

План:

1. Технология предсказательного моделирования.
2. Интеллектуальный анализ данных в предсказательном метамоделировании
3. Снижение размерности данных в предсказательном моделировании
4. Нетрадиционные постановки задач снижения размерности данных

Тезисы:

В последние годы компьютерное проектирование стало необходимым (а зачастую и важнейшим) элементом в различных областях, таких как авиакосмическая промышленность, информационные управляющие системы, телекоммуникационные системы, биотехнология, нефтегазовый сектор и многих других. Компьютерные системы предсказательного моделирования (называемые также системами поддержки принятия инженерных решений) вместе с компьютерными системами проектирования давно используются для автоматизации труда инженера-проектировщика и повышения качества принимаемых решений. Важной частью предсказательного моделирования является имитационное моделирование, используемое для исследования сложных информационно-телекоммуникационных систем (ИТКС). Для проектирования и исследования информационно-вычислительных сетей (ИВС), являющихся телекоммуникационными компонентами ИТКС, имитационное моделирование является мощным инструментом, позволяющим предсказывать вероятностно-временные и надежность характеристики сетей, проектировать и исследовать процедуры управления сетью. Как ответ на возникшие сложности, в последние годы стало активно развиваться новое, альтернативное направление математического моделирования - метамоделирование, в рамках которого математические предсказательные модели для сложных систем строятся по результатам вычислительных экспериментов с моделями, основанными на первых принципах физики, или с имитационными моделями.

В самом общем виде базовая задача построения метамоделли формулируется следующим образом. Пусть $\mathbf{O} = \{O\}$ есть множество объектов рассматриваемого класса. Для каждого объекта $O \in \mathbf{O}$ имеется его цифровое описание $X = X(O)$ размерности p , полностью определяющее этот объект. Ключевой задачей интеллектуального анализа данных в метамоделировании

является задача построения по данным (5) аппроксимационной зависимости $FSM(X \square D_n)$ (7). Величину погрешности в (8) для точек множества X_n (6) можно непосредственно вычислить и проконтролировать, а погрешность для новых (out-of-sample) точек $X_{new} \square X / X_n$ характеризует обобщающую способность построенной метамоделю $FSM(X)$. Обзоры методов параметрической и непараметрической регрессии для построения аппроксимационных зависимостей.

В приложениях размерность p цифрового описания объекта может быть очень высока. Другим примером высокой размерности входных данных являются модели, описывающие процессы передачи и обработки информации в ИВС. Пусть в ИВС циркулирует информация разных приоритетов, и информация более высокого приоритета «обгоняет» информацию более низкого приоритета при передаче по трактам передачи данных. Поэтому процесс передачи по тракту можно достаточно адекватно описать моделями систем массового обслуживания с приоритетами. Задача построения аппроксимирующего многообразия M может быть сформулирована как Задача снижения размерности, или, в современной терминологии, как Задача моделирования многообразий (Manifold Learning Problem). Мы сформулируем постановки функциональной задачи снижения размерности для двух различных случаев – при заданных целевых функционалах, близость которых должна обеспечиваться при переходе от исходного к восстановленному цифровому описанию объекта, и для неизвестных функционалов, заданных только выборками их значений.

Задача функционального снижения размерности при известных целевых функционалах. По выборке X_n из Множества данных X , лежащего в пространстве меньшей размерности, построить решение $\square = (h, g)$ (13), (15), удовлетворяющее условию и дополнительному условию $f(g(h(X))) \square f(X)$ для всех $X \square X$, $f \square \Phi$ для заданного семейства целевых функционалов Φ , определенных на Множестве данных X .