

УДК 519.226

Е. В. Бурнаев, П. В. Приходько

PreMoLab, Московский физико-технический институт (государственный университет)
Институт проблем передачи информации РАН
Datadvance

Методология построения суррогатных моделей для аппроксимации пространственно неоднородных функций

В работе предлагается общая методология для построения суррогатных моделей. Применение предложенной методологии проиллюстрировано на примере построения суррогатных моделей ограничений в задаче оптимизации веса обшивки корпуса самолета. Полученные результаты свидетельствуют об эффективности предложенного подхода.

Ключевые слова: суррогатное моделирование, регрессия, аппроксимация, оптимизация, смесь экспертов.

1. Введение

В задачах концептуального проектирования для экономии времени бывает необходимо заменить точную, но вычислительно сложную, физическую модель на быстро вычисляемую суррогатную модель [1]. Суррогатная модель аппроксимирует зависимость, реализуемую рассматриваемой физической моделью, с необходимой для последующего использования точностью, и строится по данным (конечному множеству пар “набор входных параметров” — “значения выходных характеристик физической модели для набора входных параметров”).

Примером индустриальной задачи, для решения которой востребовано суррогатное моделирование, является задача минимизации веса обшивки корпуса самолета (см. раздел 4) по параметрам, задающим структуру элементов обшивки. Целевая функция (вес обшивки) в этой задаче может быть выписана в явном виде, и поэтому возможен вычислительно эффективный подсчет ее значений. Область возможных значений оптимизируемых параметров (пространство дизайна) задается ограничениями (имеющими вид неравенств) на факторы прочности элементов обшивки. Так как факторы прочности подсчитываются с помощью специальной физической модели на основе метода конечных элементов, то решение соответствующей оптимизационной задачи является крайне времязатратным процессом, ускорение которого возможно только за счет построения суррогатных моделей факторов прочности для эффективного вычисления их значений и чувствительностей к оптимизируемым параметрам.

Важной особенностью задач построения суррогатных моделей ограничений в оптимизационном процессе является то, что в тех областях пространства дизайна, где значения факторов прочности близки к соответствующим пороговым значениям, суррогатные модели должны обеспечивать более высокую точность прогноза значений факторов прочности. В остальных областях пространства дизайна достаточно гораздо менее точного предсказания значений факторов прочности.

Другой распространенной особенностью задачи построения суррогатных моделей является пространственная неоднородность аппроксимируемой зависимости, вызванная “нестационарностью” поведения исходной физической модели. Например, в задаче предсказания коэффициента сопротивления формы характер поведения коэффициента сопротивления существенно меняется в зависимости от значения числа Маха при переходе от дозвукового режима полета к сверхзвуковому [2]. В описанной выше задаче построения суррогатных

моделей факторов прочности эта особенность также встречается и обусловлена нелинейностью процесса сминания элементов обшивки корпуса самолета.

Так как зачастую размерность пространства входных параметров велика (например, в задаче аппроксимации факторов прочности элементов обшивки самолета размерность равна 20), а аппроксимируемая зависимость имеет существенно нелинейное и нерегулярное поведение, то применение стандартных методов построения аппроксимации (Kriging, Artificial Neural Networks, Support Vector Regression, Multivariate Nonparametric Regression, Polynomial Regression и т.п. [5]) не позволяет достичь требуемой точности, см. также работы авторов [6, 7, 9].

Для учета указанных выше особенностей при построении суррогатной модели в литературе рекомендуется использовать смеси экспертов [3]. В рамках этого подхода пространство дизайна разбивается на несколько областей, далее решаются задачи построения аппроксимации рассматриваемой зависимости, ограниченной на полученные области, после чего итоговая суррогатная модель строится с помощью “сшивки” аппроксимаций.

В данной работе приводится описание предложенной авторами общей методологии построения суррогатных моделей, основанной на идеях метода смеси экспертов и позволяющей эффективно решать многие задачи инженерного проектирования, в том числе и упомянутые выше.

Данная методология дает возможность получить обладающую свойством непрерывности суррогатную модель, которая обеспечивает повышенную точность в заданных областях пространства дизайна и позволяет моделировать пространственно неоднородные функции.

Статья имеет следующую структуру.

- В разделе 2 приводится постановка решаемой задачи, а также описываются основные проблемы, которые возникают при построении суррогатных моделей.
- В разделе 3 описываются предложенная методология построения суррогатных моделей и конкретные процедуры, которые были использованы авторами на основных этапах процесса построения суррогатных моделей.
- В разделе 4 приводятся результаты применения предложенного подхода к реальной индустриальной задаче минимизации веса обшивки корпуса самолета.
- В разделе 5 приводятся основные выводы.

2. Постановка задачи

Итак, построение суррогатной модели необходимо для того, чтобы получить более быстро вычислимую аппроксимацию исходной зависимости. Поэтому в первую очередь приведем общую постановку задачи построения аппроксимации, а затем сформулируем требования, которые мы предъявляем к суррогатной модели.

