\gamma) = f(x_i; \gamma) \) is a continuous function of \( \gamma \) for \( \gamma \in \Gamma \).
A(3) \( \Gamma \) is closed and bounded.

And define
\[
B_N(\gamma, \gamma_1) = \sum_{i=1}^{N} f_i(\gamma) f_i(\gamma_1) \\
D_N(\gamma, \gamma_1) = \sum_{i=1}^{N} [f_i(\gamma) - f_i(\gamma_1)]^2
\]
In addition, we add the following assumption: 
A(4)
(a) \( N^{-1} B_N(\gamma, \gamma_1) \) converges uniformly for all \( \gamma, \gamma_1 \in \Gamma \) to a function \( B(\gamma, \gamma_1) \). This implies that \( N^{-1} D_N(\gamma, \gamma_1) \) converges uniformly to \( D(\gamma, \gamma_1) = B(\gamma, \gamma_1) + B(\gamma_1, \gamma_1) - 2B(\gamma, \gamma_1) \)
(b) It is now further assumed that \( D(\gamma, \gamma^*) = 0 \) if and only if \( \gamma = \gamma^* \)

Given the assumptions A(1) to A(3), in the literature, [9] proved that A(4) is sufficient for least squares estimator \( \hat{\gamma} \) to converge to \( \gamma^* \).

There are some papers dealing with assumption A(4a). We introduce one way we can use. Suppose \( x_1, \ldots, x_N \) are data points and \( H_N(x) = \frac{c}{2} \) is the empirical distribution function, where \( c \) is the number of \( x_i \)'s less that or equal to \( x \) (a less than or equal to \( b \) means \( a_i \leq b_i \) for all \( i \) ). [9] showed that if \( H_N(x) \) converges to a distribution function then A(4a) holds. Next, we show how to verify A(4a) based on the results in the literature.

We need to make several assumptions in this section.
Assumption 6.1: All robots will finally move around several fixed points \( \xi_k \), \( k \leq R \) clockwise on circles where \( R \) is the number of the robots and robots move in a straight line in each step.

Assumption 6.2: Second, all the trajectories of robots are in the region \([a_1, b_1] \times [a_2, b_2]\).

The first assumption is based on our observation that most robots will move around some fixed points and for simplicity, we assume all the robots perform the similar moving pattern. Then for a given \( x_p \in \mathbb{R^2} \), \( H_N(x_p) \) becomes
\[
H_N(x_p) = \frac{\text{number of data points } x_i \leq (x_{p,1}, x_{p,2}) \text{ in first } N \text{ data points}}{\text{data points number } N}
\]
We have to show that $H_N(x_p)$ converges to a distribution function as $N \to \infty$.

For $r$-th robot, $r = 1, \ldots, R$, the data points of the robot $r$ takes are $x_{i,r} \in \mathbb{R}^2$, $i = 1, 2, \ldots$. For robot $r$, according to our first assumption, $\exists I_r \in \mathbb{Z}^+ = \{1, 2, \ldots\}$ such that if $i \geq I_r$, then $x_{i,r} \in \{x : \|x - \bar{c}_{i,1}\|^2 = d_i^2\}$, where $\bar{c}_{i,1}$ is a fixed point surrounded by robot $r$.

We define a set $E_r = \{x_1 \geq \bar{c}_{r,1} - d_r, \ x_2 \geq \bar{c}_{r,2} - d_r\} \cup \{x_1 \geq \bar{c}_{r,1}, x_2 \leq \bar{c}_{r,2}\}$. For a given $x_p \in \mathbb{R}^2$, there are $I_r$ sets $E_r', r = 1, \ldots, R$ such that $x_p \in E_r'$, that is, $x_p \in E_{r_1}, E_{r_2}, \ldots, E_{r_I}$. Define several sets, $N_r = \{x_{i,r} : i = 1, 2, \ldots, X_r\}$, $C_r = \{x_{i,r} : i = 1, 2, \ldots, X_r, x_1 \leq \bar{c}_{r,1}, x_2 \leq \bar{c}_{r,2}\}$, $N_{r,n}(N) = \{x_{i,r} : i = 1, 2, \ldots, n\}$, $C_{r,n}(N) = \{x_{i,r} : i = 1, 2, \ldots, cn, x_1 \leq x_{p,1}, x_2 \leq x_{p,2}\}$. The set of points on the trajectory of robot $r$ in the first $N$ data points, $C_{r,n}(N)$ is the set of the points on the trajectory of robot $r$ satisfying $x_1 \leq x_{p,1}, x_2 \leq x_{p,2}$ in the first $N$ data points. $n_r(N)$ and $cn_r(N)$ are non-decreasing with $N$. For simplicity, we write $n_r(N), cn_r(N)$ as $n_r, cn_r$. We have $\lim_{n_r, cn_r \to \infty} N_{r,n} = N_r$. Similarly, $\lim_{n_r, cn_r \to \infty} C_{r,n} = C_r$. Since $n_r$ and $cn_r$ will go to infinity, we can always make $n_r \geq I_r$ and $cn_r \geq I_r$.

