УДК 004.02

С.М. Вовк

КРИТЕРІЙ МІНІМУМУ ПРОТЯЖНОСТІ

Анотація. Запропоновано критерій обробки даних, який полягає у вимозі мінімізувати протяжність функції, використовуваної для пошуку рішення, та, зокрема, мінімізувати протяжність області незбігу даних з їх моделлю. Наведено трактування поняття «протяжність функції» для випадків відсутності та наявності шуму, а також загальні формулювання запропонованого критерію для різних варіантів його застосування. Позначено його зв'язок з іншими критеріями обробки даних та вказані умови доцільності використання.
Ключові слова: критерій, мінімум протяжності, обробка даних.

Вступ

У розв'язку задач обробки даних ключову роль грає критерій якості обробки, який має забезпечувати отримання найкращого рішення. У низці випадків таке рішення можна отримати аналітично, наприклад, для задач обробки даних в шумі з законами розподілу Гаусса, Лапласа або Коші [1]. Але у більшості випадків цього не можна зробити з тієї причини, що математичні моделі спотворень, шумів, завад та/або оброблюваних даних невідомі або відомі неточно. В цих умовах велике значення набувають ті апріорні відомості, які дозволяють компенсувати незнання або неточне знання моделей. Так, неточне знання статистичних характеристик шумів і завад призводить до робастних методів обробки даних [2], а незнання математичних моделей даних – до формулювання критеріїв та/або обмежень, які в досить загальній формі описують бажані властивості рішення [3]. У даній роботі представлено критерій обробки даних, який заснований на вимозі мінімізувати протяжність функції [4], що використовується для пошуку рішення, і який далі буде згадуватися як критерій мінімуму протяжності.

Постановка задачі

Постановка розв'язуваної задачі полягає в отриманні математичних формулювань критерію обробки даних, заснованого на понятті «протяжність функції».

Метою даної роботи є опис критерію мінімуму протяжності і особливостей його застосування.

Огляд літератури

Основою критерію мінімуму протяжності є ідеї методів М-оцінювання [2] (або узагальненої максимальної правдоподібності [5]) і мінімуму тривалості [6]. Згідно [2], метод М-оцінювання полягає у розв'язку задачі:

$$\min_{\boldsymbol{\theta}} \sum_{i=1}^{n} \psi[x_i; \boldsymbol{\theta}], \qquad (1)$$

де $\boldsymbol{\theta}$ є невідомим параметром, x_i є i-м елементом спостережуваних даних, а ψ є довільною функцією. Функцію ψ (це позначення взяте з [7], а в [2] використовується символ ρ) називають вартісною функцією [8], функцією втрат або ваговою функцією [7] залежно від потрібного сенсу, який може й не мати імовірнісного трактування [5]. Співвідношення (1) виражає досить загальну ідею постановки будь-якої задачі у вигляді задачі оптимізації. Проте важливим є факт використання в ньому апарату вартісних функцій, які дозволяють формувати необхідні критерії обробки даних. Зокрема, використання в якості ψ логарифма спільної щільності ймовірності призводить до традиційного критерію максимальної правдоподібності [2].

Згідно [6], метод мінімуму тривалості призначений для розв'язку задачі відновлення фінітної часової функції за відомою ділянкою її спектра та полягає в мінімізації функціоналу «узагальненої тривалості»

$$D^{(\alpha,\beta,\dots)}[\hat{f}(t)] = \int_{-\infty}^{\infty} \psi^{(\alpha,\beta,\dots)}[\hat{f}(t)]dt \qquad (2)$$

її оцінки $\hat{f}(t)$ для заданого набору значень параметрів $\alpha, \beta, \dots$. Вважаючи

$$\hat{f}(t) = g(t) + s(t), \qquad (3)$$

де $g(t)$ є відомою часовою функцією з фінітним спектром в заданій смузі частот, а $s(t)$ є невідомою часовою функцією зі спектром поза межами заданної смуги частот, маємо варіаційну задачу щодо $s(t)$. Якщо $s(t)$ подати рядом:

$$s(t) = \sum_{j=1}^{M} a_j r_j(t), \qquad (4)$$

де a_j є невідомими коефіцієнтами, а $r_j(t)$ є відомими функціями зі спектром поза заданою смугою частот, то після підстановки (4) в (3) і далі (3) в (2) отри-

ISSN 1562-9945

муємо задачу мінімізації цільової функції від M невідомих a_j . При цьому в якості $r_j(t)$ для дискретного випадку запропоновано брати базисні функції дискретного перетворення Фур'є, що відповідають спектральним компонентам поза заданою смугою частот, а в якості функції $\psi^{(\alpha,\beta,...)}$ – функцію

$$\psi^{(\alpha,\beta)}(x) = [|x|^2 + \alpha^2]^{\beta} - \alpha^{2\beta}; \quad 0 < \beta < 1, \tag{5}$$

де α є параметром згладжування.

