

§7. Регуляризация

Для решения некорректных задач советским математиком Тихоновым был предложен простой, но чрезвычайно эффективный метод, называемый *регуляризацией* и основанный на привлечении дополнительной априорной информации о решении, которая может быть как качественной, так и количественной. Например, можно искать решение, обладающее определенной гладкостью, и / или максимально близкое к некоторому вектору \mathbf{y}^0 .

Концепция регуляризации сводится к замене исходной некорректной задачи

$$\mathbf{A} \cdot \mathbf{y} + \boldsymbol{\sigma} = \mathbf{b} \quad (40)$$

задачей минимизации следующей функции:

$$\Omega(\mathbf{y}, \lambda) = |\mathbf{A} \cdot \mathbf{y} - \mathbf{b}|^2 + \lambda \cdot |\mathbf{y} - \mathbf{y}^0|^2. \quad (41)$$

Здесь λ – это малый положительный *параметр регуляризации*, который необходимо подобрать определенным способом. Отметим, что, если рассматривать не дискретную, а непрерывную задачу (т. е. функцию $y(x)$ вместо вектора \mathbf{y}), то $\Omega(y(x), \lambda)$ будет представлять собой не функцию, а *функционал*, исторически называемый *функционалом Тихонова*.

Минимизируя функцию $\Omega(\mathbf{y}, \lambda)$, можно получить *регуляризованное решение* $\mathbf{y}(\lambda)$, зависящее от параметра λ . Из (41) хорошо ясен его смысл: при малых $\lambda \sim 0$ проблема поиска $\Omega(\mathbf{y}, \lambda)$ близка к (некорректной) исходной задаче, а при больших λ задача поставлена корректно, но ее решение далеко от решения исходной обратной задачи. А именно, чем больше параметр регуляризации, тем ближе решение к априорной оценке \mathbf{y}^0 . Очевидно, что на практике необходимо выбирать промежуточные значения λ .

Можно показать, что в линейном случае задача о минимизации функции $\Omega(\mathbf{y}, \lambda)$ эквивалентна к следующей системе линейных алгебраических уравнений:

$$(\mathbf{A}^T \mathbf{A} + \lambda \cdot \mathbf{I}) \cdot \mathbf{y} = \mathbf{A}^T \mathbf{B} + \lambda \cdot \mathbf{y}^0. \quad (42)$$

Повторимся, что параметр регуляризации λ необходимо подобрать некоторым оптимальным способом. Очевидно, что им будет некоторое промежуточное значение λ , устанавливающее определенный компромисс между приемлемой обусловленностью и близостью к исходной задаче. Отметим, что регуляризационный подход сводит некорректную задачу (40) к условно-корректной (по Тихонову) задаче отыскания решения системы (40), которое, в силу линейности задачи, является единственным и устойчивым.

Приведем теперь регуляризованное решение вырожденной системы двух уравнений с матрицей \mathbf{A} (см. рис. 19) и вектором правых частей \mathbf{b} :

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1.05 & 2.1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 10 \\ 10 \end{pmatrix}. \quad (43)$$

На рис. 23 представлена зависимость невязки Δ полученного решения $\mathbf{y}(\lambda)$ от параметра регуляризации λ , а само решение (пара элементов неизвестного вектора \mathbf{y}) представлено на рис. 24. Важно подчеркнуть, что точного решения исходной системы и, следовательно, системы (42) при $\lambda=0$ не существует.

Заключительным шагом регуляризации является выбор оптимального параметра регуляризации λ . Имеется, по крайней мере, два соображения, исходя из которых, можно выбрать параметр регуляризации. В рассматриваемом примере применим критерий, опирающийся на подбор нормы невязки, равной априорной оценке погрешностей задания входных данных: матрицы \mathbf{A} и вектора \mathbf{b} , т. е. величине $|\delta \mathbf{A}| + |\delta \mathbf{b}|$. Например, имея эти данные, можно выбрать норму невязки и, соответственно, параметр λ и решение $\mathbf{y}(\lambda)$, которые отмечены на рис. 23 и 24 пунктирами (а именно, $\lambda=2$).

