

Адаптация математического аппарата теории марковских цепей для вероятностного анализа рекуррентного оценивания межкадровых геометрических деформаций изображений

Г.Л. Сафина¹, А.Г. Ташлинский², М.Г. Царёв²

¹Национальный исследовательский Московский государственный строительный университет, Ярославское шоссе 26, Москва, Россия, 129337

²Ульяновский государственный технический университет, Северный Венец 32, Ульяновск, Россия 43202

Аннотация. Работа посвящена анализу возможностей использования Марковских цепей для анализа точности псевдоградиентного релейного оценивания геометрических деформаций изображений. Одним из путей сокращения вычислительных затрат при этом является дискретизации области определения исследуемых параметров. Такой подход позволяет априорно выбрать размерность матрицы переходных вероятностей. Однако такая матрица имеет достаточно сложную структуру, что не позволяет существенно сократить объем вычислений. Предложена матрица переходных вероятностей, размер которой не зависит от размерности вектора оцениваемых параметров. Полученные соотношения задают рекуррентный алгоритм расчета матрицы на итерациях оценивания. Для матрицы одношаговых переходов приведены расчетные выражения с использованием вероятностей сноса оценок параметров деформаций.

1. Введение

При оценивании параметров межкадровых геометрических деформаций изображений (МДИ) $\mathbf{Z}^{(1)}$ и $\mathbf{Z}^{(2)}$ в условиях априорной неопределенности широкое использование [1-4] получили безыдентификационные псевдоградиентные процедуры (ППГ) [5-6] вида:

$$\hat{\mathbf{a}}_t = \hat{\mathbf{a}}_{t-1} - \Lambda_t \boldsymbol{\beta}_t (Q(\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \hat{\mathbf{a}}_{t-1})), \quad (1)$$

где $\mathbf{a} = (\alpha_1, \dots, \alpha_m)^T$ - вектор оцениваемых параметров геометрических деформаций; Λ_t - матрица усиления, $\hat{\mathbf{a}}_0$ - начальное приближение вектора параметров; $\boldsymbol{\beta}_t$ - псевдоградиент целевой функции (ЦФ) $Q(\cdot)$, характеризующей качество оценивания.

Наибольшее практическое применение нашли релейные процедуры [7], когда в (1) $\boldsymbol{\beta}_t = (\text{sign}(dQ(\cdot)/d\alpha_1), \text{sign}((dQ(\cdot)/d\alpha_2)), \dots, \text{sign}((dQ(\cdot)/d\alpha_m)))^T$. При хорошей точности оценок они обладают высоким быстродействием и лучшей по сравнению с градиентными процедурами устойчивостью к помехам импульсного характера [8]. Очередная оценка параметра α_i определяется как:

$$\hat{\alpha}_{i,t} = \hat{\alpha}_{i,t-1} - \lambda_{i,t-1} \text{sign}(d\hat{Q}_i(\mathbf{Z}_1, \mathbf{Z}_2, \hat{\mathbf{a}}_{t-1})/d\alpha_i), \quad \hat{\alpha}_{i0} \in \Omega(\mathbf{a}), \quad (2)$$

где $\Omega(\mathbf{a})$ область определения \mathbf{a} . Последовательность оценок

$$\hat{\mathbf{a}}_0, \hat{\mathbf{a}}_1, \dots, \hat{\mathbf{a}}_t, \dots, \hat{\mathbf{a}}_T, \quad (3)$$

получаемая с помощью ПГП вида (1), является m -мерной последовательностью без последствия и представляет собой векторный Марковский процесс. При этом совместная плотность распределения вероятностей (ПРВ) параметров МДИ на любой итерации может быть выражена через ПРВ $w(\hat{\alpha}_0)$ начального приближения $\hat{\alpha}_0$ и одношаговые условные ПРВ $\pi_t(\hat{\alpha}_t | \hat{\alpha}_{t-1})$ переходов Марковской последовательности из $(t-1)$ -й итерации в t -ю, $t = \overline{1, T}$ [9]:

$$w(\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_T) = w(\hat{\alpha}_0) \prod_{t=1}^T \pi_t(\hat{\alpha}_t | \hat{\alpha}_{t-1}).$$

Аппарат анализа Марковских последовательностей позволяет учесть конечность области $\Omega_{\hat{\alpha}}$ определения возможных значений оценок параметров МДИ при различных правилах поведения ПГП на границах области определения параметров.