В [9]-[10] наведено розвиток ідей зазначених методів. Він полягає в побудові «супермножини» вартісних функцій, яка керується набором з трьох вільних параметрів та містить в собі характеристичну функцію множини. Зазначена «супермножина» подається вартісною функцією [9]:

$$\psi_S^{(\alpha,\beta,q)}(x) = k_S^{(\alpha,\beta,q)}[(1+|x/\alpha|^q)^{\beta/q} - 1], \tag{6}$$

де α – параметр згладжування; q – параметр ступеня цього згладжування, причому $0 < q < \infty$; β – параметр форми вартісної функції, причому $-\infty < \beta \le 1$ та $\beta < q$; $k_S^{(\alpha,\beta,q)} = 1/[(1+|x_1/\alpha|^q)^{\beta/q} - 1]$, де x_1 є точкою нормування функції (6) на одиницю, тобто $\psi_S^{(\alpha,\beta,q)}(x_1) = 1$ (звичайно $x_1 = 1$). Параметри α, β та q мають сенс вільних параметрів, які дозволяють істотно змінювати поведінку вартісної функції (6). В якості своїх граничних випадків дана супермножина вміщує багато відомих вартісних функцій (зокрема, (5)), а розроблений в [10] механізм їх перетворення дозволяє виконувати налаштування процесу обробки на поточне шумове оточення.

Основна частина

Згідно [4], протяжність абстрактної одновимірної функції можна визначити як міру множини її ненульових значень у вигляді функціоналу:

$$E[f(x)] = \int_{-\infty}^{\infty} \chi[f(x)]dx, \tag{7}$$

де $\chi[f(x)] = \begin{cases} 1; & f(x) \ne 0 \\ 0; & f(x) = 0 \end{cases}$ є характеристичною функцією множини значень функції $f(x)$, яка відіграє роль ідеальної вартісної функції. Співвідношення (7) задає строге трактування поняттю «протяжність функції», розбиваючи множину функцій на множину фінітних функцій, для яких E приймає скінченне значення, і множину нефінітних функцій, для яких $E = \infty$. Зокрема, якщо $f(x)$ є функцією відхилу відомих даних від їх моделі, то мінімізація (7) означає мінімізацію протяжності області незбігу даних з їх моделлю.

Наслідуючи (2) і використовуючи (6), уведемо нестроге трактування поняття «протяжність функції» у вигляді

$$E^{(\alpha,\beta,q)}[f(x)] = \int_{-\infty}^{\infty} \psi_S^{(\alpha,\beta,q)} \; [f(x)]dx \quad , \tag{8}$$

яке будемо згадувати далі як «квазіпротяжність функції». Оскільки для певних граничних значень вільних параметрів функція $\psi_S^{(\alpha,\beta,q)}$ перетворюється в χ [10], то функціонал (8) узагальнює (7). З цієї причини строга протяжність завжди є квазіпротяжністю, а не навпаки. Але далі в назвах замість «квазіпротяжність» буде використовуватися більш звичний термін «протяжність».

Функціонал (8) дозволяє отримувати оптимальні рішення для різних видів шуму [10]. Це обумовлено тим, що з (6) за умови $\alpha \to \infty$ й $q = 2$ отримуємо квадратичну вартісну функцію, за умови $\alpha \to \infty$ й $q = 1$ – модульну вартісну функцію, за умови $\beta = 0$ й $q = 2$ – вартісну функцію логарифмічного виду, яка реалізує оптимальну обробку шуму Коші, і т.п. Таким чином, (8) є певним універсальним інструментом, що дозволяє реалізувати різні види оптимальної обробки шляхом його налаштування.