2.1. Постановка задачи аппроксимации неизвестной зависимости

Пусть задана обучающая выборка

$$S_{learn} = \{(\mathbf{x}_i, y_i), i = 1, \dots, N_{learn}\}, \quad (1)$$

точки которой порождены случайно и независимо так, что между выходным значением (выходом) $y_i \in \mathbf{Y} \subset \mathbb{R}^1$ и входным вектором (входом) $\mathbf{x}_i \in \mathbf{X} \subset \mathbb{R}^p$ есть некоторая неизвестная функциональная зависимость $y_i = f(\mathbf{x}_i)$, при этом все входные векторы различны.

Задача восстановления неизвестной зависимости состоит в построении такой регрессионной зависимости (аппроксиматора) $\hat{f}(\mathbf{x}) = \hat{f}(\mathbf{x}|S_{learn})$ для исходной зависимости $f(\mathbf{x})$ по обучающей выборке S_{learn} , что для всех $\mathbf{x} \in \mathbf{X}$ (не только для $\mathbf{x} \in S_{learning}$) имеет место соотношение

$$\hat{f}(\mathbf{x}) \approx f(\mathbf{x}), \quad (2)$$

то есть аппроксиматор $\hat{f}(\mathbf{x})$ хорошо воспроизводит исходную зависимость. Считается, что соотношение (2) выполнено, если на независимой тестовой выборке $S_{test} = \{(\mathbf{x}_j, y_j), j = 1, \dots, N_{test}\}$ величина

$$\hat{e}_{S_{test}}(f, \hat{f}) = \frac{1}{N_{test}} \sum_{j=1}^{N_{test}} (y_j - \hat{f}(\mathbf{x}_j))^2 \quad (3)$$

мала.

Для того чтобы критерий (3) качества аппроксимации имел смысл, входные векторы выборок S_{learn} и S_{test} должны быть порождены одним и тем же распределением и достаточно плотно “покрывать” множество \mathbf{X} .

На практике при построении $\hat{f}(\mathbf{x})$ часто необходимо учесть дополнительные требования предметной области и особенности источника данных $y = f(\mathbf{x})$, породившего выборку S_{learn} .

2.2. Специфические требования по точности

Для конкретной задачи могут существовать разные требования на точность суррогатной модели в разных областях пространства дизайна \mathbf{X} . Например, в задаче построения суррогатных моделей для факторов прочности, упомянутой во введении к данной статье, требуется обеспечить высокую точность предсказания для области $\tilde{\mathbf{X}} \subset \mathbf{X}$, где $\tilde{\mathbf{X}} = \{\mathbf{x} : f(\mathbf{x}) \in (1 - \varepsilon, 1 + \varepsilon)\}$, $\varepsilon = 0.2$. По этой причине необходимо независимо строить аппроксимацию только в области $\tilde{\mathbf{X}}$ по подвыборке $\tilde{S} = \{(\mathbf{x}, y) \in S_{learn} : \mathbf{x} \in \tilde{\mathbf{X}}\}$, и потом “сшивать” ее с аппроксимацией, построенной в области $\mathbf{X} \setminus \tilde{\mathbf{X}}$ по подвыборке $S_{learn} \setminus \tilde{S}$, чтобы добиться более высокой точности предсказания именно для $\mathbf{x} \in \tilde{\mathbf{X}}$. Действительно, выделение подобласти $\tilde{\mathbf{X}}$ явным образом ограничивает изменчивость аппроксимируемой зависимости при $\mathbf{x} \in \tilde{\mathbf{X}}$, что позволяет строить более точную аппроксимацию.

2.3. Пространственная неоднородность выборки

При декомпозиции пространства дизайна и выделении областей (см. подраздел 2.2), соответствующих разным требованиям на точность суррогатной модели, может получиться, что большинство из этих областей будут представлять собой объединения некоторых несвязных множеств. Соответствующие подвыборки также будут представлять собой объединения некоторых кластеров точек.

Очевидно, что глобальная суррогатная модель, построенная по выборке, представляющей собой объединение некоторых кластеров точек, в общем случае будет иметь низкую точность. Поэтому бывает оправдано предварительно дополнительно разбить такую выборку на несколько однородных подвыборок, каждая из которых лежит в связной области пространства дизайна, после чего строить локальные аппроксимации уже только по соответствующим подвыборкам.

2.4. Избыточность в данных

Еще одним существенным моментом является то, что набор входных параметров задачи может быть избыточным в одном из двух смыслов.

- Входные данные могут быть связаны между собой. В самом простом случае это означает то, что несколько входных переменных значительно скоррелированы.
- Функция может зависеть не от всех входов. Здесь имеет смысл рассмотреть два основных сценария:
 - функция слабо зависит от нескольких входов,

– функция зависит от проекции входов на некоторое линейное подпространство.

Обнаружение и устранение такого типа избыточности в общем случае приводит к существенному улучшению качества построенных суррогатных моделей.