Если область $\Omega_{\hat{\alpha}}$ определения $\hat{\alpha}_t$ непрерывна, то последовательность (3) представляет собой простую Марковскую последовательность, если же $\Omega_{\hat{\alpha}}$ дискретна, то (3) является Марковской цепью [9]. Последнее справедливо, в частности, при релейной процедуре (2). Привлечение хорошо разработанного математического аппарата Марковских последовательностей [9, 10] и цепей [9, 11] для анализа эффективности ПГП при конечном числе итераций позволяет получить ряд полезных результатов.

Рассмотрим возможность использования аппарата теории Марковских цепей для моделирования процесса псевдоградиентного оценивания параметров МДИ.

2. Связь матриц переходных вероятностей с вероятностями сноса оценок

Если в (1) матрица усиления Λ_t диагональная, то матрицу условных вероятностей $\mathbf{P}_i(l, t) = \|P\{\hat{\alpha}_{it} = a_{ik} | \hat{\alpha}_{it} = a_{ij}\}\|$ можно выразить через вероятности сноса оценок (ВСО) параметров МДИ [12, 13]. ВСО параметра – это вероятность улучшения оценки после выполнения ПГП итерации (с учетом всех возможных изменений оценок других параметров). Тогда при изменяющемся шаге λ_t приращения в (2) элементы $\pi_{jk}(l, t)$ матрицы $\mathbf{P}(l, t)$:

$$\pi_{jk}(l, t) = \begin{cases} \rho_{\alpha}^{+}(\varepsilon_j), & \text{если } l = t-1, k = j + \Delta_t \text{ sign } \varepsilon_j, \\ \rho_{\alpha}^{o}(\varepsilon_j), & \text{если } l = t, k = j, \\ \rho_{\alpha}^{-}(\varepsilon_j), & \text{если } l = t-1, k = j - \Delta_t \text{ sign } \varepsilon_j, \\ 0, & \text{в другом случае,} \end{cases}$$

где: $\pi_{jk}(l, t)$ - вероятность того, что оценка $\hat{\alpha}_i$ примет значение a_{ik} , если на более ранней итерации $l < t$ оценка имела значение a_{ij} ; $\rho_{\alpha}^{+}(\varepsilon_j)$ - вероятность изменения оценки параметра к α_{it} ; $\rho_{\alpha}^{-}(\varepsilon_j)$ - вероятность изменения оценки от α_{it} ; $\rho_{\alpha}^{o}(\varepsilon_j)$ - вероятность не изменения оценки; $\varepsilon_{ij} = (\alpha_{it} - a_{ij})$; α_{it} - точное значение параметра α_i ; Δ_t - число возможных состояний оценки параметра α в интервале от a_j до a_k , включая состояние a_k . В дальнейшем для упрощения записи при рассмотрении одного параметра индекс « i » будем опускать.

При постоянном шаге λ_t элементы $\pi_{jk}(l, t)$ также непосредственно выражаются через ВСО:

$$\pi_{jk}(l, t) = \begin{cases} \rho_{\alpha}^{+}(\varepsilon_j), & \text{если } l = t-1, k = j + \text{sign } \varepsilon_j; \\ \rho_{\alpha}^{o}(\varepsilon_j), & \text{если } l = t, k = j; \\ \rho_{\alpha}^{-}(\varepsilon_j), & \text{если } l = t-1, k = j - \text{sign } \varepsilon_j; \\ 0, & \text{в другом случае.} \end{cases} \quad (4)$$

При этих условиях получаем однородную цепь Маркова (3), для которой справедливо

$$\mathbf{P}(t) = \mathbf{P}^t, \quad (5)$$

где \mathbf{P} - матрица одношаговых вероятностей переходов. При $t \rightarrow \infty$ такая цепь становится стационарной.

Однако при увеличении числа оцениваемых параметров применение классического математического аппарата Марковских цепей становится проблематичным из-за резкого увеличения размеров матрицы переходных вероятностей.

Одним из основных параметров, определяющих вычислительные затраты, при использовании аппарата Маковских цепей, является число возможных значений оценок параметров МДИ. Априорно выбрать размер матрицы Π позволяет дискретизации области определения оцениваемых параметров. Однако и при этом использование классического аппарата теории Марковских цепей остается целесообразным лишь при оценивании одного параметра, поскольку увеличение числа m оцениваемых параметров на единицу приводит к увеличению вычислительных затрат как минимум в K^2 раз, где K - число возможных дискретных значений оценок $(m+1)$ -го параметра. В задачах измерения параметров МДИ величина K достигает нескольких порядков. При этом нахождение ПРВ оценок параметров, основанное на использовании матрицы одношаговых переходов, становится препятствием для вероятностного моделирования процесса псевдоградиентного измерения параметров межкадровых деформаций.

