



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 34 за 2016 г.



ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

Алексеев А.К., Бондарев А.Е.

О ценной информации в
обратных задачах и ее
визуализации

Рекомендуемая форма библиографической ссылки: Алексеев А.К., Бондарев А.Е. О ценной информации в обратных задачах и ее визуализации // Препринты ИПМ им. М.В.Келдыша. 2016. № 34. 20 с. doi:[10.20948/prepr-2016-34](https://doi.org/10.20948/prepr-2016-34)
URL: <http://library.keldysh.ru/preprint.asp?id=2016-34>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Российской академии наук**

А.К.Алексеев, А.Е.Бондарев

**О ценной информации
в обратных задачах
и ее визуализации**

Москва — 2016

Алексеев А.К., Бондарев А.Е.

О ценной информации в обратных задачах и ее визуализации

Рассмотрена возможность формализации некоторых интуитивных понятий об информации, связанных с понятиями «цели» и «инструкции». Соответствующая «ценная информация» (ЦИ) трактуется как знание того, какое воздействие на управляющие параметры выбрать, чтобы с известной точностью достигнуть некоторой цели при данном состоянии системы. Она имеет векторную (градиент целевого функционала, информант) и тензорную (информационная матрица Фишера) составляющие. При решении обратных задач перенос векторной компоненты ценной информации описывается сопряженными уравнениями. Представлены иллюстрации, демонстрирующие возникновение и перенос векторной части ЦИ в задачах вычислительной газодинамики. Рассмотрена также связь ценной информации с другими существующими теориями информации.

Ключевые слова: градиент целевого функционала, информационная матрица Фишера, сопряженные уравнения, обратная задача

Alexey Kirillovich Alexeev, Alexander Evgenyevich Bondarev

On Important Information for Inverse Problems and its Visualization

The «important information» is considered as the formalization of the intuitive notions of the information related with «goal», «instruction», «control», «intention». The important information corresponds the knowledge of the movement in the control variables space that is necessary to achieve the goal with the known uncertainty under the present system state. The important information has the vector (goal functional gradient) and the tensor (Fisher's information matrix) components. For inverse problems vector component transfer for such information is described by adjoint equations. The illustrations are provided that present the generation and transfer of the vector component of important information in inverse computation fluid dynamics problems. We also consider the communication of important information to other existing information theories.

Key words: important information, goal functional gradient, Fisher's information matrix, adjoint equations, inverse problem

Работа выполнена при поддержке Российского фонда фундаментальных исследований, проекты 14-01-00769-а и 16-01-00553-а.

Введение

Значительное количество понятий об информации, используемых в современное время, носит полуформальный или полностью интуитивный характер. Рассмотрим одно из них с точки зрения возможной формализации. На бытовом уровне (повсеместно) и в научных кругах (реже) используется понятие информации, связанное с «рецептами», «инструкциями» и характеризующее такими свойствами, как «ценность информации», «важность», «смысл» [1,2]. Это понятие, вообще говоря, напрямую не связано с наиболее часто используемой теорией информации Шеннона [3] и используется только на интуитивном уровне. Однако в нескольких распространенных классах задач соответствующий формализм существует достаточно давно. Представляется, что трудности в основном связаны с его интерпретацией и «эффектом Вавилонской башни» – несовпадением терминологии и обозначений и непересечением областей научных интересов соответствующих специалистов. В частности, в задачах статистического оценивания [4,5], обратных статистических задачах [6], в задачах вариационного усвоения данных [7,8], обратных задачах в вариационной постановке [9] такое интуитивное понятие информации можно единообразно формализовать.

1. Интуитивное понятие информации, связанное с наличием цели

Наиболее часто используемая теория информации Шеннона [3] ориентирована на передачу информации (кодирование, сжатие, помехоустойчивость) без связи с содержанием. Количество информации Шеннона, содержащееся в некотором тексте длиной m знаков, имеющих вероятность появления P_i , имеет вид:

$$I_{Sh} = -\sum_i^m P_i \log_2 P_i.$$

Из этого определения видно, что при перестановке символов местами количество информации не изменится.

Трудно примириться с тем, что произвольная перестановка всех букв в романе «Война и мир» не повлияет на количество содержащейся там информации. Это связано с тем, что интуитивные понятия информации, неосознанно, но повсеместно используемые, далеко не тождественны «информации Шеннона».

Вариантов интуитивной информации достаточно много. Например, базы данных, телефонная книга, записи расходов воспринимаются как информация, как информация воспринимаются и рецепт пирога, и стрелка дорожного указателя. Рассматриваемая в настоящей работе информация относится к последнему типу. Интуитивная информация такого вида должна иметь «смысл», приносить пользу, она должна обладать «ценностью». Ясно, что в

некоторых случаях один бит информации может быть ценнее гигабайта (информация в стрелке, указывающей на выход из горящей библиотеки, может быть много ценнее, чем информация в целом книжном шкафу). Ясно также, что одна и та же формула может нести абсолютно разную информацию разным людям, т.е. информация также должна зависеть от состояния системы. Информация должна зависеть от цели, от имеющихся средств ее достижения. Интуитивная информация подразумевает также возможность существования дезинформации (имеет знак).

Существует несколько вариантов определения понятия такой информации: «*полезная информация*» [1], «*ценность информации*» [2]. Общим недостатком этих понятий является отсутствие строгих количественных определений. Однако надо отметить, что в ряде задач (например, в обратных задачах) подходящий формализм существует достаточно давно. Для его описания мы будем пользоваться термином «*ценная информация*». Под ценной информацией мы будем понимать знание того, какое воздействие на управляющие параметры выбрать, чтобы с известной точностью достигнуть некоторой цели при данном состоянии системы. Далее рассмотрим ряд известных задач с точки зрения подобного определения ценной информации.