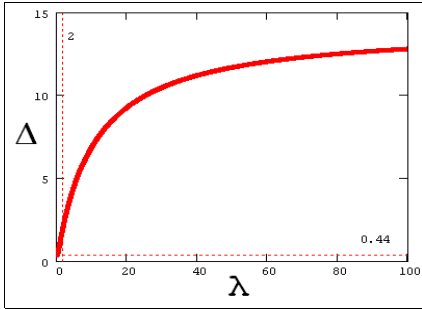


Рис. 23. Зависимость невязки Δ от параметра регуляризации λ

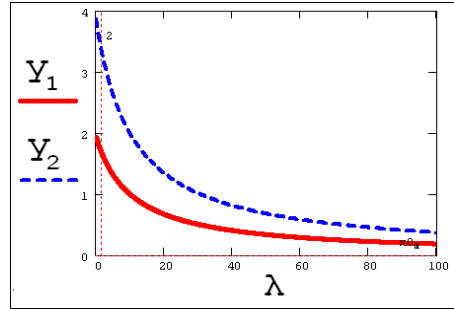


Рис. 24. Регуляризованное решение y в зависимости от λ

Другим вариантом выбора λ , не требующим никаких априорных соображений относительно погрешностей модели, является так называемый *квазиоптимальный* метод, который будет рассмотрен ниже.

В общем (нелинейном) случае минимизация $\Omega(y, \lambda)$ должна производиться по всем компонентам вектора y , что представляет собой довольно громоздкую вычислительную задачу. Как известно, поиск глобального экстремума осложняется не только благодаря многомерности задачи, но и из-за возможности существования нескольких локальных минимумов. Подчеркнем, что в случае линейной задачи (когда \mathbf{A} – это матрица) формула (42) дает тот же результат, что и численная минимизация функционала (41).

Приведем в качестве еще одного примера применения регуляризации решение некорректной линейной задачи интерпретации измерений, которая была представлена на рис. 2 и 3. В качестве модельного сигнала будем использовать квадратичную параболу, а в качестве априорной оценки – линейный профиль \mathbf{Y}_0 (рис. 25). Сделаем мы это для того, чтобы соблюсти честность, попытаюсь подогнать регуляризованное решение не к исходной (вообще говоря, неизвестной) модели, а к иной зависимости.

Поскольку эта задача также линейная, дискретизация интеграла (6) приводит к построению соответствующей матрицы A . Для построения регуляризованного решения следует решить СЛАУ (42), зависящую от λ как от параметра. В результате получим регуляризованное решение как функцию параметра регуляризации: $y(\lambda)$.

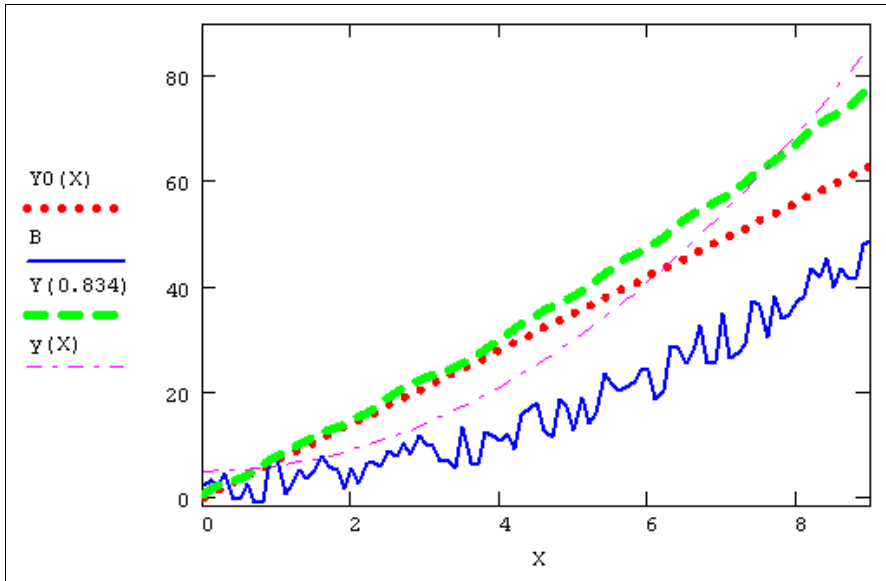


Рис. 25. Исходный сигнал y , его априорная оценка Y_0 , вектор измерений b и регуляризованное решение $Y(\lambda)$

Отметим, что соответствующая невязка системы уравнений

$$\Delta(\lambda) = |Ay(\lambda) - b| \quad (44)$$

также является функцией λ (рис. 26).