3. Адаптация аппарата Марковских цепей к решаемой задаче

Для сокращения объема вычислений, воспользуемся тем обстоятельством, что на t -й итерации вне зависимости от состояния оценок остальных параметров, переходы из j -го состояния оценки параметра α_i возможны в только известное k -е состояние, где $k \in \{j + v_{it} + 1, j + v_{it}, j, j - v_{it}, j - v_{it} - 1\}$, $v_{it} = \text{int}(\lambda_{it} / \Delta_{\alpha i})$. При этом вероятности переходов определяются состоянием оценок других параметров. Интегральная вероятность перехода оценки $\hat{\alpha}_i$ из j -го состояния ($\hat{\alpha} = a_{ij}$) в k -е состояние ($\hat{\alpha} = a_{ik}$) определяется суммой вероятностей переходов из подобластей ω_{ik} пространства параметров, $i = \overline{1, m}$, $k = \overline{1, K_i}$. Например, для использования процедур релейного типа при оценивании трех параметров α_1 , α_2 и α_3 общую вероятность $\tilde{\rho}_{ij}$ ухудшения оценки $\hat{\alpha}_1 = a_{1j}$ на t -й итерации можно записать как [14]:

$$\tilde{\rho}_{1j}^- = \sum_{k=1}^{K_1} \left(p_{2k}(t-1) \sum_{l=1}^{K_3} p_{3l}(t-1) \rho^-(\varepsilon_{1j}, \varepsilon_{2k}, \varepsilon_{3l}) \right),$$

где $p_{lk}(t-1) = P(\hat{\alpha}_l = a_{lk} | \hat{\alpha}_l = a_{lj})$ - вероятность того, что на $(t-1)$ -й итерации при $\hat{\alpha}_l = a_{lk}$, $k = \overline{1, K_l}$ величина $\hat{\alpha}_l$ имеет значение a_{1j} ; $\varepsilon_{lk} = a_{lk} - \alpha_{lr}$ - отклонение оценки $\hat{\alpha}_l$ от точного значения α_{lr} , $l = 1, 2, 3$. Соответственно, для m параметров:

$$\tilde{\rho}_{1j}^- = \sum_{k=1}^{K_2} \left(p_{2k}(t-1) \sum_{l=1}^{K_3} \left(p_{3l}(t-1) \dots \sum_{n=1}^{K_m} (p_{mn}(t-1) \rho^-(\varepsilon_{1j}, \varepsilon_{2k}, \dots, \varepsilon_{mn})) \right) \dots \right). \quad (6)$$

Аналогично могут быть записаны и выражения для вероятностей $\tilde{\rho}_{1j}^o$ и $\tilde{\rho}_{1j}^+$.

При таком подходе изменятся и матрицы одношаговых переходов, которые для этого случая обозначим $\tilde{\Pi}_i^{(t)} = \left\| \pi_{jk}^{(t)}(i, \tilde{\rho}_{ij}^*) \right\|$, где i - номер параметра; $\pi_{jk}^{(t)}(i, \tilde{\rho}_{ij}^*)$ - вероятность перехода i -й оценки из состояния a_j на $(t-1)$ -й итерации в состояние a_k на t -й итерации;

$$\tilde{\rho}_{ij}^* = \sum_{v=1}^{K_2} \left(p_{1v}(t-1) \dots \sum_{l=1}^{K_m} (p_{ml}(t-1) \rho_i^*(\varepsilon_{1v}, \dots, \varepsilon_{ij}, \dots, \varepsilon_{ml})) \right), \quad (7)$$

где $\rho_i^*(\cdot)$ - значения вероятностей $\rho_i^+(\cdot)$, $\rho_i^o(\cdot)$ и $\rho_i^-(\cdot)$ при векторе рассогласования оценок $\bar{\varepsilon}_i = (\varepsilon_{1v}, \dots, \varepsilon_{ij}, \dots, \varepsilon_{ml})^T$, $\varepsilon_{ik} = a_{ik} - \alpha_{ir}$. При этом размер матрицы при m параметрах по