3. Методология построения суррогатной модели

Пусть задана некоторая обучающая выборка S_{learn} (1). Кроме того, выберем некоторую технику построения базовых моделей g из некоторого параметрического семейства G (под семейством могут пониматься модели типа Kriging, Artificial Neural Networks, Support Vector Regression, Multivariate Nonparametric Regression, Polynomial Regression и т.п.), которая позволяет по заданной выборке S_{learn} построить аппроксимацию $y = g(\mathbf{x}, \theta)$, где θ — параметры модели. Под построением модели здесь понимается выбор для фиксированного семейства G и выборки S_{learn} -значений параметров θ , которые задают некоторый элемент $g = g(\mathbf{x}, \theta) \in G$, оптимальный в смысле критерия (3) (см. [4]).

Предлагаемый процесс построения суррогатной модели состоит из следующих этапов.

- 1) Для того чтобы избавиться от избыточности в признаках, проводится снижение размерности данных. Применение методов снижения размерности для предобработки данных не будет подробно рассмотрено в данной статье. Детальное описание задачи, а также предложенный авторами эффективный алгоритм снижения размерности приведены в [10].
- 2) Декомпозиция пространства на области $\mathbf{X} = \cup_{j=0}^{N_y-1} \mathbf{X}_j$, которым соответствуют различные диапазоны значений выходов. Такая декомпозиция полезна, если требуется для разных диапазонов значений выходов обеспечить разную точность приближения. Детальное описание процесса приведено в подразделе 3.1.
- 3) Разбиение областей \mathbf{X}_j на связные подобласти $\mathbf{X}_{j,k}$, соответствующие более регулярному поведению аппроксимируемой зависимости. Детальное описание процесса приведено в подразделе 3.2.
- 4) Построение по подвыборкам $S_{j,k} = \{(\mathbf{x}, y) \in S_{learn} : \mathbf{x} \in \mathbf{X}_{j,k}\}$ соответствующих аппроксимирующих моделей.
- 5) Построение классификатора, который “оценивает близость” от данной точки \mathbf{x} до подобластей \mathbf{X}_j . Детальное описание процесса приведено в подразделе 3.3.
- 6) Построение итоговой суррогатной модели с помощью “сшивки” полученных аппроксимирующих моделей. Детальное описание процесса приведено в подразделе 3.3.

3.1. Декомпозиция пространства дизайна на области по значениям выходов

Декомпозицию пространства дизайна на области по значениям выходов можно описать следующим образом.

- 1) Пусть y_{\min} и y_{\max} — это верхняя и нижняя границы на значение выхода $y = f(\mathbf{x})$. Интервал изменения значения $y \in [y_{\min}, y_{\max}]$ разбивается на N_y подынтервалов, т.е. $[y_{\min}, y_{\max}] \in \cup_{j=0}^{N_y-1} [y^{2j}, y^{2j+1}]$, где $y^0 = y_{\min}$, $y^{2N_y-1} = y_{\max}$ и $y_{2j} < y_{2j+1}$, $y_{2j+2} < y_{2j+1}$, $j = 0, 1, \dots, N_y - 1$. Условия на концы интервалов гарантируют непустое пересечение интервалов, что позволяет обеспечить непрерывную “сшивку” соответствующих аппроксимационных моделей (см. подраздел 3.3). Выбор соответствующего разбиения делается экспертом на основании знаний предметной области и требований задачи. Такому разбиению соответствует разбиение пространства дизайна $\mathbf{X} = \cup_{j=0}^{N_y-1} \mathbf{X}_j$, где $\mathbf{X}_j = \{\mathbf{x} : f(\mathbf{x}) \in [y^{2j}, y^{2j+1}]\}$.

- 2) Выборка S_{learn} разделяется на N_y подвыборок $S_{learn} = \bigcup_{j=0}^{N_y-1} S_j$ так, что $S_j = \{(\mathbf{x}, y) \in S_{learn} : y \in [y^{2j}, y^{2j+1}]\}$. На основе каждой подвыборки S_j строится аппроксиматор $f_{approx}^j(\mathbf{x})$, который может быть как моделью из G , так и иметь более сложную структуру, см. подраздел 3.2.

Отметим, что описанная декомпозиция пространства дизайна на области по значениям выходов полезна не только в случае, если имеются разные требования на точность суррогатной модели в разных областях пространства дизайна \mathbf{X} . Действительно, если функция $f(\mathbf{x})$ существенно пространственно неоднородна, то в областях $\mathbf{X}_j \subset \mathbf{X}$ изменчивость функции $f(\mathbf{x})$ существенно ниже, чем во всем пространстве дизайна \mathbf{X} . Значит, если построить аппроксимации для областей $\mathbf{X}_j \subset \mathbf{X}$ и “сшить” эти аппроксимации, то можно получить более точную суррогатную модель по сравнению с глобальной суррогатной моделью, построенной сразу для всего пространства дизайна \mathbf{X} .