Для реконструкции можно использовать такое значение λ , которое соответствует глобальному минимуму зависимости $\Delta(\lambda)$. Иными словами, необходимо решить еще одну задачу на минимум, но уже другой функции $\Delta(\lambda)$, задаваемой формулой (44). Подчеркнем, что само определение этой функции подразумевает вложенную минимизацию функционала Тихонова (41), которую

мы, правда, оформили посредством решения системы линейных уравнений (42). Таким образом, для построения реконструкции сигнала (она показана на том же рис. 25 штрих-пунктиром) придется решить две задачи минимизации.

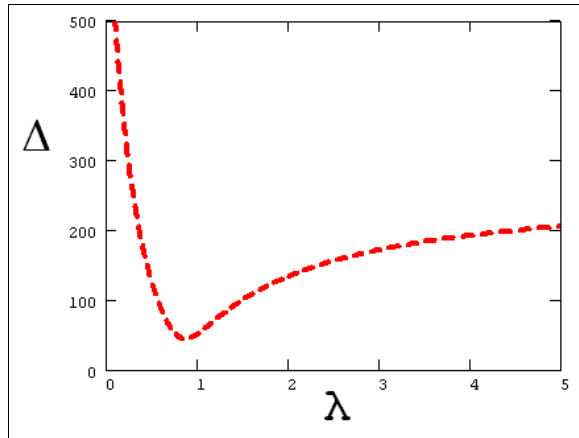


Рис. 26. Невязка $\Delta(\lambda)$, даваемая регуляризованным решением $y(\lambda)$ (41)

Вместо квазиоптимального способа определения λ можно применить разобранный в предыдущем примере принцип невязки. Напомним, что он подразумевает выбор параметра регуляризации, с которым невязка приблизительно равна сумме погрешностей измерений (т. е. заданий правой части) и аппроксимации (погрешности, заложенной в матрице \mathbf{A}).

Заметим, что конкретный вид функционала Тихонова может быть не обязательно таким, как приведено в формуле (41). Например, вместо (41) можно использовать следующее выражение:

$$\Omega(\mathbf{y}, \lambda) = |\mathbf{A}\mathbf{y} - \mathbf{b}| + \lambda \cdot |\mathbf{y} - \mathbf{y}^0|. \quad (45)$$

Опять-таки, λ — это малый параметр регуляризации. Решение той же задачи (43) для формы функционала Тихонова (45) показано на рис. 27-28.

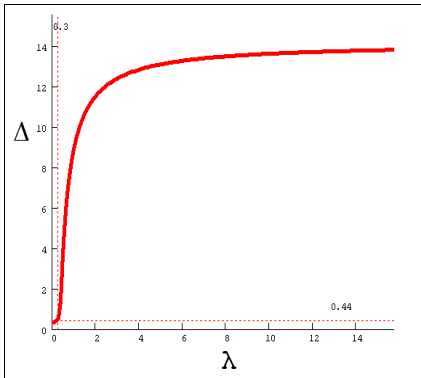


Рис. 27. Зависимость невязки от параметра регуляризации λ для функционала Тихонова в форме (45)

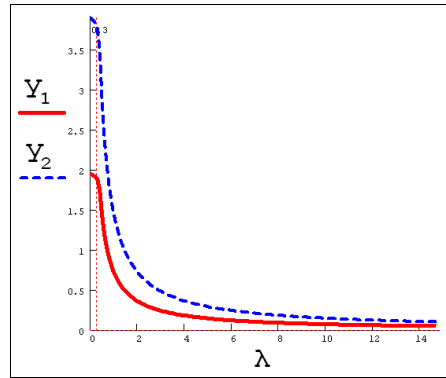


Рис. 28. Регуляризованное решение для функционала Тихонова (45)

В завершение разговора о регуляризации, обратимся к задаче томографии и покажем, как можно построить регуляризованное решение в этом случае. Напомним, что переопределенная система уравнений для веерной томографии была нами построена в §2.

Мы собираемся учитывать априорную информацию и искать решение, максимально близкое к некоторому профилю показателя преломления верхней атмосферы, т.е. к некоторому вектору \mathbf{n}^0 . Концепция регуляризации применительно к задаче томографии сводится к ее замене на задачу о минимизации функционала Тихонова (41).

Опять-таки, в силу линейности рассматриваемой задачи томографии, проблема минимизации функционала (41) сводится к системе линейных уравнений (42), решая которую, можно получить регуляризованное решение, зависящее от λ . Напомним, что при малых $\lambda \sim 0$ обусловленность системы (42) близка к плохой обусловленности исходной системы, а при больших λ система (42) обусловлена хорошо, но её решение далеко от решения искомой задачи томографии (чем больше λ , тем ближе решение к вектору \mathbf{n}^0). Регуляризационный подход сводит некорректную задачу томографии к условно-корректной (по

Тихонову) задаче отыскания решения системы (42), которое, в силу линейности задачи, является единственным и устойчивым.

