Применение метода ядерной гребневой оценки к задаче расчета аэродинамических характеристик самолета

Червоненкис А. Я.
Институт Проблем Управления РАН,
г. Москва,
Россия
chervnks@ipu.rssi.ru

Чернова С. С.
Институт ситемного анализа РАН
chernova@cpt-ran.ru

Зыкова Т. В.
Международный научно-исследовательский
институт проблем управления
zytanya@yandex.ru

Аннотация

При проектировании новых сложных технических объектов приходится просматривать очень большое число вариантов их компоновки, поэтому важнейшей проблемой, является проблема построения моделей, позволяющих быстро проводить массовые расчеты с требуемой точностью. В качестве таких моделей используется, так называемые, суррогатные модели (метамодели) основанные на данных.
В настоящем докладе описаны результаты применения метода Ядерной гребневой регрессии к задаче быстрого расчета аэродинамических характеристик самолета. Полученные результаты сравниваются с результатами, полученными с помощью суррогатных моделей, построенных с помощью искусственных нейронных сетей.

1 Введение

При проектировании самолета приходится сравнивать очень большое число вариантов их компоновки. Методы традиционно используемые для таких расчетов (CFD-коды, натуральные эксперименты в аэродинамических трубах) требуют значительного времени [3] и не могут использованы на этапе предварительного проектирования, когда рассматривается большое количество вариантов. Эффективным способом решения проблемы массовых расчетов является использование суррогатных моделей основанных на данных [1], [2]. Пример применения суррогатных моделей в задаче расета аэродинамических характеристик самолета описан в [5] и [6]. В этих работах для построения суррогатных моделей используются, в основном, искусственные нейронные сети.
В настоящем докладе к данным, которые были использованы для построения аэродинамических моделей в работах [5], [6] был применен метод Ядерной гребневой регрессии (ЯГР).
Метод Ядерной гребневой регрессии использует фактически те же формулы, что и давно известный метод кригинга [4]. Но в методе кригинга задача ставится как оценка значений случайной функции многих аргументов, тогда как в методе ЯГР от вероятностной интерпретации отказываются.

2 Кригинг

Метод кригинга решает следующую задачу: по заданным значениям случайной функции в точках $x_1, x_2, \dots, x_l$,оценить значение функции $u(x)$ в произвольной текущей точке x (в пределах ее области определения). Если критерием служит средний квадрат отклонения оценки от фактического значения функции, то наилучшей оценкой будет условное математическое ожидание $E(u(x)/u(x_1), u(x_2), \dots, u(x_l))$. В гауссовом случае, когда совместное распределение входных параметров и выходной величины является нормальным, условное математическое ожидание $u^*(x)$ значения функции $u(x)$ в текущей точке x будет линейной функцией от измеренных значений $u(x_1), \dots, u(x_l)$:

$$u^*(x) = \sum_{j=1}^{l} a_i u(x_i),$$

и проблема состоит в том, чтобы найти правильные значения весов как функцию взаимного расположения точек $x_1, x_2, \dots, x_l$ и текущей точки x. В гауссовом случае наша случайная зависимость полностью характеризуется математическим ожиданием $M(x) = Eu(x)$ и ковариационной функцией $R(x_1, x_2) = Eu(x_1)u(x_2)$. Математическое ожидание, не ограничивая общности, можно принять равным нулю. Достаточно часто в ковариационной функции выделяют регулярную составляющую

$R_0(x,y)$ непрерывную при $x=y$ и некоррелированную составляющую $D_0\delta(y-x)$, где $\delta(y-x)=1$ при $x=y$ и $\delta(y-x)=0$ в остальных случаях:

$$R(x,y)=R_0(x,y)+D_0\delta(y-x).$$

Обозначим через a вектор неизвестных коэффициентов $a=(a_1,\dots,a_l)$, через u вектор измеренных значений $u=(u(x_1),\dots,u(x_l))$, через r вектор коэффициентов ковариации между значениями функции в точках измерения и в текущей точке $r=(R(x_1,x),\dots,R(x_l,x))$ и через R_0 матрицу размерности $(l\times l)$ с элементами $r_{ij}=R_0(x_i,x_j)$, равными регулярной составляющей ковариации функции в точках измерения.
Тогда искомое значение вектора коэффициентов задается формулой

$$a_{\text{опт}}=r^T(R_0+D_0I)^{-1},$$

а наилучшее предсказание значения поля в точке x формулой

$$u^*(x)=r^T(R_0+D_0I)^{-1}u, \qquad (1)$$

где I – единичная матрица.
Если же отказаться от гипотезы нормальности, то эта оценка будет наилучшей линейной (по значениям $u(x_1),\dots,u(x_l)$) оценкой функции в текущей точке x. Как функция же координат точек $x_1,x_2,\dots,x_l$ эта оценка будет существенно нелинейной.

