ТЕОРЕТИЧЕСКАЯ И МАТЕМАТИЧЕСКАЯ ФИЗИКА

УДК 519:25

ВОЗМОЖНОСТИ ДИАЛОГА ПРИ РЕДУКЦИИ ИЗМЕРЕНИЙ

Б. И. Волков, Ю. П. Пытьев, Д. В. Юдин

(кафедра математики)

Принято считать, что влияние субъективного мнения исследователя на решение задачи интерпретации экспериментальных данных нежелательно, и его по мере возможности следует исключать из системы обработки измерений. При этом, к сожалению, полностью игнорируются богатый научный опыт и конкретная информация об объекте исследования, которой обладает ученый. Но для того чтобы включить его в систему обработки без риска внести в результат субъективные ошибки, необходим критерий, позволяющий контролировать непротиворечивость представлений исследователя и результатов данного эксперимента. В настоящей работе рассматриваются основные вопросы, возникающие при решении задач обработки и интерпретации в диалоге с исследователем.

В первой части в рамках различных моделей процесса измерения формулируется задача интерпретации. Приводится метод сравнения моделей, определяющий качество интерпретации. Предлагается способ формализации априорных знаний исследователя.

Во второй части более детально обсуждаются вопросы, связанные с прогнозом влияния дополнительной информации на качество результатов обработки. Например, при графическом представлении искомой зависимости исследователь часто может уточнить поведение некоторой части экспериментальной кривой, и в этом случае полезно знать, как будет сказываться такое локальное вмешательство на других участках.

В третьей части обсуждается понятие надежности модели эксперимента, в которой, в частности, учтены априорные представления исследователя об изучаемом объекте.

Приводятся результаты тестовых расчетов и обсуждаются возможности диалога в задачах интерпретации.

1. **Модели с дополнительной информацией. Формализация априорных знаний.** Рассмотрим линейную схему измерений $\xi=Af+\nu$, где ξ — результат измерения сигнала f, проведенного на приборе A с шумом ν. Ограничимся конечномерным случаем, причем ξ, $\nu\in\overline{\mathcal{R}}$, $f\in\mathcal{R}$, A — линейный оператор. В зависимости от имеющейся информации о шуме, приборе и сигнале выделим следующие модели схемы измерений.

а) *Модель* $[A, \Sigma]$. Здесь считается, что задан оператор A и корреляционный оператор шума $\Sigma:\Sigma x=E\nu(x, \nu)$, $x\in\overline{\mathcal{R}}$, причем $E\nu=0$. Простейшая задача редукции для модели $[A, \Sigma]$ ставится так: найти линейный оператор R, для которого при заданном линейном операторе U (приборе) и любом f

$$E\,\|R\xi-Uf\|^2=\inf\{E\,\|R'\xi-Uf\|^2\,|\,ER'\xi=Uf\}. \qquad (1)$$

Если задача (1) разрешима, редукцию $R\xi=Uf+R\nu$ можно интерпретировать как «выход» заданного прибора U, искаженный шумом $R\nu$.

Теорема 1 [1]. Условие $U \in D_A = \{U', U'(I-A^-A)=0\}$* необходимо и достаточно для разрешимости задачи (1). При этом решение имеет вид $R=U(A^*\Sigma^{-1}A)^-A^*\Sigma^{-1}$, корреляционный оператор шума $R\nu$ равен

$$\widehat{\Sigma}=(A^*\Sigma^{-1}A)^-, \quad \text{а} \quad h(U)=E\|R\xi-Uf\|^2=\|U\widehat{\Sigma}^{-1/2}\|_2^{2**}$$

определяет полную ошибку редукции.

Определение. Будем говорить, что качество модели $[A, \Sigma]$ не ниже, чем модели $[A', \Sigma']$, и писать $[A, \Sigma] \prec [A', \Sigma']$, если $D_A \supset D_{A'}$ и $h(U) \leqslant h'(U)$ для любых приборов $U \in D_A$ [2]. Согласно определению лучшей модели соответствуют более широкие возможности редукции и менее интенсивный шум.

Далее на основе результатов, полученных для модели $[A, \Sigma]$, рассмотрим модели, включающие дополнительные измерения и априорную информацию.

б) *Модель* $[\widetilde{A}, \widetilde{\Sigma}]$ объединяет N независимых измерений $\xi_i = A_i f + \nu_i$, $i=1, 2, \ldots, N$. При этом

$$\widetilde{A}=\begin{pmatrix} A_1 \\ \vdots \\ A_N \end{pmatrix}, \quad \widetilde{\Sigma}=\begin{pmatrix} \Sigma_1 & & 0 \\ & \ddots & \\ 0 & & \Sigma_N \end{pmatrix}.$$

В [2] показано, что $[\widetilde{A}, \widetilde{\Sigma}] \prec [A_i, \Sigma_i]$, $i=1, 2, \ldots, N$, т. е. редукция совокупности измерений позволяет получать результаты лучшего качества.