2. Вариационная постановка обратных задач

Вариационная постановка обратных задач и задач оптимального управления является достаточно универсальным подходом [7,8,9], рассмотрим ее подробнее на примере задач усвоения данных [8,9]. Эти задачи посвящены определению вектора управляющих параметров u некоторой системы (например, атмосферы) по вектору наблюдения f^{obs} (давление, температура, компоненты скорости) при наличии ковариационной матрицы погрешности наблюдений W^{-1} . Они сводятся к решению вариационной обратной задачи с целевым функционалом вида

$$\varepsilon(u) = ((Au - f^{obs}), W(Au - f^{obs}))/2, \quad (1)$$

где A – оператор эволюции задачи (пропагатор)

$$f = Au, \quad (2)$$

а f – прогнозируемый вектор состояния системы.

Целевой функционал (1) обычно содержит скалярное произведение в L_2 , но в некоторых случаях возможно использование других пространств, например H_1 . В ряде задач [9] погрешность наблюдений не учитывается, и целевой функционал имеет более простую форму $\varepsilon(u) = ((Au - f^{obs}), (Au - f^{obs}))/2$. В случаях, когда оператор A представляется в виде матрицы $f_k = A_{kj}u_j$, функционал (1) имеет вид $\varepsilon(u) = (A_{ij}u_j - f_i^{obs})W_{ik}(A_{kj}u_j - f_k^{obs})/2$.

В том случае если динамика описывается системой частных дифференциальных уравнений (ЧДУ), чаще всего точный вид оператора A неизвестен и он используется неявно. Численное решение системы ЧДУ позволяет определить действие этого оператора $f = Au$. В качестве иллюстрации рассмотрим решение системы $du/dt = Bu$ с помощью конечно-разностной схемы $(u^{n+1} - u^n)/\tau = Bu$, которое дает следующую аппроксимацию действия оператора: $Au_0 = (E + \tau B) \cdot (E + \tau B) \dots (E + \tau B) \cdot u_0$, для момента времени $t = n\tau$: $Au = (E + \tau B)^n \cdot u_0$ (здесь E – единичная матрица, а B – матричный генератор полугруппы эволюции). Формально можно построить и матричную аппроксимацию оператора $A = (E + \tau B) \cdot (E + \tau B) \dots (E + \tau B)$, но в случае ЧДУ она имеет неприемлемые размеры и, как правило, не используется.

Если мы можем рассчитать действие оператора (и сопряженного оператора), то во многих случаях можно обойтись без использования оператора A в явном виде. С учетом этого мы будем далее пользоваться формальным видом оператора A , что позволяет существенно сократить выкладки.

Функционал (1) имеет смысл невязки расчета и измерения, нормированной на погрешность измерения. Его градиент имеет вид

$$\nabla \varepsilon = A^* W (Au - f^{obs}). \quad (3)$$

Шаг ньютоновской итерации, основанный на том, что в точке минимума $\nabla \varepsilon = 0$, дает решение вида

$$u = (A^* W A)^{-1} b, \quad (4)$$

где

$$F = A^* W A, \quad (5)$$

а

$$b = A^* W f^{obs}. \quad (6)$$

В большинстве задач усвоения данных оператор A задан неявно с помощью системы частных дифференциальных уравнений (ЧДУ) и $A^* W f^{obs}$ в явном виде не известен. К счастью, существует развитая техника определения $\nabla \varepsilon = A^* W (Au_0 - f^{obs})$ с помощью решения сопряженных уравнений [9,10]. Поэтому на практике иногда используются ньютоновские итерации вида

$$\Delta u = u - u_0 = -(A^* W A)^{-1} A^* W (Au_0 - f^{obs}), \quad (7)$$

или, много чаще, градиентные методы, неявно использующие некоторую примитивную аппроксимацию матрицы Фишера и решающие задачу итерационными методами типа наискорейшего спуска

$$u_i^{n+1} = u_i^n - \tau^n \nabla \varepsilon_i^n. \quad (8)$$

Итерации (7,8) в отличие от выражения (4) работоспособны и в нелинейном случае.

3. Статистическая постановка обратных задач

Выражения (4-6) используются в статистической теории оценивания [4,5]. Так, например, выражение $F = A^*WA$ является информационной матрицей Фишера. Величина $\nabla \varepsilon = A^*W(Au_0 - f^{obs})$ именуется вкладом, информантом (informant, score). Следует отметить, что английское слово score среди многочисленных значений («счет, задолженность, набранные очки в игре и др.») имеет и такое – «причина».

В том случае, когда мы имеем данные о погрешности измерений, возможно применение статистического метода решения обратных задач [6], являющегося вариантом статистической теории оценивания [4], а точнее метода максимального правдоподобия. Этот метод позволяет определить решение обратной задачи и его погрешность в статистическом виде. Ниже мы рассмотрим этот метод в соответствии с [11].

Пусть оператор задачи A известен в матричном виде. Считаем, что из анализа измерений случайного вектора $f_i' \in R^m$ у нас есть среднее $\langle f_i' \rangle = f_i^{obs}$ и ковариационная матрица погрешностей эксперимента W^{-1} . Представим результаты эксперимента в виде $f_i' = f_i^{obs} + \mathcal{J}_i'$. Предположим, что ошибка измерений \mathcal{J}_i' является статистически независимым ($\langle \mathcal{J}_i' \rangle = 0$) и нормально распределенным случайным вектором. Тогда она имеет следующую плотность распределения:

$$P(\mathcal{J}') = C_f \exp(-(\mathcal{J}', W\mathcal{J}') / 2), \quad (9)$$

где $C_f = (\det W)^{-1/2} (2\pi)^{-m/2}$, m – размерность пространства наблюдаемых величин.

Нас интересует случайный вектор u_i' , такой, что $Au_i' = f_i' = f_i^{obs} + \mathcal{J}_i'$ и, соответственно,

$$Au_i' - f_i^{obs} = \mathcal{J}_i'. \quad (10)$$

Обратная задача в статистическом виде может быть сформулирована как поиск вероятности $P(u' | f^{obs})$ того, что мы имеем u' при измеренном среднем f^{obs} .