3 ЯГР

В методе Ядерной гребневой регрессии отказываются от статистической постановки задачи. Дело в том, что в многомерном случае ковариационную функцию $R(x,y)$ на самом деле приходится считать заданной априори, и взять ее не откуда. Поэтому в методе ЯГР просто задаются некоторым положительно определенным ядром $R(x,y)$ (обычно в форме $R(x-y)$), и ищут оценку значения функции в текущей точке в виде

$$u(x)=\sum_{i=1}^{l}a_iR(x,x_i)u_i,$$

где a_i – неизвестные коэффициенты, x_i - точки, в которых функция была измерена или посчитана, u_i - значения функции в этих точках. Считая далее, что поле было измерено с помехой, дисперсия которой равна D_0, полагают $R(x,y)=R_0(x,y)+D_0\delta(y-x)$ и применяют формулу метода кригинга (1).

Этот метод получил название ядерная гребневая регрессия (Kernel Ridge Regression). Добавочная матрица D_0I играет роль регуляризатора. Вид и параметры ядра и коэффициент регуляризации D_0 выбирают из тех или иных соображений.

4 Эксперимент

В качестве объекта рассматривается компоновка современного пассажирского самолета. Поскольку целью настоящей работы было сравнение результатов, полученных с применением ИНС с результатами получаемыми средствами ЯГР, то были использованы те же исходные данные, с помощью которых строились модели в [5], [6]. Эти данные содержали описание поверхности самолета в виде вектора большой размерности. С помощью генератора объектов было построено около 80 тыс. вариантов - 5 тыс. компоновок, для каждой их которых проводились расчеты для разных режимов полета, с помощью CFD-кодов, которые реализуются методами решения краевых залач. В настоящей работе использовалась только одна выходная характеристика - коэффициент подъемной силы (CL).
Кроме того, каждый объект x был снабжен значением $u_{\text{инс}}(x)$ оценки выходной величины, полученной ранее средствами ИНС. Исходное описание объекта содержало 444 параметров, большая часть которых относилась к описанию формы крыла и хвостового оперения. Применением метода главных компонент с удалением мало значимых компонент к этим данным число параметров в

описании объекта удалось сократить до 52.
В настоящей работе метод ЯГР применялся локально. Прежде всего, все данные были разбиты на 30 групп по 10300 объектов. Для этого было случайно выбрано 30 объектов, и каждая группа формировалась как 10300 точек ближайших в пространстве описаний к одному из выбранных объектов. В каждой группе проводилось сокращение размерности описания путем применения метода главных компонент. Все входные параметры центрировались и нормировались на среднеквадратичное отклонение (для каждой группы отдельно). Далее работа велась по этим группам раздельно.
В каждой группе было случайно выделено 300 объектов на тест. В тестовых точках значение выходной величины оценивалось по значениям остальных объектов группы (без использования тестовых точек). Для этого у каждой тестовой точки x в пространстве нормированных значений входных параметров находилась сферическая окрестность минимального радиуса, содержащая ровно 300 объектов группы и запоминался радиус $R(x)$ этой окрестности. Для оценки значения выходной величины $u^*(x)$ в точке x использовались только эти 300 объектов.
Самой грубой оценкой служило просто среднее арифметическое значение $u_{\text{ср}}(x)$ выходной величины по всем объектам окрестности. Более точная оценка $u_{\text{лин}}(x)$ получалась путем построения линейной (по входным параметрам) регрессии по данным обучения и подстановки в нее координат точки x. Метод Ядерной гребневой регрессии использовался для учета нелинейных эффектов. Для этого в каждой точке x_i, соответствующей i-тому объекту окрестности, вычислялось отклонение $\Delta(x_i)$ линейной оценки $u_{\text{лин}}(x_i)$ от фактического значения выходной величины $u(x_i)$:

$$\Delta(x_i) = u(x_i) - u_{\text{лин}}(x_i).$$

Метод ЯГР применялся для оценки этих отклонений от линейной регрессии по формуле (1). В результате получалась оценка отклонения $\Delta^*(x)$. Само же значение выходной величины в точке оценивалось как

$$u_{\text{ягр}}(x) = u_{\text{лин}}(x) + \Delta^*(x).$$

В качестве ядра $R_0(x, y)$ использовалась функция RBF

$$R_0(x, y) = exp[-|x - y|^2/r_0]$$

при значении параметра r_0 равном радиусу $R(x)$ выбранной окрестности. Попытки уменьшить или увеличить значение r_0 приводили только к ухудшению результата (в среднем).
Параметр регуляризации D_0 был подобран на первой группе данных и применялся без изменения для всех остальных. Оптимальным значением D_0 оказалось 10^{-11}. Столь малое значение параметра, соответствующего дисперсии помехи, видимо, связано с тем, что значение выходной величины в исходных данных вычислялось, а не измерялось, и дисперсия помехи связано только с погрешностями вычислений.