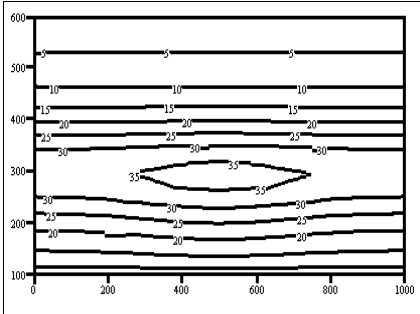


Рис. 29. Модель атмосферы – поле показателя преломления $n(x,z)$

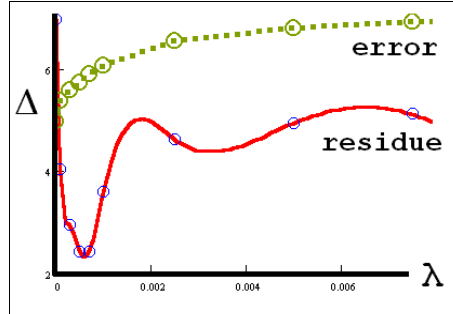


Рис. 30. Зависимость невязки от параметра регуляризации

Проиллюстрируем сказанное конкретным вычислительным примером. Зададим некоторую модель атмосферы (рис. 29), которая образована наложением на регулярный высотный профиль (с единственным максимумом) неоднородности, расположенной в центре, и промоделируем процесс измерений. Высота ИСЗ в компьютерном эксперименте составляла 900 км, два приёмника располагались в точках с координатами 0 км и 500 км по поверхности Земли в плоскости пролёта ИСЗ.

Дискретизация интегральных уравнений проводилась методом кусочно-планарной аппроксимации на сетке 21×21 узел. Такая крупная (по сравнению с реальными задачами томографии) сетка была взята нами для того, чтобы проверить выдвигаемую методику с помощью точного метода решения (42), не прибегая к приближённым итерационным алгоритмам. Поскольку число арифметических операций, требующихся компьютеру для завершения точного метода, порядка $(M_K)^6$, то для расчёта на сетке 100×100 узлов необходимо время в 15,000 раз большее. Решение системы (42) осуществим точным методом Гаусса с выбором главного элемента.

Задав реалистичный профиль $n^0(z)$ в качестве априорной информации, представим его в дискретной форме в виде вектора

n^0 . Для того, чтобы определить оптимальный параметр регуляризации, проведём серию расчётов с различными λ , контролируя невязку. На рис. 30 изображена зависимость нормы невязки решения регуляризованной задачи от параметра регуляризации λ . На том же рисунке изображена суммарная ошибка (т.е. норма отклонения решения от модельного распределения, которое в данном случае известно априори).

Как уже было сказано, для реконструкции можно использовать «квазиоптимальное» значение λ , соответствующее глобальному минимуму зависимости $\varepsilon(\lambda)$ (рис. 30), либо применить т.н. принцип невязки, который требует выбора λ , с которым невязка приблизительно равна сумме погрешностей измерений (т.е. заданий правой части) и аппроксимации. Результаты реконструкции для разных λ изображены на рис.31-32. Видно, что наилучшее совпадение с модельным полем (рис. 29) демонстрирует «квазиоптимальное» решение (рис. 31).

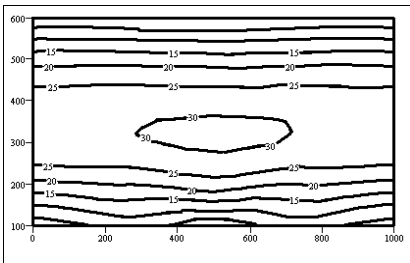


Рис. 31. Реконструкция $n(x,z)$
для $\lambda=0.001$

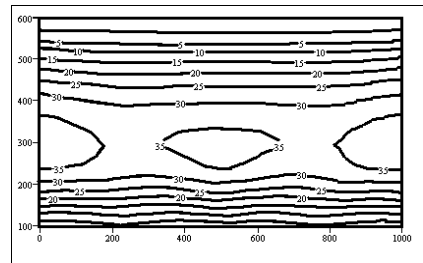


Рис. 32. Реконструкция $n(x,z)$
для $\lambda=0.006$