3.2. Декомпозиция областей на подобласти по значениям входов

Итак, аппроксиматор $f_{approx}^j(\mathbf{x})$ для $\mathbf{x} \in \mathbf{X}_j$ (см. подраздел 3.1) может быть построен по выборке S_j как базовая модель из G , однако если выборка S_j существенно пространственно неоднородна или даже представляет собой несколько разделенных облаков (кластеров) точек, то построенный по такой неоднородной выборке аппроксиматор в общем случае будет иметь низкую точность. Для того чтобы увеличить точность, предлагается

- дополнительно разбить области \mathbf{X}_j на такие связные подобласти $\mathbf{X}_j = \bigcup_{k=1}^{N_x} \mathbf{X}_{j,k}$, что
 - подвыборки из выборки S_j , принадлежащие этим связным подобластям, более однородны,
 - в пределах подобластей $\mathbf{X}_{j,k}$ аппроксимируемая зависимость имеет, вообще говоря, более регулярное поведение;
- построить отдельную локальную аппроксимационную модель $g_{j,k} \in G$ для каждой подобласти $\mathbf{X}_{j,k}$, а итоговую аппроксимационную модель $f_{approx}^j(\mathbf{x})$ представить в виде смеси этих локальных моделей.

В данной работе для построения разбиения $\mathbf{X}_j = \bigcup_{k=1}^{N_x} \mathbf{X}_{j,k}$ на основе выборки S_j , $j = 0, \dots, N_y - 1$, предлагается использовать смесь гауссовских распределений [5] следующим образом. Предполагается, что точки $(\mathbf{x}, y) \in S_j$ порождены в соответствии с распределением

$$\text{Law}(\mathbf{x}, y) = \sum_{k=1}^{N_x} \alpha_k \mathcal{N}(\mu_k^j, \Theta_k^j), \quad (4)$$

где μ_k^j и Θ_k^j — среднее и матрица ковариации k -го нормального распределения смеси для j -й выборки S_j . Следовательно, если рассматривать только входной вектор \mathbf{x} , то $\text{Law}(\mathbf{x}) = \sum_{k=1}^{N_x} \alpha_k \mathcal{N}(\mu_{k,x}^j, \Theta_{k,x}^j)$, где $\mu_{k,x}^j$ и $\Theta_{k,x}^j$ — подвектор средних значений вектора μ_k^j и подматрица ковариаций матрицы Θ_k^j соответственно, задающие безусловное распределение вектора входов $x \in \mathbf{X}_j$.

Параметры гауссовской смеси оцениваются по выборке S_j , например, с помощью EM-алгоритма [5], и в рамках полученной модели осуществляется дополнительная кластеризация выборки S_j . Отметим, что использование при кластеризации не только значений входов, но и значений выходов позволяет учитывать возможную пространственную неоднородность функции при $\mathbf{x} \in \mathbf{X}_j$. При этом оцененные параметры гауссовской смеси будут фактически задавать разбиение $\mathbf{X}_j = \bigcup_{k=1}^{N_x} \mathbf{X}_{j,k}$. Действительно, определим расстояние в метрике Махаланобиса от центра k -го кластера до точки \mathbf{x} в соответствии с формулой $d_{j,k}(\mathbf{x}) = (\mathbf{x} - \mu_{k,x}^j)^T (\Theta_{k,x}^j)^{-1} (\mathbf{x} - \mu_{k,x}^j)$, тогда множество $\mathbf{X}_{j,k} = \{\mathbf{x} \in \mathbf{X}_j : d_{j,k}(\mathbf{x}) \leq \chi_{97\%}^2(p)\}$,

где $\chi_{97\%}^2(p)$ является 97%-квантилью распределения χ^2 с p степенями свободы. Локальные аппроксимационные модели $g_{j,k} \in G$ для каждой подобласти $\mathbf{X}_{j,k}$ строятся по подвыборкам $S_{j,k} = \{(\mathbf{x}, y) \in S_{learn} : \mathbf{x} \in \mathbf{X}_{j,k}\}$, $k = 1, \dots, N_x$.

Определим степень вхождения точки \mathbf{x} в k -й кластер согласно формуле

$$w(k|\mathbf{x}, j) = \frac{1}{2} \left(\tanh \left(1 - 2 \frac{d_{j,k}(\mathbf{x}) - \chi_{97\%}^2(p)}{\chi_{99\%}^2(p) - \chi_{97\%}^2(p)} \right) + 1 \right), \quad (5)$$

где $\chi_{99\%}^2(p)$ и $\chi_{97\%}^2(p)$ являются 99%- и 97%-квантилями распределения χ^2 с p степенями свободы соответственно. В рамках модели (4) классификатор, определяющий степень принадлежности точки \mathbf{x} к k -му кластеру, может строиться в соответствии с формулой

$$\hat{w}(k|\mathbf{x}, j) = \frac{w(k|\mathbf{x}, j)}{\sum_r w(r|\mathbf{x}, j)}. \quad (6)$$

Итоговое предсказание получается в виде гладкой сшивки локальных моделей $g_{j,k}$ по всем кластерам, к которым точка принадлежит, и вычисляется согласно формуле

$$f_{approx}^j(\mathbf{x}) = \sum_{k=1}^{N_x} \hat{w}(k|\mathbf{x}, j) g_{j,k}(\mathbf{x}),$$

где $g_{j,k}(\mathbf{x})$ — локальный аппроксиматор в k -м кластере, построенный по подвыборке $S_{j,k}$.