Therefore, $H_N(x_p)$ becomes

$$H_N(x_p) = \frac{\sum_{r=1}^R \#N_{r,n}}{\sum_{r=1}^R \#N_{r,n}}$$

where $\#N_{r,n}, \#C_{r,n}$ are the number of elements in sets $N_{r,n}, C_{r,n}$ respectively. We define some other sets, $N_{r,n} = \{x_{i,r} : i = 1, \ldots, n\}$, $C_{r,n} = \{x_{i,r} : i = 1, \ldots, cn, x_1 \leq x_{p,1}, x_2 \leq x_{p,2}\}$. That is, points in $N_{r,n}$ and $C_{r,n}$ are those on the circle $\|x - \bar{c}_{i,1}\|^2 = d_i^2$. $H_N(x_p)$ is

$$H_N(x_p) = \frac{\sum_{r=1}^R \#C_{r,n} + \sum_{r=1}^R \#(C_{r,n} \setminus C_{r,n})}{\sum_{r=1}^R \#N_{r,n} + \sum_{r=1}^R \#(N_{r,n} \setminus N_{r,n})}$$

It is not difficult to show that if $F_N(x_p) = \frac{\sum_{n=1}^N \#N_{r,n}}{\sum_{n=1}^N \#N_{r,n}}$ converges to a distribution function $F(x_p)$ then $H_N(x_p)$ will also converge to a distribution function.

Based on the definitions and introduction to our problem, we give the following theorem.

**Theorem 6.1:** Given Assumptions 6.1 and 6.2, $\lim_{N \to \infty} H_N(x_p) = H(x_p)$, where $H(x_p)$ is a Cumulative Distribution Function (CDF).

To verify Assumption 4(b), we first give an assumption.

**Assumption 6.3:** Define $Q(N) = \sum_{i=1}^N \{|f_i(\gamma) - f_i(\gamma^*)|^2\}$ where $\gamma \neq \gamma^*$. Assume $Q(N) \in \Theta(N)$ which means that $Q(N)$ is bounded above and below by $N$ asymptotically.

Next, we show the estimates in the fixed-source case is convergent.

**Theorem 6.2:** With Assumption 6.3 and Theorem 6.1, the least squares estimator $\hat{\gamma}$ is convergent to $\gamma$ in the fixed-source case.

Next, we consider the convergence of the fixed-source case with autocorrelated errors

$$y_i = f(x_i; \gamma) + \epsilon_i, i = 1, \ldots, N$$

The only difference from the fixed-source case is that $\epsilon_1, \ldots, \epsilon_N$ are autocorrelated, $\epsilon \sim (0, \sigma^2 I)$. Since the parameters we are interested are only $\gamma$, we only consider the convergence of $\gamma$. The convergence analysis can be conducted by converting generalized least squares estimation to ordinary least squares estimation and then the convergence conditions in ordinary least squares estimation will apply. More specifically, according to Cholesky decomposition $\sigma^2 I = (\sigma_x B)(\sigma_x B)^\top$, if we multiply both sides of the matrix form of the equation (16) with $(\sigma_x B)^\top$, we get

$$y^* = \Gamma(X; \gamma) + \epsilon^*$$

where $y^* = (\sigma_x B)^\top y, \Gamma(X; \gamma) = (\sigma_x B)^\top R(X; \gamma) + \epsilon^* = (\sigma_x B)^\top \epsilon \sim (0, I)$, where $I$ is an identity matrix. Then we can use the same methods and assumptions as in the fix-source case to conduct the convergence analysis.

**References**


