Содержание

1 Математическая постановка задачи

1.1 Постановка задачи

В экспериментальных исследованиях результатом измерения является конечный набор значений $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N)$, где ξ_i – наклонное содержание NO_2 для зинитного угла направления визирования z_i. По набору этих значений требуется оценить значения вертикального распределения NO_2 $n(h)$ в заданном наборе точек области определения.

В задаче восстановления NO_2 результат измерения имеет вид:

$$\xi_i = \int_0^H m(z_i, h) n(h) dh + \nu_i = \sum_{j=1}^{M} m(z_i, h_j) n(h_j) \triangle h + \nu_i = (a_i, n) + \nu_i \qquad (1)$$

где a_i – послойная воздушная масса, которая рассчитывается из модели переноса излучения. Она связывает наклонное содержание NO_2 $\boldsymbol{\xi}$ с вертикальным распределением концентраций $n(h)$.

Схему измерений можно представить в виде:

$$\boldsymbol{\xi} = A\boldsymbol{n} + \boldsymbol{\nu} \qquad (2)$$

На основании измерения вектора $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N)$ требуется оценить значение вектора $Un \in \mathbb{R}^N$, где $U : \mathbb{R}^N \to \mathbb{R}^N$ – оператор, которому сопоставляется матрица с учётом априорной информации о значениях координат вектора $\boldsymbol{n}$.

1.2 Обратные задачи

Рассмотрим неизвестный сигнал $f \in \mathcal{L}^2(\mathbb{R}^d)$ и предположим, что последний проверяется каким-либо сенсорным устройством, что приводит к вектору данных $\mathbf{y} = [y_1, \dots, y_L]$ из L измерения. Восстановление f из вектора данных $\mathbf{y}$ называется обратной задачей.

Следующие допущения являются стандартными:

- Чтобы учесть неточности датчиков предполагается, что вектор данных $\mathbf{u}$ это результат случайного вектора $\mathbf{Y} = [Y_1, \dots, Y_L] = [\langle f, \phi_1 \rangle, \dots, \langle f, \phi_L \rangle]$, колеблющийся в соответствии с некоторым распределением шума. Записи $\mathbb{E}[\mathbf{Y}] = \widetilde{\mathbf{y}}$ называются идеальными измерениями(измерения, которые были бы получены в отсутствие шума).

- Измерения предполагаются несмещенными и линейными, т.е. $\mathbb{E}[\mathbf{Y}] = \Phi^* f$, для некоторых функционалов выборки $\{\phi_1, \dots, \phi_L\} \subset \mathcal{L}^2(\mathbb{R}^d)$ моделирующих систему сбора данных.

1.3 Некорректность обратных задач

Для решения обратной задачи можно аппроксимировать среднее $\mathbb{E}[Y]$ по его одновыборочной эмпирической оценке $\mathbf{y}$ и решить линейную задачу:

$$\mathbf{y} = \mathbf{G}\alpha \qquad (3)$$

К сожалению, такие задачи в целом некорректны:

- **Решений может не быть**. Если $\mathbf{G}$ не сюръективен, $\mathcal{R}(\mathbf{G}) \subseteq \mathbb{R}^L$. Следовательно, зашумленный вектор данных $\mathbf{y}$ не гарантируется принадлежность к $\mathcal{R}(\mathbb{G})$.

- **Может существовать более одного решения**. Если $L < N$ действительно (или, в более общем смысле, если $\mathbf{G}$ не инъективен), $\mathcal{N}(\mathbf{G}) \neq \{\mathbf{0}\}$. Следовательно, если α^* является решением (3), тогда $\alpha^* + \beta$ также является решением $\forall \beta \in \mathcal{N}(\mathbf{G})$:

$$\mathbf{G}(\alpha^* + \beta) = \mathbf{G}\alpha^* + \mathbf{G}\beta = \mathbf{G}\alpha^* = \mathbf{y}$$

- **Решения могут быть численно неустойчивыми**. Если $\mathbf{G}$ например, сюръективно, то $\mathbf{G}^\dagger = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}$ является правой инверсией $\mathbf{G}$, а также $\alpha^*(\mathbf{y}) = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}$ является решением (3). У нас есть тогда

$$||\alpha^*(\mathbf{y})|| \leq ||\mathbf{G}||_2||(\mathbf{G}^T\mathbf{G})^{-1}||_2||\mathbf{y}||_2 = \underbrace{\frac{\sqrt{\lambda_{\max}(\mathbf{G}^T\mathbf{G})}}{\lambda_{\min}(\mathbf{G}^T\mathbf{G})}}_{\text{может быть очень большим}} ||\mathbf{y}||_2, \qquad \forall \mathbf{y} \in \mathbb{R}^L.$$

1.4 Регуляризация обратных задач

Линейная система (3) не только некорректна, но и бессмысленна: точное совпадение измерений нежелательно, так как последние на практике искажаются инструментальными шумами.

Вместо этого более разумный подход состоит в решении обратной задачи с помощью штрафной задачи оптимизации, сопоставляя физические данные с априорными представлениями аналитика о решении (например, гладкость, разреженность) с помощью терминов точности данных и регуляризации соответственно:

$$\min_{\alpha \in \mathbb{R}^N} F(\mathbf{y}, \mathbf{G}\alpha) + \lambda \mathcal{R}(\alpha)$$

Различные величины, участвующие в приведенном выше уравнении, можно интерпретировать следующим образом:

- $F : \mathbb{R}^L \times \mathbb{R}^L \to \mathbb{R}_+ \cup \{+\infty\}$ представляет собой функционал стоимости/верности данных/потери, измеряющий несоответствие между наблюдаемыми и прогнозируемыми измерениями $\mathbf{y}$, а также $\mathbf{G}\alpha$ соответственно.