Скалярное произведение в (9) имеет вид $(\mathcal{J}', W\mathcal{J}') = ((Au' - f^{obs}), W(Au' - f^{obs}))$. Это выражение совпадает с (1), однако в (9) подразумеваются средние по ансамблю измерений, в то время как в выражении (1) f^{obs} представляет один элемент ансамбля измерений.

Пользуясь симметричностью матрицы W , получим

$$(\mathcal{J}', W\mathcal{J}') = (u', Fu') - 2(b, u') + (f^{obs}, Wf^{obs}), \quad (11)$$

где $F = A^*WA$ и $b = A^*Wf^{obs}$ соответствуют выражениям (5) и (6).

Объединяя сомножители, не зависящие от u'_i , и подставляя в (9), получим стандартное выражение для нормально распределенного случайного вектора u'_i :

$$P(\mathcal{J}') = C_1 \exp(-(u', Fu')/2 + (b, u')), \quad (12)$$

где $C_1 = C_f \exp(-(f^{obs}, Wf^{obs})/2)$.

Исходя из $Au' - f^{obs} = \mathcal{J}'$ мы его интерпретируем как $P(u' | f^{obs})$. Это определит интересующее нас решение (максимум плотности вероятности), совпадающее с (4):

$$u = \langle u' \rangle = F^{-1}b = (A^*WA)^{-1}A^*Wf^{obs}. \quad (13)$$

При этом ковариационная матрица ошибки результата (обратная матрица Фишера) имеет вид

$$F^{-1} = (A^*WA)^{-1} = Cov_u. \quad (14)$$

Следует отметить, что в случае некорректно поставленной задачи матрицы Фишера $F = (A^*WA)$ недостаточно для однозначного определения решения и его погрешности. С алгебраической точки зрения дело заключается в том, что матрица Фишера может быть близка к вырожденной, а ее обращение может привести к сильной неустойчивости. Для регуляризации задачи используется матрица $D = Cov_b^{-1}$, содержащая априорную информацию об искомом решении (в виде обратной ковариационной матрицы априорного распределения): $P(u) = C \exp(-(u', Du')/2)$. Используя теорему Байеса, можно получить решение вида

$$u = (A^*WA + D)^{-1}A^*Wf^{obs} \quad (15)$$

с корреляционной матрицей погрешности

$$(F + D)^{-1}. \quad (16)$$

Таким образом, статистический метод позволяет решить обратную задачу в смысле получения наиболее вероятного решения и ковариационной матрицы ошибки результата.

Легко видеть, что статистическая постановка частично соответствует решению вариационной обратной задачи с целевым функционалом, который содержит регуляризирующий член, соответствующий априорной информации (матрица D , если $D_{ij} = \alpha E_{ij}$ (α – коэффициент регуляризации), описывает тихоновскую регуляризацию нулевого порядка [12]).

$$\varepsilon(u) = (A_{ij}u_j - f_i^{obs})W_{ik}(A_{km}u_m - f_k^{obs})/2 + u_i D_{ij}u_j / 2. \quad (17)$$

Минимум $\varepsilon(u)$ в вариационном подходе соответствует максимуму плотности вероятности u в статистическом методе (для нормального распределения). Так как $\Delta\varepsilon(u) = (\nabla\varepsilon_j \Delta u_j) = A_{ij}\Delta u_j W_{ik}(A_{km}u_m - f_k^{obs}) + u_i D_{ij}\Delta u_j$, то в точке минимума $\nabla\varepsilon_j = A_{ij}W_{ik}(A_{km}u_m - f_k^{obs}) + u_i D_{ij} = 0$ и шаг итерации Ньютона имеет вид:

$$u = (A^*WA + D)^{-1}A^*Wf^{obs}, \quad (18)$$

который формально совпадает с выражением для среднего решения в статистическом подходе (15).

4. Структура информации в статистических и вариационных обратных задачах

Таким образом, в статистических и вариационных обратных задачах ценная информация как знание того, какое воздействие на управляющие параметры выбрать для достижения цели при данном состоянии системы и с какой точностью цель будет достигнута, имеет векторную Δu и тензорную $(A^*WA)^{-1}$ (в ряде работ, например [13], матрица Фишера рассматривается как метрический тензор) составляющие.

Тензорная компонента ценной информации – некий «эллипсоид рассеивания», задающий неопределенность достижения цели – обратная информационная матрица Фишера.

Векторная компонента ценной информации – координаты цели в пространстве управляющих параметров Δu .

В линейном случае векторная часть ценной информации Δu связана с градиентом целевого функционала $\nabla\varepsilon$ с помощью линейного преобразования (7) $\Delta u = (A^*WA)^{-1}\nabla\varepsilon$. В этом смысле они эквивалентны, и ниже мы будем использовать в качестве векторной части ценной информации именно $\nabla\varepsilon$. Это

связано с тем, что уравнения в терминах расстояния до цели Δu неизвестны, а на практике используются сопряженные уравнения, позволяющие рассчитывать $\nabla \varepsilon$.

Строго говоря, $\nabla \varepsilon$ дает точное направление на цель, только если $F = E$. Однако итерации типа наискорейшего спуска в конечном итоге приводят к цели. В случае нелинейных задач вообще можно вычислить только $\nabla \varepsilon$.

Векторная компонента ценной информации зависит от невязки прогноза и наблюдения, тензорная определяется оператором эволюции и погрешностью наблюдений.

5. Векторная компонента ценной информации в сопряженных уравнениях

Таким образом, и при решении обратной задачи оптимизационным методом, и при решении обратной задачи в статистической постановке необходимым элементом является вычисление градиента целевого функционала.