Критерій мінімуму протяжності полягає у вимозі мінімізувати протяжність функції, використовуваної для пошуку рішення. Якщо при цьому брати (7), то цей критерій можна застосувати лише до обмежених за аргументом і вільних від шуму функцій. Якщо ж використовувати (8), то його область застосовності значно розширюється за рахунок включення в неї як зашумленних функцій, так і функцій, які є майже обмеженими за аргументом (наприклад, гауссових функцій). Розширення формулювань критерію на випадок багатокомпонентних даних (наприклад, коли дані є комплекснозначними або являють собою багатоканальне зображення) вимагає подальшого узагальнення поняття «протяжність». Можливі варіанти такого узагальнення полягають у визначенні протяжності багатокомпонентної функції як суми протяжностей її окремих компонент або як її загальнокомпонентної протяжності.

При формулюванні задач можливі, принаймні, два варіанти застосування критерію мінімуму протяжності. У першому варіанті функція, до якої застосовується критерій, розглядається як рішення задачі. Це рішення може складатися з суми відомої й невідомої компоненти або бути невідомою функцією із заданого класу. Тоді критерій мінімуму протяжності формулюється як вимога мінімізувати протяжність шуканого рішення, що далі реалізується або шляхом варіації невідомої компоненти функції, або шляхом варіації власне самої невідомої функції. У другому варіанті функція, до якої застосовується критерій, розглядається як функція відхилу, яка подає результат наближення даних моделлю. Тоді критерій мінімуму протяжності формулюється як вимога мінімізувати протяжність відхилу рішення. Якщо при цьому функція відхилу залежить від певного числа невідомих параметрів, то критерій мінімуму протяжності реалізується шляхом варіації цих невідомих параметрів. Крім зазначених варіантів, є можливим також змішаний варіант, коли висувається вимога одночасно мінімізувати як протяжність шуканого рішення, так і протяжність функції відхилу.

Умови доцільності застосування критерію мінімуму протяжності розглянемо на прикладі задачі наближення кривої заданої форми $f(x) \neq 0$; $x \in X$ з невідомим амплітудним параметром $A \in \boldsymbol{R}$, де $\boldsymbol{R}$ є множиною дійсних чисел, до відомої функції $g(x)$; $x \in X$. Нехай частина значень $g(x)$; $x \in X_1 \subset X$ відома точно, а решта значень спотворена випадковими викидами з випадковими амплітудами, причому елементи множини X_1 апріорно невідомі. Оскільки за відсутності спотворень має виконуватися: $g(x) = A_{\text{true}} f(x)$, де A_{true} є справжнім значенням A, то шукане рішення $A = A_{\text{true}}$ може бути отримане діленням $g(x)$ на $f(x)$ для довільного $x \in X_1$. Але оскільки множина X_1 є апріорно невідомою, то виникає колізія, пов'язана з вибором її елемента з множини X. Саме цю колізію і розв’язує критерій мінімуму протяжності. Так як множина X_1 складається з тих значень x, для яких відношення $g(x) / f(x)$ дорівнює A_{true}, то тоді значення функціоналу строгої протяжності (7), що розглядається як функція параметра A, буде сягати мінімуму, який дорівнює протяжності множини $X \setminus X_1$, що відповідає спотворенням. Для інших значень $A = g(x) / f(x)$, які за умовою є випадковими і не будуть дорівнювати A_{true}, значення функціоналу строгої протяжності буде дорівнювати протяжності всієї множини X. Отже, рішення цієї задачі досягається в точці мінімуму функціоналу строгої протяжності за параметром A, і має сенс постановка задачі

$$\min_A E[g(x) - Af(x)]. \qquad (9)$$

Для дискретної версії задачі (9) метод її рішення полягає в перебиранні всіх значень відношення $g(x)/f(x)$ з метою пошуку того значення, для якого досягається мінімум функціоналу строгої протяжності. Продовжуючи розгляд цієї задачі для випадку, коли на множині X_1 значення функції відомі з малою похибкою через шум, а на множині $X \setminus X_1$ як і раніше випадково спотворені великими викидами, можна сформулювати задачу

$$\min_A E^{(\alpha,\beta,q)}[g(x) - Af(x)], \qquad (10)$$

де значення вільних параметрів α, β й q повинні забезпечувати малу похибку на X_1 й обмежувати велику похибку на $X \setminus X_1$. Необхідне для цього налаштування вільних параметрів можна отримати методом навчання.