в) *Модели с априорной информацией* $[A, F, \Sigma]$, $[A, f_0, F_0, \Sigma]$. Здесь в схеме измерений f считается случайным вектором. Для модели $[A, F, \Sigma]$ помимо A и Σ известен ковариационный оператор $F: Fx=Ef(x, f)$, $x \in \mathscr{R}$. Задача редукции формулируется следующим образом: для заданного линейного оператора U определить линейный оператор R такой, что

$$E\|R\xi-Uf\|^2=\inf\{E\|R'\xi-Uf\|^2 \mid R'\}. \tag{2}$$

Теорема 2. Пусть $AFA^*+\Sigma>0$, тогда задача (2) имеет единственное решение $R=UFA^*(AFA^*+\Sigma)^{-1}$ для любого U. Ошибка редукции $h(U)$ для модели $[A, F, \Sigma]$ не больше, чем для модели $[A, \Sigma]$ [3]. В модели $[A, f_0, F_0, \Sigma]$ кроме A и Σ известны среднее значение $f_0=Ef$ и корреляционный оператор $F_0: F_0x=E(f-f_0)(x, f-f_0)$, $x \in \mathscr{R}$. Редукция в этом случае имеет вид $R(\xi-Af_0)+Uf_0$.

г) *Связь моделей.* Легко показать, что все результаты, полученные для моделей $[A, f_0, F_0, \Sigma]$ и $\left[\begin{pmatrix} A \\ A' \end{pmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma' \end{pmatrix}\right]$, совпадают, если положить $A'=I$, $\Sigma=F_0$, а схему измерений переписать в виде

$$\begin{pmatrix} \xi \\ f_0 \end{pmatrix}=\begin{pmatrix} A \\ I \end{pmatrix} f+\begin{pmatrix} \nu \\ \nu' \end{pmatrix}.$$

Это формально гарантирует эквивалентность дополнительных измерений и априорной информации об f [3].

В связи с этим можно предложить следующий подход к учету априорной информации. Исследователь с определенной уверенностью предлагает свой вариант решения, который формализуется как допол-

---

* Для любого оператора A псевдообратный A^- равен $\lim\limits_{\omega \to 0} A^*(AA^*+\omega I)^{-1}$.

** Имеется в виду операторная норма Гильберта—Шмидта, $\|B\|_2^2=\mathrm{Tr}B^*B$.

нительное измерение $f_0=A'f+\nu'$. Если $A'=I$, то f_0 — информация обо всем f; если $A'=P$, где P — ортогональный проектор на $\mathscr{R}(P)$, то f_0 — локальная информация. Будем характеризовать ошибку ν' корреляционным оператором αF_0. Если $F_0=I$, то возможны любые отклонения f от f_0, определяемые параметром $\alpha=\sigma^2$, равным дисперсии ν'. Однако исследователь нередко располагает более определенной информацией, в частности о гладкости решения на некотором интервале $[x_1, x_2]$. В этом случае, например, в качестве ν' можно взять решение краевой задачи

$$\eta^{(n)}(x)=\zeta(x),$$

$$\eta^{(i)}(x_1)=\eta_1^i, \quad i=0, 1, \ldots, k-1, \quad \eta^{(j)}(x_2)=\eta_2^{(j)}, \quad j=0, 1, \ldots, n-k-1,$$

где $\zeta(x)$ — случайный вектор, $E\zeta=0$, корреляционный оператор которого равен αI, $\alpha>0$. Тогда $\eta(x)$ будет иметь определенную гладкость и характеризоваться корреляционным оператором $\alpha F_0=\alpha GG^*$, где G — интегральный оператор Грина соответствующей краевой задачи.

2. **Влияние дополнительной информации на качество результатов редукции.** Приведем результаты, позволяющие оценивать возможное влияние дополнительных данных на результат редукции. При графическом представлении решения исследователь может заранее знать поведение искомой кривой на отдельных участках. Желательно выяснить, как повлияет учет такой информации на результат редукции для всей кривой и от чего зависит такое влияние.

