

ОПЫТ ВЫСОКОТОЧНОГО МОДЕЛИРОВАНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ НА МАССОВИДНЫХ СТАТИСТИЧЕСКИХ МАТЕРИАЛАХ

© 2008 г. Г. А. Гольц
(Москва)

Во многих технико-экономических, прогнозных и управленческих задачах требуется точность расчета в пределах среднего отклонения 5–7%. Главная цель подобных расчетов — определять, что произойдет с многокомпонентной системой при существенном изменении величины формирующих ее факторов. Отличительной особенностью подобных систем является смешение детерминированных и вероятностных свойств. Часто существенным ограничением применению стандартного вероятностного подхода является отсутствие либо слабое соответствие требованию аксиом теории вероятностей: случайности, независимости и нормальности.

В процессе многолетней научной, производственной и преподавательской деятельности автору пришлось решать многие крупные задачи народнохозяйственного уровня на количественной основе. За это время были выработаны определенные правила работы с исходной информацией и ее обработкой: вне зависимости от территориального уровня систем и исторической глубины процессов везде соблюдается принципиальное единство *к подходу и технике расчетов*, а также массовидность исходной информации.

Феномен случайного события. Когда рассматривается ряд показателей по историческому времени, то можно вполне считать, что их значения по каждому году, стране, региону случайны. Но как только делается переход к содержательным переменным, упорядоченным во времени, и между ними ищется связь, феномен случайности пропадает. Он остается только в пределах случайной ошибки в исходных данных. Тогда можно считать анализируемые ряды исходных содержательных переменных почти детерминированными. И в этом случае имеются ограничения в критериях стандартной математической статистики и особенно ее разнообразных оценок соответствия теории фактам. Опыт моделирования показал, что в массе регионов страны поиски методов сжатия исходной многомерной информации не дают желаемого приближения теории к факту для получения универсальной зависимости от содержательных переменных при большом и даже исключительно большом разнообразии условий применения стандартных методов математической статистики.

В физике для обработки детерминированных многомерных данных применяют разнообразные методы комбинирования и подгонки в целях получения универсальных формул. Однако общего метода сжатия многомерной информации в таких сложных условиях до сих пор не выработано.

Мы предлагаем в таких случаях использовать модельный фильтр. С его помощью в порядке экспериментальных испытаний можно отбрасывать результаты, которые по модулю выходят за границы относительного максимального отклонения теории от факта по каждой точке-наблюдению более чем на 15%. Тогда при симметричном распределении отклонений среднее отклонение составит соответственно 5%. Для моделирования социально-экономических процессов такой точности вполне достаточно.

Почему по каждой точке? Дело в том, что используемые в математической статистике методы выявления трендов для реального прогнозирования и управления мало пригодны. Нужны теоретические модели связи содержательных переменных, проходящие в окрестностях каждой точки. Только в этом случае можно увидеть реальный ход социально-экономических процессов, а тренды можно оставить для рассуждений философам, политэкономистам и политикам. Иными словами, ретроспективная информация должна быть использована для выработки достаточно точных рекомендаций. А главное — адекватно предвидеть, к каким конкретным последствиям приведет их реализация. Этого можно достичь лишь с помощью метода высокоточного моделирования.

Выявление реального диапазона существования многомерной социально-экономической системы позволяет в пределах него строить оптимальные и всякие другие траектории развития и

применять оптимальные модели и рациональные подходы. Главное — не выходить за пределы этого диапазона, чтобы не разрушать социально-экономическую систему как таковую.

Технологические этапы обработки массива исходных данных:

- создается электронный банк статистических показателей (не менее 70);
- предлагается содержательная гипотеза связи показателей;
- апробация версии производится в координатах XYZ , обработка значений в статистическом пространстве точек-наблюдений ведется с помощью программы Statistica (по нашему опыту наилучшей аппроксимацией является квадратичная поверхность, которая обладает свойством гладкости);
- для подбора переменных помимо гипотетических связей строится корреляционная матрица, по ней отбираются связи, которые по величине $r \geq 0.9$; не исключается нагрузка каждой оси дополнительными переменными (как правило, показателями одной размерности внутри оси);
- по полученной аппроксимирующей поверхности вычисляются расчетные (условно теоретические) связи трех переменных (в простейшем случае — плоскость с двумя переменными);
- затем определяются отклонения расчетных значений от фактических, строится гистограмма отклонений. Если по первоначальной версии отклонения на всем массиве точек выходят за рамки максимального отклонения 15–20% и среднего 5–7%, она отбрасывается, а расчетный процесс продолжается с другими содержательными переменными;
- достоинством такого последовательно направленного по фильтру отбора решений является то, что при таком переборе могут возникнуть неожиданные комбинации и новое содержательное знание получится как бы автоматически.

Идея предлагаемого принципа состоит в том, чтобы описывать значение функций нескольких переменных с необходимой точностью в пределах отклонения факта от теории *по каждой точке* (году, региону) максимально в 15–20 и в среднем 5–7%. Эта точность достигается благодаря использованию соответствующего критерия отклонения и перебору содержательных переменных. Метод применим для восстановления динамических рядов и регионального анализа. В настоящем докладе основное внимание уделяется практическим результатам его применения, полученным за последние 5 лет.

Перечень накопленного и используемого для обработки массового статистического материала:

- динамические ряды по России (в ее границах на соответствующий год) — 150 показателей за каждый год на XVIII–XX вв.;
- динамические ряды по 79 регионам России за каждый год (по 50 показателям за период 1994–2005 гг.);
- динамические ряды по 10 странам с шагом в 5 лет (по 26 показателей за период 1900–2000 гг.);
- выборочные ряды показателей по внутригородским, пригородным, городским и межрегиональным транспортным связям и потокам.

Основные методические результаты многомерных динамических исследований. Пожалуй, самым главным явилось то, что с помощью чисто дедуктивных построений исходных гипотез для динамических многомерных пространств в исключительных случаях можно построить адекватную и хорошо работающую модель. На первый план здесь выходит последовательный анализ эмпирических данных, что, по нашему мнению, является следствием по большей части иррационального поведения людей и больших коллективов, которое обычно в принципе не совпадает с рациональными и формально описываемыми конструкциями. Обобщение опыта состоит в возможности находить границы многомерного диапазона существования социально-экономических, социально-технических, социально-природных, социально-политических систем. Выбор конкретной траектории в найденном диапазоне является областью хорошо разработанных оптимальных, рациональных, логистических и т.п. конструкций. А вот определение нижних и верхних границ возможного существования сложных динамических систем определяется с помощью предлагаемого метода высокоточного моделирования социально-экономических процессов.

Перечислим важные нетривиальные результаты, полученные независимыми исполнителями описанным методом в самое последнее время (апрель 2007 г.): обратная зависимость между рождаемостью и обеспеченностью жилым фондом; зависимость ВРП от строительства и числа организаций; зависимость ВВП от ВРП и продукции отраслей природно-ресурсного комплекса. Число новых зависимостей нарастает, как снежный ком, по мере освоения накопленного массива информации.