Для задач, описываемых ЧДУ, наиболее эффективный метод расчета градиента целевого функционала основан на использовании сопряженных уравнений [10]. Обычно градиент целевого функционала $\nabla \varepsilon$ выражается через комбинацию физических и сопряженных параметров или даже просто через сопряженные параметры. В качестве эвристического примера рассмотрим модельную задачу определения $u(t, x)$ по наблюдениям $f^{obs}(t, x)$ в системе, заданной уравнением конвективного переноса [14]

$$\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + u(t, x) = 0. \quad (19)$$

Целевой функционал (для простоты записанный без учета погрешности) имеет вид скалярного произведения функций в L_2

$$\varepsilon(u) = \int (f(t, x) - f^{obs}(t, x))^2 / 2 dx dt. \quad (20)$$

Необходимые выражения можно получить из условий стационарности лагранжиана $L(u) = \varepsilon(u) + \int (\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + u(t, x)) \psi(x, t) dx dt$. На решении основной и сопряженной задачи

$$\frac{\partial \psi}{\partial t} + a \frac{\partial \psi}{\partial x} + (f - f^{obs}) = 0, \quad (21)$$

$\Delta L(u) = \Delta \varepsilon(u) = \int \Delta u(t, x) \psi(x, t) dx dt$ и градиент целевого функционала в L_2 равен сопряженной функции

$$\nabla \varepsilon = \psi(t, x). \quad (22)$$

Таким образом, при постановке обратных задач для систем ЧДУ можно считать, что сопряженные уравнения описывают перенос векторной компоненты ценной информации. При этом источниковый член $f - f^{obs}$ соответствует невязке расчета и наблюдения и описывает образование информации. Надо отметить, что в сопряженных уравнениях эволюция идет в обратном направлении по времени.

Сопоставление векторной компоненты ценной информации с градиентом целевого функционала позволяет более точно определить некоторые интуитивные понятия.

Цели можно сопоставить точку минимума целевого функционала в пространстве управляющих параметров.

Градиенту, который указывает, куда двигаться для достижения цели, можно соотнести понятия инструкции, рецепта.

Зависимость градиента от начального состояния системы u_0 соответствует зависимости информации от «предварительной осведомленности», «исходной информированности». Естественно, что при одном состоянии можно обойтись слабым (в смысле величины $\|\nabla \varepsilon\|$) управляющим воздействием, при другом – потребуется сильное.

Отношение типа «цена/качество» может быть определено при тихоновской регуляризации нулевого порядка. В точке минимума выполняются $\alpha \Delta u + \nabla \varepsilon = 0$ [15] и $((\alpha \Delta u + \nabla \varepsilon), \Delta u) = 0$. Соответственно, $\alpha(\Delta u, \Delta u) + (\nabla \varepsilon, \Delta u) = \alpha \|\Delta u\|^2 + \Delta \varepsilon$ и соотношение выигрыш/затраты (качество/цена) имеют вид коэффициента регуляризации $\alpha = \frac{\Delta \varepsilon}{\|\Delta u\|^2} = \frac{(\nabla \varepsilon, \Delta u)}{(\Delta u, \Delta u)}$.

Ценность данной информации (с точки зрения заданной цели) можно определить как максимальное изменение целевого функционала при движении вдоль градиента $\Delta \varepsilon = \tau \cdot \nabla \varepsilon$.

Если мы знаем точную информацию $\nabla \varepsilon_{true}$, то мы можем судить о векторе информации M (например, некоем информационном послании) по скалярному произведению $(M, \nabla \varepsilon_{true})$. Это позволяет определить степень неточности информации (расстояние) $\|M - \nabla \varepsilon_{true}\|$, нулевую информацию $(M, \nabla \varepsilon_{true}) = 0$, наличие информации с отрицательной ценностью (дезинформации) $(M, \nabla \varepsilon_{true}) < 0$, не уменьшающей, а увеличивающей целевой функционал. Степень новизны информации может быть оценена через ортогональную к $\nabla \varepsilon_{true}$ часть информации M_{\perp} .

Соотнесение ценной информации и сопряженных параметров (22) позволяет пролить свет на одну старую проблему. Сопряженные уравнения [10] используются для решения широчайшего круга практических задач, включая многие обратные задачи и задачи апостериорной оценки погрешности, при этом

соответствующая библиография содержит тысячи работ. Однако природа сопряженных уравнений остается неясной. С математической точки зрения в одних случаях они получаются из тождества Лагранжа для скалярного произведения, в других – из условия стационарности некоторого лагранжиана, в третьих – из функции Грина, что, вообще говоря, не указывает на наличие какой-либо общей природы. Тем не менее, при выводе из внешне разных предпосылок уравнения получаются идентичными. В прикладных дисциплинах в настоящее время возобладал сугубо прагматичный подход, в котором все внимание обращено на методы расчета сопряженных параметров, вообще без обсуждения их природы. Однако представляется очень странным, что объект, используемый стандартным образом в широчайшем классе приложений, не имеет единой (а, строго говоря, вообще никакой) физической интерпретации. Вследствие этого регулярно появляются работы, предлагающие некоторую трактовку «физического смысла сопряженных уравнений», например [16,17]. Настоящая работа также стимулирована поиском места сопряженных уравнений и параметров в общенаучной системе знаний.

На раннем этапе исследований один из основоположников теории сопряженных уравнений академик Г.И. Марчук обозначал сопряженную функцию как «функцию ценности информации» [10]. Смысл этого понятия достаточно расплывчат, и в настоящее время оно почти не употребляется. Тем не менее идея придать сопряженным параметрам информационный смысл выглядит исключительно перспективной. Как показано выше, градиент целевого функционала связан с сопряженной функцией (или даже тождественен ей (22)), соответственно, сопряженные параметры имеют смысл векторной компоненты ценной информации, а сопряженные уравнения описывают ее перенос. Трудности, связанные с поиском физического смысла сопряженных параметров, вызваны не столько проблемами трактовки собственно сопряженных уравнений и их решений, сколько отсутствием общепринятой количественной формулировки широко используемого на интуитивном уровне понятия информации.