Пусть $\xi=Af+\nu$ — исходное измерение, а локальная информация представлена в виде $f_0=Pf+\nu'$. Предположим, что модель $[A, \Sigma]$ такова, что возможна редукция к идеальному прибору $U=I$. Обозначим $\hat{f}=R\xi$ решение задачи (1) при $U=I$ для модели $[A, \Sigma]$, а $\hat{\hat{f}}=\tilde{R}\begin{pmatrix}\xi\\ f_0\end{pmatrix}$ — решение той же задачи для модели $\left[\begin{pmatrix}A\\ P\end{pmatrix}, \begin{pmatrix}\Sigma & 0\\ 0 & \alpha F_0\end{pmatrix}\right]$, учитывающей информацию, предложенную исследователем. В данном случае оператор P выделяет область на кривой, относительно которой сформулировал свои предположения (f_0) исследователь; параметр $\alpha>0$ определяет, насколько точны эти предположения. Если $F_0>0$, то при $\alpha\to 0$ $P\hat{\hat{f}}\to Pf_0$, а при $\alpha\to\infty$ $\hat{\hat{f}}\to\hat{f}$ для любого f_0. Понятно, что влияние информации f_0 на кривую f тем больше, чем меньше α. Максимальный эффект можно получить предельным переходом при $\alpha\to 0$. Подчеркнем, что при любом $\alpha>0$ влияние, очевидно, не зависит от (неизвестного) f и может быть вычислено априори.

Теорема 3. Пусть невырожден оператор $S=A^*\Sigma^{-1}A+PF_0^{-1}P$, тогда максимальное влияние информации f_0, αF_0 на оценку $\hat{\hat{f}}$ для модели $[\tilde{A}, \tilde{\Sigma}]$ определяется оператором $V=S^{-1/2}Q-QS^{-1/2}$, где $Q=S^{-1/2}PF_0^{-1}PS^{-1/2}$. При этом редукция имеет вид $\hat{\hat{f}}=\hat{f}-VA^*\Sigma^{-1}\nu$, а корреляционный оператор шума для $\hat{\hat{f}}$ равен $\hat{\Sigma}-V$, где $\hat{\Sigma}=(A^*\Sigma^{-1}A)^-$.

Таким образом, область, на которую распространяется действие локальной информации о сигнале f, определяется оператором «влияния» V и зависит только от модели.

В предположении, что задача редукции разрешима и $A^*\Sigma^{-1}A+PF_0^{-1}P>0$, полная ошибка для объединенной модели $\left[\begin{pmatrix}A\\ P\end{pmatrix}, \begin{pmatrix}\Sigma & 0\\ 0 & \alpha F_0\end{pmatrix}\right]$ равна

$$h_\omega(U) = E\,\|R\xi - Uf\|^2 = \operatorname{Tr} U\,(A^*\Sigma^{-1}A + \omega P F_0^{-1} P)^{-1}\,U^*,$$

где $\omega = 1/\alpha$. При $\omega \to 0$ получаем

$$h_0(U) = \lim_{\omega\to 0} h_\omega(U) = \operatorname{Tr} US^{-1/2}\,(S^{-1/2}\,A^*\Sigma^{-1}AS^{-1/2})^{-}\,S^{-1/2}U^* -$$

это случай отсутствия дополнительной информации. Значение $h_\infty(U) = \lim\limits_{\omega\to\infty} h_\omega(U) = \operatorname{Tr} US^{-1/2}\,(I - Q^{-}Q)\,S^{-1/2}U^*$ определяет энергию шума при достоверном знании части сигнала (предельные переходы следуют из леммы 7 работы [1]). Для исследования зависимости $h_\omega(U)$ от ω полезна формула $\dfrac{dh_\omega(U)}{d\omega} = -\operatorname{Tr} US_\omega^{-1}PF_0^{-1}PS_\omega^{-1}U^*$, где $S_\omega = A^*\Sigma^{-1}A + \omega PF_0^{-1}P$.

3. **Надежность модели.** Из теорем 1 и 2 следует, что точность решения задачи редукции полностью определяется принятой моделью схемы измерений, в которую может входить и дополнительная информация, и не зависит от конкретного измерения. Однако объективность принятой модели до сих пор никак не контролировалась. В [4] построен критерий, позволяющий судить, насколько принятая модель отвечает реальному положению вещей.