Эксперименты показывают, что такой подход оказывается устойчивее/точнее по сравнению с подходом, где в качестве весов (6) используются апостериорные вероятности $\hat{w}(k|\mathbf{x}, j) = P(k|\mathbf{x}, j)$ точки \mathbf{x} принадлежать соответствующим кластерам.

3.3. Построение классификатора и вычисление выходного значения суррогатной модели

В качестве классификатора, с помощью которого будут выбираться аппроксимационные модели $f_{approx}^j(\mathbf{x})$, $j = 0, \dots, N_y - 1$, для построения прогноза $\hat{f}(\mathbf{x})$, предлагается использовать аппроксиматор $f_{class}(\mathbf{x}) \in G$, построенный по всей выборке S_{learn} . В таком случае прогноз $\hat{f}(\mathbf{x})$ для заданного входного вектора \mathbf{x} получается следующим образом.

- Подсчитываем значение $f_{class}(\mathbf{x})$. Обозначим через $\#(A)$ — мощность множества A , $J(\mathbf{x}) = \{j \in (0, \dots, N_y - 1) : f_{class}(\mathbf{x}) \in [y_{2j}, y_{2j+1}]\}$.
- Если $\#(J(\mathbf{x})) = 1$, то подсчитываем прогноз согласно формуле $\hat{f}(\mathbf{x}) = f_{approx}^{J(\mathbf{x})}(\mathbf{x})$.
- Если $\#(J(\mathbf{x})) = 2$, то, обозначив через $j \in J(\mathbf{x})$ наименьшее из двух значений индексов, принадлежащих множеству $J(\mathbf{x})$, подсчитываем прогноз согласно формуле

$$\hat{f}(\mathbf{x}) = f_{approx}^j(\mathbf{x}) \left(1 - w \left(\frac{f_{approx}^{j+1}(\mathbf{x}) - y_{2j+2}}{y_{2j+1} - y_{2j+2}} \right) \right) + f_{approx}^{j+1}(\mathbf{x}) w \left(\frac{f_{approx}^{j+1}(\mathbf{x}) - y_{2j+2}}{y_{2j+1} - y_{2j+2}} \right),$$

где $w(x) = 3x^2 - 2x^3$. Так определенные веса позволяют обеспечить гладкую “сшивку” моделей $f_{approx}^j(\mathbf{x})$ и $f_{approx}^{j+1}(\mathbf{x})$.

4. Пример

Аэродинамические структуры (обшивка корпуса самолета) в основном состоят из панелей, укрепленных ребрами жесткости как в продольном, так и в поперечном направлениях. Пример аэродинамической структуры приведен на рисунке 1.

Для удобства такая структура представляется в виде набора базовых элементов, состоящих из условного объединения ребра жесткости и двух боковых панелей. Эти базовые

элементы подвержены таким нелинейным эффектам, как коробление, потеря устойчивости и т.п. Для того чтобы определить оптимальный размер базовых элементов (базовые элементы должны иметь минимальный вес), необходимо контролировать устойчивость элементов, подсчитывая статические механические критерии прочности с применением специального программного обеспечения, основанного на нелинейных вычислениях.

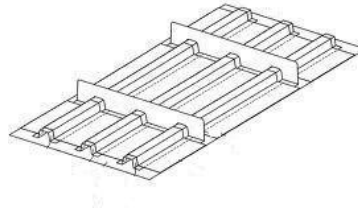


Рис. 1. Панель с ребрами жесткости

Таким образом, на каждой итерации алгоритма оптимизации массы аэродинамической структуры необходимо проводить анализ размеров всей структуры в соответствии со следующей двухуровневой схемой: на глобальном уровне модель на основе метода конечных элементов позволяет рассчитать внутренние нагрузки, приложенные к каждому базовому элементу; на локальном уровне вычисленные нагрузки используются для того, чтобы вычислить статические механические критерии и их чувствительности к параметрам базовых элементов.

Формальная постановка решаемой задачи может быть записана в виде

$$M(\mathbf{z}) \rightarrow \min_{\mathbf{z} \in \mathbb{R}^n}, \begin{cases} \mathbf{z}_{\min} \leq \mathbf{z} \leq \mathbf{z}_{\max}, \\ RF_{i,j,l}(N(\mathbf{z}), \mathbf{z}) \geq 1, \quad i = 1 \dots N_e, \quad j = 1 \dots N_l, \quad k = 1 \dots N_{fin}, \\ d_l(\mathbf{z}) \geq 1, \end{cases}$$

где

- \mathbf{z} — вектор длины n , задающий параметры конструкции,
- $M(\mathbf{z})$ — целевая функция, выражающая вес структуры,
- $N(\mathbf{z})$ — вектор, описывающий силы, действующие на структуру.