- $\mathcal{R} : \mathbb{R}^N \to \mathbb{R}_+ \cup \{+\infty\}$ является функционалом регуляризации/штрафа, отдающим предпочтение простым и хорошим решениям (обычно с конечным числом степеней свободы).

- $\lambda > 0$ является параметром регуляризации/штрафа, который контролирует степень регуляризации, помещая функционал регуляризации и функционал стоимости в одинаковую шкалу.

1.5 Выбор функционала потерь

Функционал потерь можно выбрать как отрицательное логарифмическое правдоподобие данных **y**:

$$F(\mathbf{y}, \mathbf{G}\alpha) = -\ell(\alpha|\mathbf{y}) = -\log p_{Y_1,\ldots,Y_L}(y_1,\ldots,y_L|\alpha).$$

Когда распределение шума не известно полностью или вероятность слишком сложна, можно также использовать общие ℓ_p функционалы стоимости:

$$F(\mathbf{y}, \mathbf{G}\alpha) = ||\mathbf{y} - \mathbf{G}\alpha||_p^p = \sum_{i=1}^{L} \left| y_i - \sum_{n=1}^{N} G_{in}\alpha_n \right|^p .$$

где $p \in [1, +\infty]$ обычно выбирается в соотвествие с поведеднием хвоста распределения шума.

1.6 Проксимальные алгоритмы

Большинство задач оптимизации, используемых на практике для решения обратных задач, имеют вид:

$$\min_{\mathbf{x}\in\mathbb{R}^N} \mathcal{F}(\mathbf{x}) + \mathcal{G}(\mathbf{x}) + \mathcal{H}(\mathbf{K}\mathbf{x}) \tag{4}$$

где:
*$\mathcal{F}: \mathbb{R}^N \to \mathbb{R}$ выпукла и дифференцируема, с $\beta-$Липшицевым непрерывный градиентом,
*$\mathcal{G}: \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ и $\mathcal{H}: \mathbb{R}^N \to \mathbb{R} \cup \{+\infty\}$ две собственные полунепрерывные снизу и выпуклые функции с простыми проксимальными операторами.

- $\mathbf{K}: \mathbb{R}^N \to \mathbb{R}^M$ является линейным оператором.

Задачи вида (4) могут быть решены с помощью итерационных **проксимальных алгоритмов**:

- **Primal-dual splitting**: решает проблемы вида
$$\min_{\mathbf{x}\in\mathbb{R}^N} \mathcal{F}(\mathbf{x}) + \mathcal{G}(\mathbf{x}) + \mathcal{H}(\mathbf{K}\mathbf{x})$$
- **Chambolle Pock splitting**: решает проблемы вида
$$\min_{\mathbf{x}\in\mathbb{R}^N} \mathcal{F}(\mathbf{x}) + \mathcal{H}(\mathbf{K}\mathbf{x})$$
- **Douglas Rachford splitting/ADMM**: решает проблемы вида
$$\min_{\mathbf{x}\in\mathbb{R}^N} \mathcal{F}(\mathbf{x}) + \mathcal{H}(\mathbf{x})$$
- **Forward-Backward splitting/APGD**: решает проблемы вида
$$\min_{\mathbf{x}\in\mathbb{R}^N} \mathcal{F}(\mathbf{x}) + \mathcal{G}(\mathbf{x})$$

2 Расчёты

Приведём

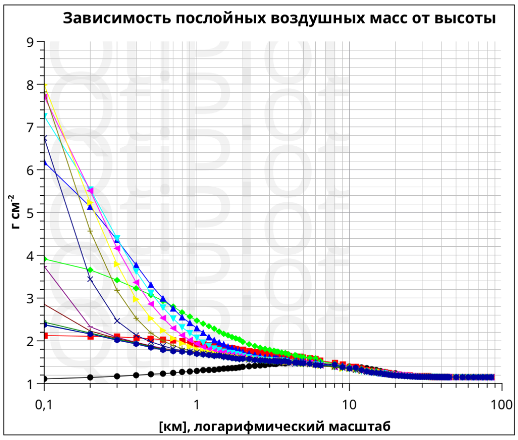
Зависимость послойных воздушных масс от высоты
г см$^{-2}$
[км], логарифмический масштаб
0,1
1
10
100

Рис.1

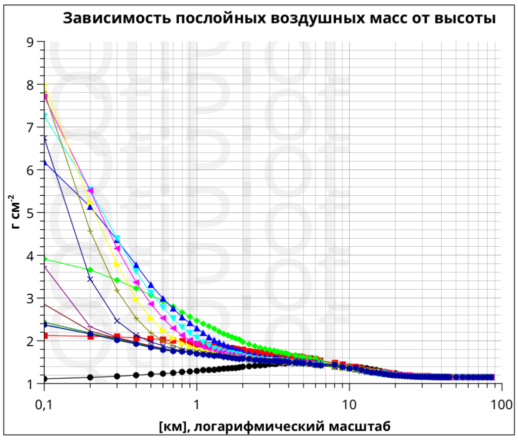
Зависимость послойных воздушных масс от высоты
г см$^{-2}$
[км], логарифмический масштаб
0,1
1
10
100

Рис.2