Рассмотрим модель $[A, \Sigma]$. Пусть вектор шума ν имеет нормальное распределение с нулевым средним и корреляционным оператором $\Sigma > 0$; последний будем считать точно известным (в ряде экспериментов есть возможность непосредственно измерять Σ). Таким образом, все сомнения относительно модели будут касаться лишь оператора A. Задача проверки гипотезы о верности модели $[A, \Sigma]$ сводится к аналогичной задаче о представимости ξ в виде $\xi = a + \nu$, где $a \in \mathcal{R}(A)$ либо $a \overline{\in} \mathcal{R}(A)$. Разумность такой замены исходной задачи очевидна, так как при условии $a \overline{\in} \mathcal{R}(A)$ модель должна быть отвергнута. Критерий отношения правдоподобия в данном случае эквивалентен условию

$$t(\xi) = \|\Sigma^{-1/2}(\xi - A\hat{f})\|^2 = \|Q_m\Sigma^{-1/2}\xi\|^2 \geqslant t_\alpha, \tag{3}$$

где $\hat{f} = (\Sigma^{-1/2}A)^{-}\,\Sigma^{-1/2}\xi$, $Q_m = I - \Sigma^{-1/2}A\,(\Sigma^{-1/2}A)^{-}$,

t_α определяется из условия $\int\limits_{t_\alpha}^{\infty} p_{m,0}(z)\,dz = \alpha$, где $p_{m,0}$ — плотность χ^2-распределения с m степенями свободы. Здесь используются следующие свойства статистики $t(\xi)$: при $a \in \mathcal{R}(A)$ $t(\xi)$ имеет χ^2-распределение с m степенями свободы, при $a \overline{\in} \mathcal{R}(A)$ $t(\xi)$ имеет нецентральное $\chi^2_{m,\theta}$-распределение с m степенями свободы и параметром нецентральности $\theta^2 = \|Q_m\Sigma^{-1/2}a\|^2$. Критерий (3) будет отвергать модель $[A, \Sigma]$, когда ξ удовлетворяет неравенству $t(\xi) \geqslant t_\alpha$, и в случае верной модели в среднем в 100 α% случаев будет приводить к ошибочным решениям.

В работе [4] надежность модели определена как статистика $\alpha(\xi) = \inf\{\alpha,\ t(\xi) \geqslant t_\alpha\}$. Поскольку, как показано в [4], критерий (3) является равномерно наиболее мощным в некотором естественном классе критериев, согласно такому определению, надежность — минимальная вероятность ошибочно отвергнуть модель, основываясь на измерении ξ. В случае верности модели надежность $\alpha(\xi)$ имеет равномерное распределение на [0, 1]. Если же $a \overline{\in} \mathcal{R}(A)$, то плотность вероятности $\alpha(\xi)$ неограничена в нуле (и только в нуле). Поэтому если данному измерению ξ соответствует малое значение $\alpha(\xi)$, то скорее всего модель неверна, и ее необходимо отвергнуть. Если же $\alpha(\xi)$ велико, то основа-

ний отвергнуть модель нет, и ее следует принять. Разумеется, надежность можно построить и для объединенной модели эксперимента, включающей дополнительную информацию. В таком случае $\alpha\begin{pmatrix}\xi\\ f_0\end{pmatrix}$ позволит оценить непротиворечивость этой объединенной модели и результатов эксперимента.

Для иллюстрации возможности диалога при редукции измерений на ЭВМ были проведены следующие модельные эксперименты. Тестовый П-образный сигнал f искажался оператором A, имитирующим реальный прибор, а к результату Af прибавлялся случайный шум, разыгранный на ЭВМ в виде реализации нормально распределенного $N(0, \Sigma)$ случайного вектора (рис. 1, *а*). В рамках модели $[A, \Sigma]$ решалась за-

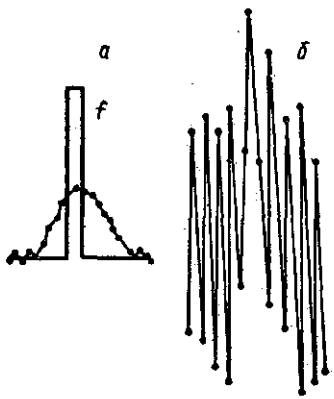

Рис. 1

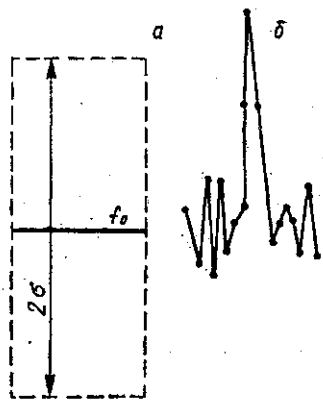

Рис. 2

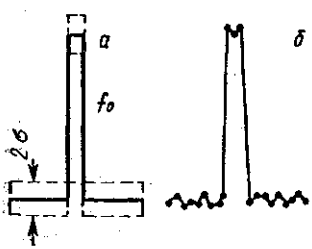

Рис. 3

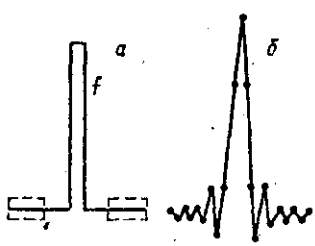

Рис. 4

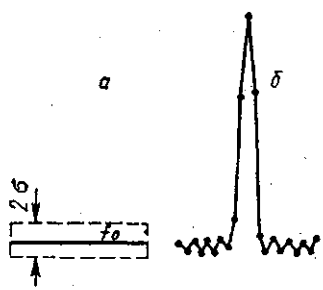

Рис. 5

дача редукции (1) к $U=I$, результат которой $f+Rv=\hat{f}$ приведен на рис. 1, *б*. Такой результат неудовлетворителен из-за слишком большого шума Rv.