Ограничения в оптимизационной задаче (7) делятся на три типа:

- область определения параметров: $\mathbf{z}_{\min} \leq \mathbf{z} \leq \mathbf{z}_{\max}$,
- механические критерии устойчивости: $RF_{i,j,l}(N(\mathbf{z}), \mathbf{z}) \geq 1$,
- ограничения на допустимую геометрию конструкции: $d_l(\mathbf{z}) \geq 1$.

Индексы i, j, k для механических критериев указывают на то, что во время каждой итерации алгоритма оптимизации необходимо подсчитывать значения механических критериев устойчивости для каждой из N_e панелей, N_l действующих сил и N_{fin} типов отказа.

Большинство статических механических критериев задается факторами прочности: структура считается устойчивой, если значения всех факторов прочности больше 1. Как уже было сказано выше, детальный дизайн обшивки корпуса самолета требует двухуровневого цикла [6, 7]. Сначала изменения в локальной геометрии на предыдущей итерации приводят к новому распределению нагрузок по всей структуре. Затем изолированно для каждой панели вычисляются механические критерии прочности и их чувствительности в зависимости от внутренних нагрузок и геометрических размеров. Этот двухуровневый локально-глобальный процесс повторяется до тех пор, пока распределение нагрузок не

стабилизируется на всей структуре. Для оптимизации параметров панелей применяются локальные градиентные методы оптимизации. Использование градиентных методов обусловлено огромной размерностью задачи ($O(10^4)$ переменных и $O(10^5)$ ограничений).

Применение локальных градиентных методов требует знания градиентов ограничений, которые при наличии лишь “закрытой” физической модели (реализованной в виде черного ящика для подсчета значений факторов прочности) можно получить только методом конечных разностей. Одно обращение к черному ящику может занимать до 1 секунды времени, и, как следствие, необходимость применять метод конечных разностей для каждой из многочисленных панелей очень существенно увеличивает время одного шага полного цикла оптимизации всей структуры. Это приводит к тому, что такого рода оптимизация является времязатратным процессом.

Вычисление значения выхода соответствующей суррогатной модели занимает на порядок меньше времени, чем обращение к черному ящику, поэтому использование суррогатных моделей может привести к значительной экономии времени [6, 7]. Кроме того, суррогатные модели дают непрерывно дифференцируемую аппроксимацию факторов прочности (которые сами, вообще говоря, могут демонстрировать нерегулярное поведение). Такое свойство суррогатных моделей обеспечивает лучшую сходимость оптимизационного процесса. В итоге использование суррогатных моделей позволяет достичь две цели, важные для инженеров:

- уменьшение времени проведения предварительного проектирования изделия,
- получение гладкого приближения рассматриваемой зависимости.

Предметная область налагает жесткие ограничения на время работы суррогатной модели и ее точность: для того чтобы применение модели было оправдано, 95%-квантиль относительной ошибки предсказания не должна превышать 5% для значений факторов прочности порядка 1, а время вычисления предсказания модели в новой точке должно быть меньше 0.001 с.

В связи с высокой размерностью задачи (значение фактора прочности зависит от 20 параметров), а также существенно нелинейным и нерегулярным поведением зависимости, применение стандартных методов построения аппроксимации (Kriging, Artificial Neural Networks, Support Vector Regression, Multivariate Nonparametric Regression, Polynomial Regression и т.п. [5]) не позволяет достичь требуемой точности (см. также в [6, 7, 9] результаты сравнения точности суррогатной модели, построенной с помощью предложенного подхода, с точностью суррогатных моделей, построенных с помощью стандартных методов).

Построенная авторами суррогатная модель (MACROS) была использована в оптимизационном процессе предварительного дизайна композитных участков A350XWB, где она заменяла вычислительно дорогой инженерный код PS3, позволяющий рассчитывать механические критерии устойчивости для каждой из панелей на основе метода конечных элементов.

В результате оказалось, что использование модели MACROS позволяет получить более быструю сходимость к хорошему решению за меньшее количество итераций, а также приводит к уменьшению общего времени оптимизации с нескольких дней до нескольких часов.

Изменение целевой функции (массы конструкции) в зависимости от итерации приведено на рисунке 2:

- черная сплошная кривая (с маркерами в виде ромбов) соответствует запуску процесса оптимизации с использованием суррогатной модели MACROS,
- темно-серая сплошная кривая (с маркерами в виде треугольников) получена с использованием инженерного кода PS3 при запуске процесса оптимизации из точки оптимума, найденного при оптимизации с использованием суррогатной модели MACROS,

- черная пунктирная кривая (с маркерами в виде квадратов) соответствует запуску процесса оптимизации с использованием только инженерного кода PS3.

Можно видеть, что применение как суррогатной модели MACROS, так и кода PS3 дает хорошую сходимость, хотя процесс оптимизации с использованием суррогатной модели MACROS сходится быстрее (поскольку суррогатная модель имеет более регулярное поведение по сравнению с поведением исходной физической модели). Найденные в итоге оптимальные решения так же имеют близкие значения целевой функции (веса).