Висновки

Критерій мінімуму протяжності полягає у вимозі мінімізувати протяжність множини значень функції, використовуваної для пошуку рішення. Така функція може являти собою як шукане рішення, так і функцію відхилу даних від їх моделі.

Критерій мінімуму протяжності має гнучке застосування через можливість його перетворення в інші традиційні критерії обробки шляхом зміни значень вільних параметрів. Для випадку функції відхилу доцільність застосування критерію мінімуму протяжності виникає за наявності в даних грубих помилок.

Формалізація постановок типових задач дозволить виявити загальні риси і особливості побудови методів обробки даних на основі зазначеного критерію.

ЛІТЕРАТУРА

1. Вовк С.М., Гнатушенко В.В. Обробка даних за наявності шуму і грубих помилок // Вісник Дніпропетровського університету. Серія «Фізика. Радіоелектроніка». – 2017. – Вип. 24(2) . – Т.25. – С. 117 – 144.

2. Huber P., Ronchetti E.M. Robust statistics. 2nd ed.–Hoboken:Wiley,2009.370p.

3. Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г. Регуляризирующие алгоритмы и априорная информация. – М.: Наука, 1983. – 200 с.

4. Titchmarsh E. C. The theory of functions. – New York: Oxford Press, 1939. – 454 p.

5. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. – М.: Мир, 1989. – 512 с.

 ISSN 1562-9945

6. Vovk S.M., Borulko V.F. Minimum duration method for recovery of finite signals // Radioelectronics and Communications Systems. – 1991. – Vol. 34. – P. 67-69.

7. Демиденко Е. З. Оптимизация и регрессия. – М.: Наука, 1989. – 296 с.

8. Aysal T. C. Meridian filtering for robust signal processing / T. C. Aysal, K. E. Barner // IEEE Trans. on Signal Processing. – 2007. – V. 55. – N. 8. – P. 3949–3962.

9. Borulko V.F., Vovk S. M. Minimum-duration filtering // Радіоелектроніка, інформатика, управління.– 2016.– №1. –С.7-14. DOI:10.15588/1607-3274-2016-1-1.

10. Vovk S.M. General approach to building the methods of filtering based on the minimum duration principle // Radioelectronics and Communications Systems. – 2016. – V. 59. – N.7. – P. 281-292. DOI: 10.3103/S0735272716070013.

REFERENCES

1. Vovk S.M., Gnatushenko V.V. Obrobka danih za nayavnosti shumu i grubih pomilok // Visnik Dnipropetrovskogo universitetu. SerIya «FIzika. Radioelek-tronika». – 2017. – Vip. 24(2) . – T.25. – S. 117 – 144.

2. Huber P., Ronchetti E.M. Robust statistics. 2nd ed. – Hoboken: Wiley, 2009. – 370 p.

3. Tihonov A.N., Goncharskiy A.V., Stepanov V.V., Yagola A.G. Regulyariziruyuschie algoritmyi i apriornaya informatsiya. – M.: Nauka, 1983. – 200 s.

4. Titchmarsh E.C. The theory of functions.–New York: Oxford Press,1939.–454 p.

5. Hampel F., Ronchetti E., Rausseu P., Shtael V. Robastnost v statistike. Podhod na osnove funktsiy vliyaniya. – M.: Mir, 1989. – 512 s.

6. Vovk S.M., Borulko V.F. Minimum duration method for recovery of finite signals // Radioelectronics and Communications Systems. – 1991. – Vol. 34. – P. 67-69.

7. Demidenko E.Z. Optimizatsiya i regressiya. – M.: Nauka, 1989. – 296 s.

8. Aysal T. C. Meridian filtering for robust signal processing / T. C. Aysal, K. E. Barner // IEEE Trans. on Signal Processing. – 2007. – V. 55. – N. 8. – P. 3949–3962.

9. Borulko V.F., Vovk S.M. Minimum-duration filtering // Radioelektronika, informatika, upravlinnya.– 2016.– #1. –S.7-14. DOI:10.15588/1607-3274-2016-1-1.

10. Vovk S.M. General approach to building the methods of filtering based on the minimum duration principle // Radioelectronics and Communications Systems. – 2016. – V. 59. – N.7. – P. 281-292. DOI: 10.3103/S0735272716070013.