

# Содержание

<b>1</b>	<b>Математическая постановка задачи</b>	<b>2</b>
1.1	Обратные задачи . . . . .	2
1.2	Некорректность обратных задач . . . . .	2
1.3	Регуляризация обратных задач . . . . .	3
1.4	Выбор функционала потерь . . . . .	3
1.5	Проксимальные алгоритмы . . . . .	3

# 1 Математическая постановка задачи

## 1.1 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  $L$  измерений. 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  $\mathbb{E}[\mathbf{Y}] = \tilde{\mathbf{y}}$  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.

## 1.2 Некорректность обратных задач

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

$$\mathbf{y} = \mathbf{G}\alpha \tag{1}$$

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  $\mathcal{R}(\mathbf{G})$ .
- **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  $\alpha^*$  является решением (1), 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  $\alpha^*(\mathbf{y}) = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{y}$  является решением (1). У нас есть тогда

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

### 1.3 Регуляризация обратных задач

Линейная система (1) 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:

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

### 1.4 Выбор функционала потерь

Функционал потерь можно выбрать как отрицательное логарифмическое правдоподобие данных  $\mathbf{y}$ :

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

Когда распределение шума не известно полностью или вероятность слишком сложна, можно также использовать общие  $\ell_p$  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 соответствие с поведением хвоста распределения шума.

### 1.5 Проксимальные алгоритмы

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

$$\min_{\mathbf{x} \in \mathbb{R}^N} \mathcal{F}(\mathbf{x}) + \mathcal{G}(\mathbf{x}) + \mathcal{H}(\mathbf{K}\mathbf{x}) \quad (2)$$

где:

$*\mathcal{F} : \mathbb{R}^N \rightarrow \mathbb{R}$  выпукла и дифференцируема, с  $\beta$ -Липшицевым непрерывный градиентом,

$*\mathcal{G} : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  и  $\mathcal{H} : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  две собственные полунепрерывные снизу и выпуклые функции с простыми проксимальными операторами.

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

Задачи вида (2) 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 формы

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

- **Douglas Rachford splitting/ADMM**: решает проблемы формы

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

- **Forward-Backward splitting/APGD**: решает проблемы формы

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