Statistica Sinica Preprint No: SS-2015-0249R2	
Title	Gaussian Process Modeling with Boundary Information
Manuscript ID	SS-2015-0249.R2
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202015.0249
Complete List of Authors	Matthias Hwai Yong Tan
Corresponding Author	Matthias Hwai Yong Ta
E-mail	matthtan@cityu.edu.hk
Notice: Accepted version subject to English editing.	

3. GP Modeling with Boundary Information

This section describes our proposed modification of the GP model to incorporate boundary information given by

$$\lim_{\mathbf{x}^i \to \mathbf{c}^i} [f(\mathbf{x}) - a^i(\mathbf{x})] = 0, i = 1, \ldots, k, \qquad (3.1)$$

where $\mathbf{x}^i = (x_{i_1}, \ldots, x_{i_{l(i)}}), \{i_1, \ldots, i_{l(i)}\} \subset \{1, \ldots, d\}$, $\mathbf{c}^i = (c^i_1, \ldots, c^i_{l(i)}) \in (\mathbb{R}_+ \cup \{\infty\})^{l(i)}$, and a^i is continuous. We assume all $\mathbf{c}^i, i = 1, \ldots, k$ are *distinct*, i.e., if $x_{i_h} = x_{j_m}$ for $i \neq j$, then $c^i_h \neq c^j_m$. This is typically on a boundary of the region $\mathbb{R}_+^{l(i)}$ in which $\mathbf{x}^i$ is defined, i.e., it has components that equal 0 and ∞. Thus, $\{\mathbf{x} \in (\mathbb{R}_+ \cup \{\infty\})^d : \mathbf{x}^i = \mathbf{c}^i\}$ is a subset of the boundary of $\mathbb{R}_+^d$. In some engineering problems, an input may be defined over a region of the form $[0, \infty)$ (e..g, the percentage of a substance in a mixture). In those cases, $\mathbf{c}^i$ is also often on the boundary of the region in which $\mathbf{x}^i$ is defined. The case $\lim_{\mathbf{x}^i \to \mathbf{c}^i} f(\mathbf{x}) = \infty$ can be transformed into a special case of (3.1) by using the transformation $1 - \exp(-ey)$ for some e that may be estimated with data. Unlike $1/y$, this transformation allows us to handle cases where there exists a vertical asymptote $\lim_{\mathbf{x}^i \to \mathbf{c}^i} f(\mathbf{x}) = \infty$ and a zero horizontal asymptote $\lim_{\mathbf{x}^j \to \mathbf{c}^j} f(\mathbf{x}) = 0$. Although a $\mathbf{c}^i$ with a component that equals ∞ can be far from the region of interest $\mathcal{X}$, it is often the case that f converges quickly enough so that the information in (3.1) is useful for improving prediction within $\mathcal{X}$. Moreover, the information is useful for improving extrapolation accuracy.

As reviewed in Section 2, stationary GP models can incorporate information about known function values at a finite set of points (which is the kind of information obtained in a computer experiment). Thus, when $a^i = 0$ and $\mathbf{c}^i \in \mathbb{R}_+^d$ (does not have ∞ as a component), we can incorporate the information in (3.1) by including $(\mathbf{c}^i, a^i(\mathbf{c}^i)), i = 1, \ldots, k$ as data points. However, the information in (3.1) cannot be easily incorporated in the stationary GP model when $\mathbf{x}^i$ is a proper subset of $\mathbf{x}$ or when some components of $\mathbf{c}^i$ is ∞. For example, when $d = 2$ and $\lim_{x_1 \to \infty} f(x_1, x_2) = a$, the value of y on an entire ray is known. Even if we update the GP model with such data points $(\mathbf{x}, y)$ of the form $((0, x_2), a)$, the GP model can still fail to predict accurately on the edge $x_1 = 0$. Moreover, since we know y converges to a constant as $x_1 \to \infty$, we have decreasing uncertainty about the function f as $x_1 \to 0$. This fact is not taken into account in a stationary GP model because the prior variance is constant. As another example, suppose x_1 represents time, and $\lim_{x_1 \to \infty} f(x_1, x_2) = a$, i.e., the steady state value of y is known. There is no existing method for exploiting this information. If a stationary GP model is used, the posterior mean when $x_1 \to \infty$ is the prior mean and the posterior variance when $x_1 \to \infty$ is the prior variance. This gives poor extrapolation behavior.

We propose to incorporate the information given by (3.1) into the BMGP model by