5 Результаты

Результаты, даваемые различными предсказателями, сравнивались с фактическим значением выходной характеристики для всех тестовых объектов. В качестве предсказателей использовались $u_{\text{ср}}(x)$ – среднее значение выходной величины в окрестности тестовой точки, $u_{\text{лин}}(x)$ – оценка по линейной (в окрестности) регрессии, $u_{\text{ягр}}(x)$ – оценка с применением метода ЯГР, а также $u_{\text{инс}}(x)$ – оценка, полученная ранее средствами ИНС.
Квадрат отклонения усреднялся по всем тестовым объектам каждой группы. В таблице 1 приведены значения среднеквадратичного уклонения предсказанного значения от фактического по всем предсказателям (столбцы) и первым 10 группам объектов (строки). В последней строке таблицы приведены результаты, усредненные по всем группам объектов, использованным в эксперименте.

Таблица 1: Среднеквадратичные уклонения предсказанных значений от фактического.

Номер выборки	$u_{\text{ср}}(x)$	$u_{\text{лин}}(x)$	$u_{\text{ягр}}(x)$	$u_{\text{инс}}(x)$
1	0.0379665690687	0.0108905744585	0.00310092999078	0.00482075104216
2	0.0351949752019	0.0103035552388	0.00300223050086	0.0527766368141
3	0.0431088036407	0.0105080793242	0.00230109234928	0.0622057557917
4	0.0403138553725	0.0102624036241	0.00265150316641	0.0570152403284
5	0.0404425439911	0.011505840531	0.00302884894609	0.0578286831701
6	0.0419549477253	0.0118239416413	0.00398996650142	0.0579758601408
7	0.0392895945949	0.0102540658186	0.00149332247267	0.0634876661128
8	0.0417724033723	0.011432503227	0.00202150624428	0.0621627641124
9	0.0362941536185	0.010417038905	0.00260492296544	0.0522648938623
10	0.0340004988966	0.0101537182782	0.00189431655157	0.0582985038951
среднее по 30	0.040435724	0.011511524	0.003425719	0.05871626

Из приведенных результатов видно, что метод ЯГР практически по всем группам объектов дал значительно более точный результат, чем другие методы, включая Искусственные нейронные сети.

6 Заключение

Несмотря на значительное увеличение точности благодаря применению метода ЯГР в данной работе не решался важнейший вопрос скорости вычислений. Построение окрестности текущей точки, обращение матриц и др. – все это требует значительных временных затрат, хоть и существенно меньших, чем при использовании CFD-кодов. Однако в дальнейшей работе предполагается решить эту проблему путем надлежащего структурирования данных обучения и вынесения трудоемких операций в область предварительной обработки данных.

Список литературы

[1] Кулешов А.П. Когнитивные технологии в адаптивных моделях сложных объектов. *Информационные технологии и вычислительные системы*, в. 1, 2008, с. 18-29.

[2] Бернштейн А.В., Кулешов А.П. Математические методы в когнитивных инженерных технологиях. Обозрение прикладной и промышленной математики, сер. *Вероятность и статистика.* 2008, т. 15, №3, с. 451-452.

[3] Вышинский В.В., Свириденко Ю.В. (2006). Применение технологии быстрого вычисления характеристик сложных объектов для расчета аэродинамических характеристик самолета. В приложении к журналу *Информационные технологии* № 3/2006, *Проекты в области информационных технологий Международного научно-исследовательского института проблем управления.*

[4] Gammerman A., Kalnashkan Y. and Vovk V. On-line prediction with kernels and complexity approximation principle. In Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence, pp. 170-176. AUAI Press, 2004.

[5] Bernstein A.V., Kuleshov A.P., Sviridenko Yu.N., Vyshinskiy V.V. Fast Aerodynamic Model for Design Technology. Proceedings of West-East High Speed Flow Field Conference. Nov. 19-22, 2007, Moscow, Russia, http://wehsff.imamod.ru/pages/s7.htm.

[6] Бернштейн А.В., Вышинский В.В., Кулешов А.П., Свириденко Ю.Н. Быстрый метод аэродинамического расчета для задач проектирования. *Труды Центрального аэрогидродинамического института им. проф. Жуковского.* Выпуск №2678 *Применение искусственных нейронных сетей в задачах прикладной аэродинамики.* М.: ЦАГИ, 2008, с. 35-45.