Для повышения качества результатов использовалась следующая дополнительная информация: $f_0=0$, $F_0=I$, $\alpha=\sigma^2=1$ в схеме $f_0=f+v$ (рис. 2, *а*). Надежность объединенной модели оказалась достаточно большой ($\sim 0{,}87$), а качество полученной оценки существенно выше, чем в предыдущем случае (рис. 2, *б*). Если использовать более полную информацию (рис. 3, *а*), то качество результата значительно повышается (рис. 3, *б*). Пусть теперь дополнительная информация имеет «локальный» характер и соответствует схеме $f_0=Pf+v'$, где P — ортогональный проектор на первые и последние координаты вектора

(рис. 4, *а*). Надежность соответствующей объединенной модели еще достаточно велика — примерно 0,6, а качество редукции удовлетворительное (рис. 4, *б*). Отметим, что область влияния охватывает и центр графика, где удается снизить шум редукции.

Наконец, предположим, что П-образного импульса не существует, что формализуем в виде $f_0=0$, $F_0=I$, $\alpha=0,01$ для схемы $f_0=f+\nu'$ (рис. 5, *а*). Оказывается, что в этом случае надежность модели, объединяющей исходную схему измерений и такие дополнительные данные, крайне мала ($\sim 0{,}001$), и принятое предположение об f необходимо отвергнуть как противоречащее измерению ξ. Такая ситуация может возникнуть, например, когда исследователь ошибается в определении местоположения или формы импульса. Заметим, что результат, полученный с использованием ошибочной дополнительной информации (рис. 5, *б*), при достаточно высокой точности сильно отличается от тестового сигнала в центральной области.

СПИСОК ЛИТЕРАТУРЫ

[1] Пытьев Ю. П. // Матем. сб. 1982. **118(160)**, № 1(5). С. 19—49. [2] Пытьев Ю. П. // Матем. сб. 1983. **120(162)**, № 2. С. 240—272. [3] Пытьев Ю. П. Возможности измерительно-вычислительного комплекса при планировании, анализе и интерпретации измерений: лекции для молодых ученых ОИЯИ. № Р 10-84-525. Дубна, 1984. [4] Пытьев Ю. П. // Вестн. Моск. ун-та. Сер. 3, Физ. Астрон. 1986. **27**, № 3. С. 14—19.

Поступила в редакцию
18.09.85

ВЕСТН. МОСК. УН-ТА. СЕР. 3, ФИЗИКА. АСТРОНОМИЯ. 1987. Т. 28, № 1

УДК 530.145:530.12

ИНТЕГРИРОВАНИЕ НЕГАМИЛЬТОНОВЫХ СИСТЕМ МЕТОДОМ УДВОЕНИЯ ПЕРЕМЕННЫХ

Ю. Г. Павленко, С. И. Зеленский

(кафедра теоретической физики)

Гамильтоновы системы занимают весьма важное место в теоретической и математической физике. С одной стороны, они описывают множество явлений в классических и квантовых теориях. С другой стороны, канонический формализм позволяет развить универсальные методы интегрирования систем нелинейных уравнений [1—8].

В общепринятом подходе уравнения Лагранжа и Гамильтона эквивалентны, поскольку преобразование Лежандра реализует переход от обобщенных координат и скоростей к фазовому пространству координат и импульсов [2, 8—10]. Лагранжевы уравнения представляют экстремали функционала действия. Однако произвольные системы не являются уравнениями Лагранжа—Эйлера вариационной задачи на безусловный экстремум для некоторого функционала [11, 12]. Поэтому сложилось убеждение, что если система не имеет лагранжевой формы, то переход к гамильтоновым уравнениям невозможен. Например, в механике системы, учитывающие силы трения, считаются негамильтоновыми. То же относится к уравнениям, описывающим химические реакции или различные экологические модели, сингулярно-возмущенным уравнениям в теориях движения плазмы, жидкости и газа. Негамильтонов характер этих систем заставляет в каждом конкретном случае искать специальные методы решения.