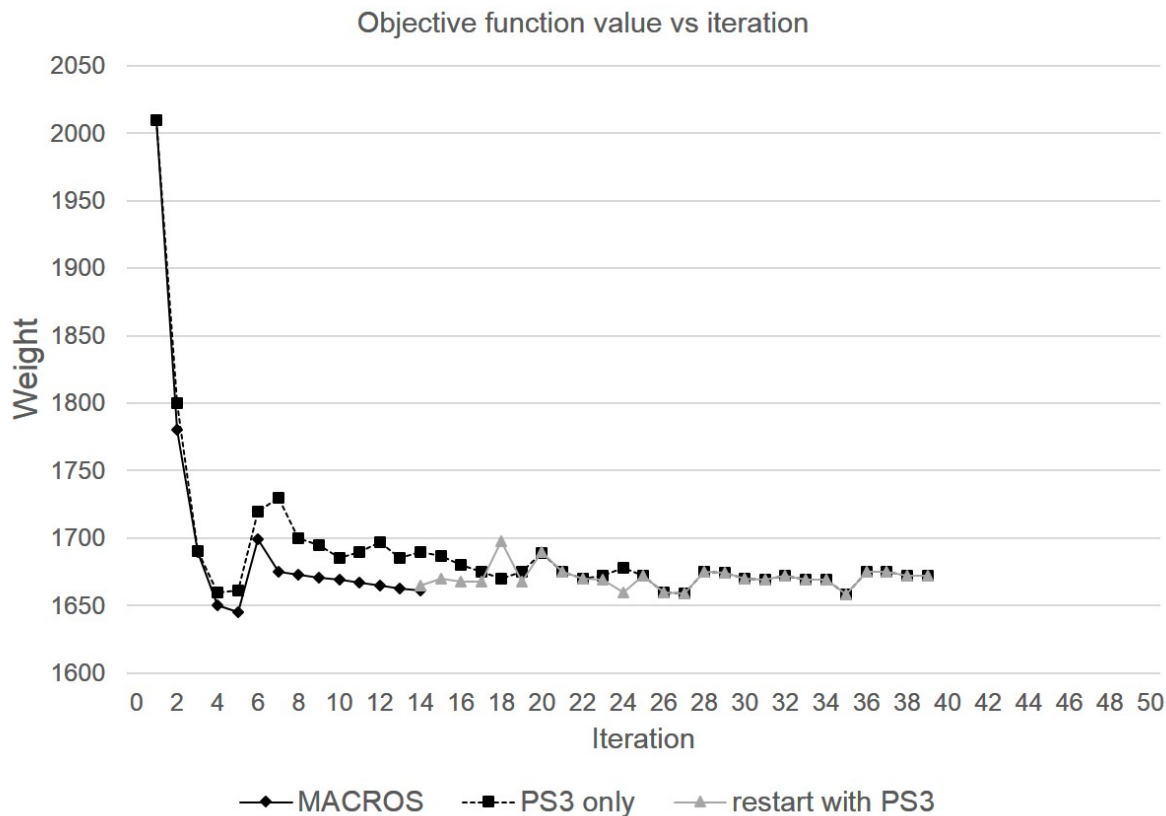


Рис. 2. Сравнение сходимости оригинального инженерного кода и построенной суррогатной модели

В точке оптимума, полученной в результате оптимизации с использованием суррогатной модели MACROS (точка оптимума MACROS), было произведено сравнение значений факторов прочности, предсказанных с помощью суррогатной модели MACROS, со значениями факторов прочности, вычисленных с использованием инженерного кода PS3. Результаты приведены на рис. 3: здесь на схеме слева приведены значения ограничений на разных участках крыла, полученные в точке оптимума MACROS, а на схеме справа приведены значения ограничений в этой же точке, пересчитанные в PS3. Из рисунков видно, что построенная модель обладает хорошей точностью, достаточной для применения модели на этапе предварительного проектирования (значения ограничений в точке оптимума MACROS, пересчитанные в PS3, соответствуют требованию, что ошибка суррогатной модели должна быть не больше 0.05 для 95% случаев). Достаточная точность модели также была подтверждена специалистами предметной области.

5. Вывод

В работе предложена общая методология построения суррогатных моделей. Применение предложенной методологии проиллюстрировано на примере решения задачи оптимизации веса обшивки корпуса самолета при заданных статических ограничениях на факторы проч-