Следует также отметить, что сообщества исследователей, использующих сопряженные уравнения [10,16,17], информационную динамику [13], статистическое решение обратных задач [4,6,11], скалярные инварианты тензорной информации [3,18] в основном не осознают наличия некоторой общей информационной основы.

Сопряженные уравнения и их решение (сопряженные параметры) представляют уникальную возможность визуализации переноса ценной информации [19]. На рис. 1 представлено поле плотности в сверхзвуковом течении при пересечении скачков уплотнения, на рис. 2 представлено поле сопряженной плотности. Течение описывается параболизированными уравнениями Навье-Стокса, целью задачи является оценка погрешности расчета плотности в точке, расположенной справа по центру.

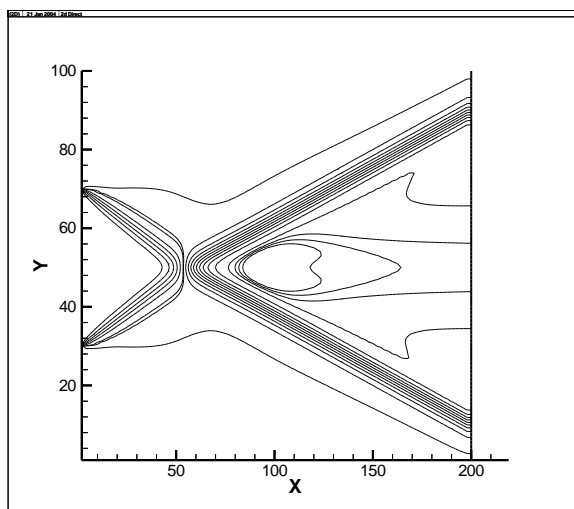


Рис. 1. Изолинии плотности

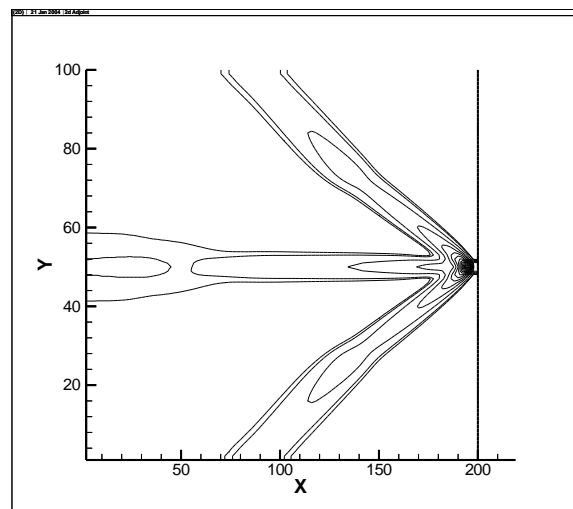


Рис. 2. Изолинии сопряженной плотности

В данной задаче вариация целевого функционала (погрешность расчета плотности в оцениваемой точке) в зависимости от локальной ошибки дискретизации имеет вид $\Delta\rho = \Psi_{\rho,ij} \delta\rho_{ij} + \Psi_{U,ij} \delta U_{ij} + \Psi_{V,ij} \delta V_{ij} + \Psi_{e,ij} \delta e_{ij}$. Локальные погрешности ($\delta\rho_{ij} \dots$) могут оцениваться, например, с помощью дифференциального приближения разностной схемы.

Сопряженная плотность $\Psi_{\rho,ij}$ соответствует весу, с которым локальная погрешность аппроксимации уравнения неразрывности входит в общую погрешность расчета плотности в оцениваемой точке. Соответственно, зоны, в которых вклад погрешности аппроксимации велик (и, например, нужно сетку мельчить) или, соответственно, мал (сетку можно сделать грубее), легко различимы.

Количественная оценка ошибки расчета является необходимым элементом верификации результатов расчетов, поле ценной информации позволяет оптимизировать данную оценку, используя визуальные средства.

Легко также видеть, что ценная информация распространяется в обратном направлении, если сравнивать с направлением эволюции течения.

В качестве более подробного примера рассмотрим обтекание сверхзвуковым потоком цилиндра при падении на него наклонного скачка уплотнения, описываемое нестационарными двумерными уравнениями Эйлера.

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho U_k)}{\partial x^k} = 0, \quad (23)$$

$$\frac{\partial(\rho U_i)}{\partial t} + \frac{\partial(\rho U_k U_i + P \delta_{ik})}{\partial x^k} = 0, \quad (24)$$

$$\frac{\partial(\rho E)}{\partial t} + \frac{\partial(\rho U_k h_0)}{\partial x^k} + q(x, y) = 0. \quad (25)$$

Здесь $U_1 = U, U_2 = V$, $\gamma = \frac{C_p}{C_v}$, $h = \frac{\gamma}{\gamma-1} \frac{P}{\rho} = \gamma e$, $e = \frac{RT}{\gamma-1}$, $E = \left(e + \frac{1}{2}(U^2 + V^2) \right)$,
 $h_0 = (U^2 + V^2)/2 + h$, $P = (\gamma-1)\rho e$.

При таком обтекании реализуется схема взаимодействия скачков четвертого типа по Эдни [20] (рис. 3), для которой характерно резкое повышения давления на теле. Интересно оценить возможность снижения давления на теле с помощью источников тепла (например, лазерного нагрева газа). Одним из вариантов является минимизация функционала

$$\varepsilon(q) = 1/2 \int_{\Gamma} p^n d\Gamma. \quad (26)$$

Соответствующие сопряженные уравнения имеют вид

$$\frac{\partial \Psi_{\rho}}{\partial t} + \frac{\partial \Psi_i}{\partial t} U^i + \frac{\partial \Psi_e}{\partial t} E + U^k \frac{\partial \Psi_{\rho}}{\partial X^k} + U^k U^i \frac{\partial \Psi_i}{\partial X^k} + (\gamma-1) \frac{\partial \Psi_k}{\partial X^k} e + U^k h_0 \frac{\partial \Psi_e}{\partial X^k} = 0, \quad (27)$$

$$\frac{\partial \Psi_i}{\partial t} \rho + \frac{\partial \Psi_e}{\partial t} \rho U^i + U^i \frac{\partial \Psi_k}{\partial X^i} + U^i \frac{\partial \Psi_i}{\partial X^k} + \frac{\partial \Psi_{\rho}}{\partial X^k} + h_0 \frac{\partial \Psi_e}{\partial X^k} + U^i U^k \frac{\partial \Psi_e}{\partial X^i} = 0, \quad (28)$$

$$\frac{\partial \Psi_e}{\partial t} + \gamma U^k \frac{\partial \Psi_e}{\partial X^k} + (\gamma-1) \frac{\partial \Psi_k}{\partial X^k} = 0. \quad (29)$$

Начальные условия $\Psi_{\rho, U, V, e}|_{t=0} = 0$; граничные – $\Psi_{\rho, U, V, e}|_{\partial\Omega} = 0$,
 $\Psi_{\rho}|_{\Gamma} = n P^{n-1} (\gamma-1) e$, $\Psi_e|_{\Gamma} = n P^{n-1} (\gamma-1) \rho$.

Градиент целевого функционала определяется через сопряженную энергию в виде

$$\nabla \varepsilon_q = \Psi_e(t, x, y). \quad (30)$$

Таким образом, и в данной задаче сопряженные параметры имеют смысл векторной компоненты ценной информации. На рис. 4 представлены изолинии сопряженной внутренней энергии $\nabla \varepsilon_q = \Psi_e(t, x, y)$ в некоторый момент времени.

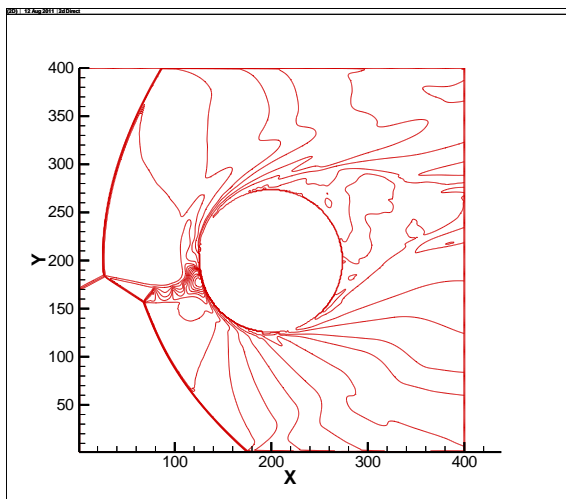


Рис. 3. Изолинии плотности

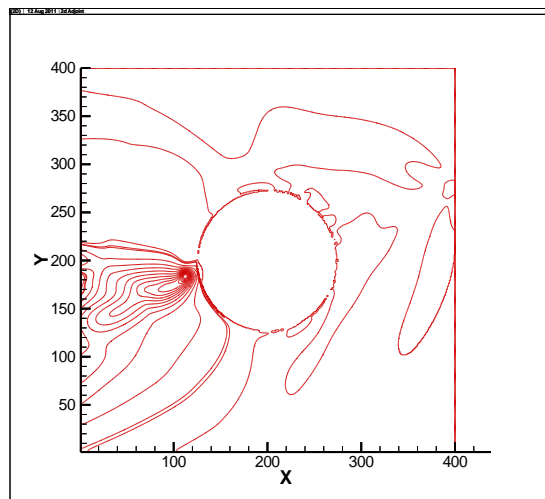


Рис. 4. Изолинии $\nabla \varepsilon_q = \Psi_e(t, x, y)$

Здесь вариация целевого функционала (интеграла давления) в зависимости от управляющего возмущения (источников тепла) имеет вид $\Delta \varepsilon = \Psi_{e,ij} \delta q_{ij}$.

Сопряженная энергия $\Psi_{e,ij}$ соответствует весу, с которым локальный источник тепла входит в общую вариацию давления. Соответственно, зоны, в которых влияние источника тепла на давление на поверхности велико, легко различимы, что позволяет осуществлять некоторое визуальное управление полем течения. Векторная часть ценной информации и здесь соответствует некоторому «рецепту», «инструкции», с помощью которых можно добиться цели (в случае рассматриваемой задачи – снижения давления на поверхности).

6. Тензорная информация

Тензорная компонента ценной информации (информационная матрица Фишера) является наиболее сложным элементом с точки зрения вычислений. Формально она может быть рассчитана с помощью уравнений информационной динамики типа Эйнштейна [13], но этот подход создает исключительные по своей сложности вычислительные проблемы.

Однако для $\varepsilon(u) = (A_{ij}u_j - f_i^{obs})W_{ik}(A_{km}u_m - f_k^{obs})$ вместо информационной матрицы Фишера $F = A^*WA$ ($F_{jm} = \frac{\partial T_i}{\partial u_j} W_{ik} \frac{\partial T_k}{\partial u_m}$, $A_{ij} = \partial f_i / \partial u_j$) можно использовать

гессиан невязки $H_{jm} = \frac{\partial^2 \varepsilon}{\partial u_j \partial u_m}$. Это связано с тем, что гессиан в окрестности

решения аппроксимирует матрицу Фишера. Действительно,

продифференцировав градиент $\nabla \varepsilon_j = \frac{\partial \varepsilon}{\partial u_j} = \frac{\partial T_i}{\partial u_j} W_{ik} (f_k - f_k^{obs})$, получим

$$H_{jm} = \frac{\partial^2 \varepsilon}{\partial u_j \partial u_m} = \frac{\mathcal{F}_i}{\partial u_j} W_{ik} \frac{\mathcal{F}_k}{\partial u_m} + \frac{\partial^2 f_i}{\partial u_j \partial u_m} W_{ik} (f_k - f_k^{obs}), \quad (31)$$

и, т.к. в окрестности оптимального решения $(f_k - f_k^{obs}) \approx 0$, то $H_{jm} \approx F_{jm}$.