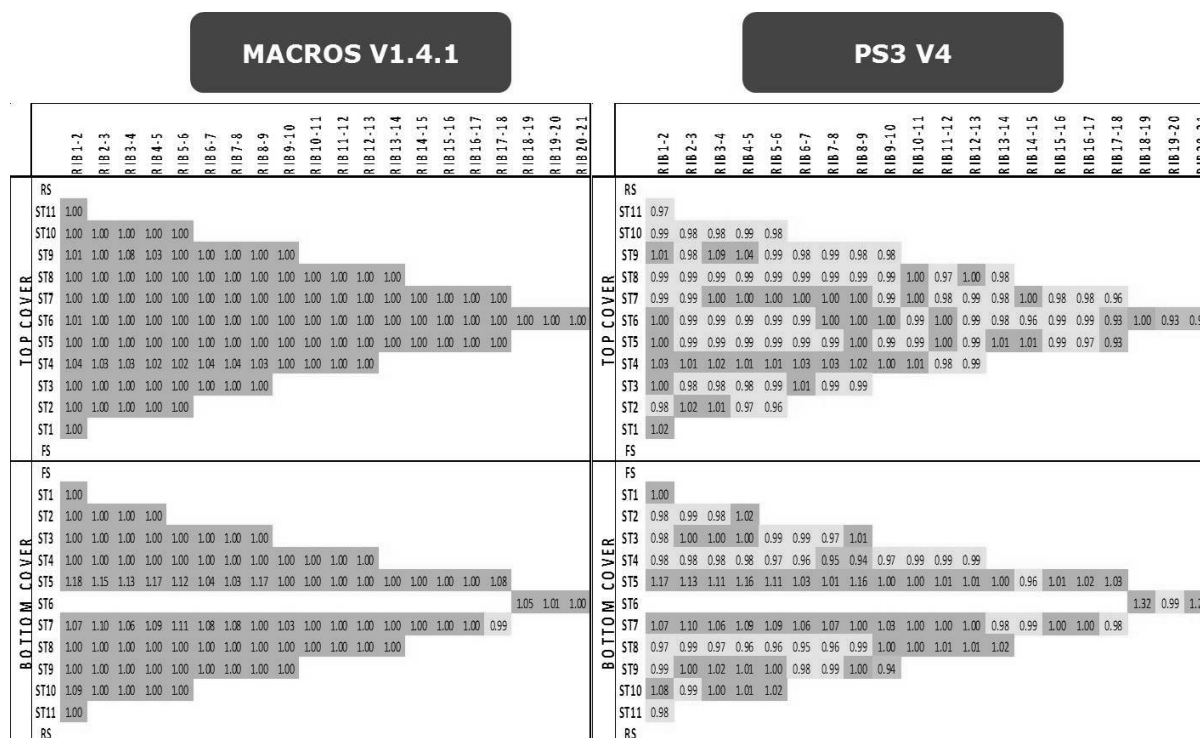


Рис. 3. Сравнение решений в точках оптимума для оригинального инженерного кода и построенной суррогатной модели

ности. Использование суррогатной модели, построенной с помощью предложенной методологии, позволило

- получить более быструю сходимость к хорошему оптимуму (с более гладким распределением толщин/размеров панелей) за меньшее число итераций. Нарушения ограничений на факторы прочности незначительные и могут быть легко исправлены на детальной стадии проектирования обшивки корпуса самолета;
- существенно уменьшить время оптимизации с нескольких дней до нескольких часов.

Работа выполнена при поддержке Лаборатории структурных методов анализа данных в предсказательном моделировании, МФТИ, грант правительства РФ, дог. 11.G34.31.0073.

Литература

1. *Kuleshov A.P., Bernstein A.V., Burnaev E.V.* Adaptive models of complex systems based on data handling // Proceedings of the 3rd International Conference on Inductive Modelling. Kyiv, Ukraine, 2010. P. 64–71.
2. *Schlichting H.* Boundary Layer Theory. McGraw-Hill, New York, 1968.
3. *Yuksel S.E., Wilson J. N., Gader P. D.* Twenty Years of Mixture of Experts // IEEE Transactions on Neural Networks and Learning Systems. — 2012. — V. 23, I. 8. — P. 1177–1193.
4. *Anthony M., Bartlett P.L.* Neural network learning: Theoretical foundations. Cambridge: Cambridge University Press, 1999.
5. *Hastie T., Tibshirani R., Friedman J.* The elements of statistical learning: data mining, inference, and prediction. Springer, 2008.

6. *Burnaev E.V., Grihon S.* Construction of the metamodels in support of stiffened panel optimization // Proceedings of the International Conference on Mathematical Methods in Reliability. Moscow, Russia. — 2009. — P. 124–128.
7. *Grihon S., Alestra S., Burnaev E., Prihodko P.* Surrogate Modeling of Buckling Analysis in Support of Composite Structure Optimization // Proceedings of the 1st International Conference on Composite Dynamics. Arcachon, France. — 2012.
8. *Burnaev E.V., Belyaev M.G., Prihodko P.V.* Algorithm for tuning parameters of approximation based on linear expansions in parametric dictionaries // Proceedings of the 8th International Conference on Intellectualization of Information Processing. Cyprus. — 2010. — P. 204–207.
9. *Grihon S., Alestra S., Bettebghor D., Burnaev E., Prihodko P.* Comparison of different techniques for surrogate modeling of stability constraints for composite structures // Proceedings of the 1st International Conference on Composite Dynamics. Arcachon, France. — 2012.
10. *Burnaev E., Prihodko P.* Approaches to dimensionality estimation in the effective dimensionality reduction based on gaussian processes // Information Technologies and Systems 2012. Petrozavodsk, August 2012. — P. 19–25.

Поступила в редакцию 11.12.2012