Расчет матрицы Фишера с использованием гессиана существенно проще, чем [13], и может быть выполнен с помощью решения сопряженных уравнений второго порядка [21] либо дифференцированием градиента, полученного из решения сопряженных уравнений. Однако прямое численное дифференцирование градиента, полученного из сопряженной задачи, требует решения $2p$ прямых задач (где p – число управляющих параметров), что часто неприемлемо с точки зрения вычислительных затрат. Поэтому представляет больший интерес расчет спектра гессиана с помощью итерационного подхода (типа Ланцоша/Арнольди [22,23]). Он использует действие гессиана на вектор $H\delta u$. Оно может быть получено численным дифференцированием градиента $\nabla \varepsilon$ (a – параметр дифференцирования)

$$H\delta u = (\nabla \varepsilon(u + a\delta u) - \nabla \varepsilon(u)) / a. \quad (32)$$

Так как гессиан содержит относительно немного существенных собственных векторов, у которых собственные числа отличаются от нуля, то опираясь на них, можно построить аппроксимацию гессиана [23] $H_{reduced} = V_r \Lambda_r V_r^*$. Здесь матрица $V_r \in R^{p \times r}$ построена из r существенных собственных векторов H , соответствующих r наибольших собственным числам, которые собраны в диагональную матрицу $\Lambda_r = \text{diag}(\lambda_i) \in R^{r \times r}$.

Этот прием позволяет существенно снизить вычислительные усилия при расчете гессиана, если $r \ll p$.

7. Скалярные инварианты матрицы Фишера

Информационная матрица Фишера имеет скалярные инварианты: детерминант $\det(F)$ и след $\text{tr}(F)$.

В случае нормального распределения детерминант матрицы Фишера связан с количеством информации Шеннона [24]:

$$I_{sh} = 1/2 \log_2 \det(F), \quad (33)$$

а след матрицы соответствует количеству информации Фишера

$$I_F = \text{tr}(F). \quad (34)$$

Эти инварианты имеют множество приложений, в частности, в работе [24] они использованы для оценки количества информации, содержащейся в измерениях.

Количество информации Шеннона является основой современной теории информации [3] и используется в огромном количестве приложений.

Количество информации Фишера является менее популярным объектом, в ряде работ [18] оно используется в качестве инструмента для выведения уравнений физики, при этом ищется экстремум количества информации Фишера на решении уравнения типа Гамильтона-Якоби. В частности, так получают уравнения Шредингера в форме Маделунга [25,26]. Широко известный фильтр Калмана [27] определяет точку на траектории с эллипсоидом рассеяния и алгоритмически основан на поиске экстремума именно количества информации Фишера.

Следует отметить, что количества информации Шеннона и Фишера являются скалярами, поэтому они содержат относительно мало сведений о тензорной компоненте ценной информации и никак не связаны с векторной компонентой ценной информации. Количества информации Шеннона и Фишера достаточны для описания системы, в которой градиент равен нулю, а матрица Фишера имеет диагональную форму типа единичной матрицы (эллипсоид рассеивания имеет вид сферы). Состояние такой системы должно быть близко к экстремуму целевого функционала, а сам функционал должен быть изотропен в пространстве параметров u . Термодинамические системы, описываемые информацией/энтропией Шеннона/Больцмана, могут служить в качестве примера.

8. Обсуждение

В настоящее время понятие «информация» применяется в исключительно широком наборе контекстов разной степени формализации. Соответственно, степень соответствия рассмотренной выше ценной информации с другими «информациями» варьируется очень широко. Несколько родственных подходов здесь стоит упомянуть.

В работе [28] для динамики систем, обладающих целью, указано на важность сопряженных уравнений, описывающих поле «управляющей информации», в которой участвуют действующее лицо («деятель») и цель. Выдвинуто предположение о том, что успешное целенаправленное поведение возможно в случае самосопряженных уравнений, т. е. совпадения уравнений динамики и сопряженных уравнений. В работе утверждается, что такая самосопряженность реализуется в квантовой механике.

В работе [29] информация определяется как квантовый потенциал, возникающий в уравнении Шредингера, если оно выражено в гидродинамической форме (Маделунга). Этот квантовый потенциал получается из вариации информации Фишера.

В работах [30,31] использовано понятие прагматической информации, в которой подчеркивается воздействие информации на получателя. Ключевыми элементами являются новизна информации и ее подтверждение, они могут быть выражены через скалярное произведение или норму векторной ценной информации.

Формальная структура ценной информации очень близка к структуре информации в задачах статистического оценивания [4,5]. Различия касаются обозначений, интерпретации, используемой техники, особенно в областях, где применяются ЧДУ.

Выводы

В ряде задач [4,6-9] можно ввести формальное понятие ценной информации, обладающее многими свойствами интуитивных определений информации, таких как «полезная информация» [1], «ценность информации» [2].

В рассматриваемом в работе подходе ценная информация трактуется как знание того, какое воздействие на управляющие параметры выбрать, чтобы с известной точностью достигнуть некоторой цели при данном состоянии системы. Ценная информация имеет векторную и тензорную составляющие.

Векторной компоненте ценной информации соответствует градиент целевого функционала, указывающий некоторое направление к цели (точке минимума целевого функционала) в пространстве управляющих параметров.

В том случае если динамика задается системой частных дифференциальных уравнений, градиент целевого функционала выражается через комбинацию физических и сопряженных параметров, получается при решении сопряженных уравнений и эффективно визуализируется.

Тензорная компонента ценной информации задает неопределенность достижения цели («эллипсоид рассеивания») и соответствует информационной матрице Фишера F , которая в случае ЧДУ также может быть рассчитана с помощью сопряженных уравнений.

Литература

1. Бонгард М.М. Проблема узнавания. М.: Наука, 1967. 319 с.
2. Чернавский Д.С. Синергетика и информация. Динамическая теория информации. М.: УРСС, 2004. 287 с.
3. Shannon C. E. A Mathematical Theory of Communication / The Bell System Technical Journal. Vol. 27. July, October, 1948. P. 379-423, 623-656.
4. Fisher R.A. Theory of statistic estimation / Proc. Cambridge Phil. Soc. Vol.25. 1925. PP. 700-725.
5. Закс Ш. Теория статистических выводов, М.: Мир, 1975. 776 с.

6. Теребиж В.Ю. Введение в статистическую теорию обратных задач. М.: Физматлит, 2005. 375 с.
7. Le-Dimet F.X., Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations / *Tellus*. V. 38A . 1986. P. 97-110.
8. Lorenc A.C. Analysis methods for numerical weather prediction / *Q J R Meteorol Soc*. 1986. 112 (474):1177-1194.
9. Алифанов О.М., Артюхин Е.А., Румянцев С.В. Экстремальные методы решения некорректных задач. М.: Наука, 1988.
10. Марчук Г.И. Сопряженные уравнения и анализ сложных систем. М.: Наука, 1992. 336 с.
11. Устинов Е.А. К решению задачи численного дифференцирования методом статистической регуляризации // *Космические исследования*. Т.28. № 4. 1990. С. 545-554.
12. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1979.
13. Calmet X., Calmet J. Dynamics of the Fisher information metric / *Physical Review E* 71 (5), 056109, 2005.
14. Алексеев А.К. О ценной информации в задачах усвоения данных наблюдений / Труды IV Всероссийской конференции “Нелинейная динамика в когнитивных исследованиях – 2015”, Нижний Новгород, 2015. С. 16-18.
15. Лионс Ж.-Л. Оптимальное управление системами, описываемыми уравнениями с частными производными. М.: Мир, 1972.
16. Пупко В.Я., Зродников А.В., Лихачев Ю.И. Метод сопряженных функций в инженерно-физических исследованиях. М.: Энергоатомиздат, 1984.
17. Wang Q. and Gao J.-H. The drag-adjoint field of a circular cylinder wake at Reynolds numbers 20, 100 and 500 / *Journal of Fluid Mechanics*. V. 730. 2013. P. 145-161.
18. Frieden B.R., Soffer B.H. Lagrangians of physics and the game of Fisher-information transfer / *Phys. Rev. E*. 1995. V.52. P. 2274-2286.
19. Алексеев А.К., Бондарев А.Е. Применение сопряженных уравнений в задачах верификации и валидации расчетов // *Препринты ИПМ им. М.В. Келдыша*. 2011. № 49. 15 с.
URL: http://keldysh.ru/papers/2011/prep49/prep2011_49.pdf
20. Боровой В.Я. Течение газа и теплообмен в зонах взаимодействия ударных волн с пограничным слоем. М.: Машиностроение, 1983. 128 с.
21. Alekseev A.K., and Navon I.M. On Estimation of Temperature Uncertainty Using the Second Order Adjoint Problem // *Int. J. of Comput. Fluid Dynamics*. V. 16 (2). 2002. P. 113-117.
22. Flath H.P. et al. Fast algorithms for Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations // *SIAM J. Sci. Comp*. V. 33(1). 2011. P. 407-432.

23. Lehoucq R.B., Sorensen D.C., Yang C. ARPACK Users Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods // SIAM. 1998.
24. K. Singh, A. Sandu, M. Jardak, K. W. Bowman, and M. Lee. A practical method to estimate information content in the context of 4D-Var data assimilation // SIAM/ASA Journal of Uncertainty Quantification. V. 1. 2013. P. 106-138.
25. Madelung E. Quantum Theory in Hydrodynamical form // Zeit. f. Phys. V.40. 1927. P. 322- 325.
26. Reginatto M. Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information // Phys. Rev. A. V. 58. №. 3. 1998. P. 1775-1778.
27. Kalman R. E. A new approach to linear filtering and prediction problems // J. Basic Eng., Ser. D, 82, 34-45 (1960).
28. Shaw R., Kadar E., & Kinsella-Shaw J. Modelling systems with intentional dynamics: A lesson from quantum mechanics. In K. Pribram (Ed.) Appalacia II: Origins of self-organization // The Report of the Second Annual Appalachian Conference on Neurodynamics. Hillsdale NJ: Lawrence Erlbaum & Associates. 1995.
29. Hiley B. J., Pylkkanen P. Can Mind Affect Matter Via Active Information? // Mind&Matter V.3(2). 2005. P. 7-15.
30. Gernert D. Pragmatic Information: Historical Exposition and General Overview // Mind & Matter. V4(2). 2006. P. 141-167.
31. Weizsacker E. von and Weizsacker C. von. Wiederaufnahme der begrifflichen Frage: Was ist Information? // In Nova Acta Leopoldina 206, Band 37/1, J.A. Barth, Leipzig, 1972. P. 535-555.

Оглавление

Введение	3
1. Интуитивное понятие информации, связанное с наличием цели	3
2. Вариационная постановка обратных задач.....	4
3. Статистическая постановка обратных задач	6
4. Структура информации в статистических и вариационных обратных задачах	8
5. Векторная компонента ценной информации в сопряженных уравнениях.....	9
6. Тензорная информация	14
7. Скалярные инварианты матрицы Фишера.....	15
8. Обсуждение.....	16
Выводы	17
Литература	17