сравнению с традиционным подходом сокращается с $\sum_{i=1}^m K_i \times \sum_{i=1}^m K_i$ до $K_i \times K_i$. Примерно настолько же сокращаются и вычислительные затраты. Это сокращение происходит за счет потери информации о вероятности принадлежности оценки $\hat{\mathbf{a}}$ вектора параметров каждой из подобластей пространства параметров. Сохраняются только информация о проекциях распределения в этом пространстве. При этом этой информации достаточно для расчета ПРВ оценок МДИ при конечном числе итераций. Для расчета ПРВ оценки i -го параметра на t -й итерации ППП необходимо знать ПРВ оценок всех параметров на $(t-1)$ -й итерации:

$$\mathbf{p}_i^T(t) = \mathbf{p}_i^T(0) \prod_{s=1}^t \tilde{\mathbf{\Pi}}_i(s). \quad (8)$$

Таким образом, нахождение матрицы $\tilde{\mathbf{\Pi}}_i(t)$ допускает только рекуррентный способ вычисления. Отметим также, что при использовании $\tilde{\mathbf{\Pi}}_i(t)$ даже при $\lambda_i = const$ Марковскую цепь оценок, формируемых ППП, уже нельзя считать однородной. Соответственно выражение (5) для этого случая не справедливо. Для расчета дискретного распределения вероятностей $\mathbf{p}_i^T(t)$ оценок i -го параметра при использовании матрицы $\tilde{\mathbf{\Pi}}_i(t)$ получаем рекуррентную процедуру:

$$\mathbf{p}_i^T(t) = \mathbf{p}_i^T(t-1) \tilde{\mathbf{\Pi}}_i(t), \quad i = \overline{1, m}. \quad (9)$$

При переменном шаге λ_{it} при принятых упрощениях матрицу $\tilde{\mathbf{\Pi}}_i(t)$ для параметра α_i можно найти как:

$$\tilde{\mathbf{\Pi}}_i(t) = \begin{pmatrix} \tilde{\rho}_{i_1}^- + \tilde{\rho}_{i_1}^o & 0 & \dots & \tilde{\rho}_{i_{1-1}}^+ & \tilde{\rho}_{i_{1-2}}^+ & \dots & 0 & 0 \\ \tilde{\rho}_{i_2}^- & \tilde{\rho}_{i_2}^o & \dots & 0 & \tilde{\rho}_{i_{2-1}}^+ & \dots & 0 & 0 \\ \tilde{\rho}_{i_{3-2}}^- & \tilde{\rho}_{i_{3-1}}^- & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \tilde{\rho}_{i_{4-2}}^- & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & \tilde{\rho}_{i_{Ki-1}}^o & \tilde{\rho}_{i_{Ki-1}}^+ \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & \tilde{\rho}_{i_{Ki}}^o + \tilde{\rho}_{i_{Ki}}^+ \end{pmatrix}. \quad (10)$$

При постоянном λ_i выражение, определяющее элемент матрицы, существенно упрощается:

$$\pi_{jk}^{(i)}(t, \tilde{\rho}_{i_j}^*) = \begin{cases} \tilde{\rho}_{i_j}^- + \tilde{\rho}_{i_j}^o, & \text{если } j = k = 1, \\ \tilde{\rho}_{i_j}^-, & \text{если } j = k - 1, 1 < j < K_i, \\ \tilde{\rho}_{i_j}^o, & \text{если } j = k, 1 < j < K_i, \\ \tilde{\rho}_{i_j}^+, & \text{если } j = k + 1, 1 < j < K_i, \\ \tilde{\rho}_{i_j}^o + \tilde{\rho}_{i_j}^+, & \text{если } j = k = K_i, \\ 0 & \text{в другом случае.} \end{cases} \quad (11)$$

Соотношения (7)-(11) определяют рекуррентные алгоритмы расчета матрицы $\tilde{\mathbf{\Pi}}_i(t)$ и ПРВ погрешностей оценок параметров для требуемой итерации оценивания, начиная с начального приближения. При этом размер $\tilde{\mathbf{\Pi}}_i(t)$ не зависит от размерности вектора $\bar{\alpha}$ и определяется только параметрами дискретизации области определения конкретного параметра α_i . Вычислительные затраты при увеличении размерности вектора параметров растут пропорционально $\sum_{i=1}^m K_i$, что позволяет найти компромисс между точностью расчета ПРВ и требованиями к вычислительным ресурсам. Дальнейшее сокращение вычислительных затрат возможно за счет наложения ограничений на область допустимых значений оцениваемых параметров и введения правил учета вероятностей за границами этой области.

4. Заключение

Показано, что последовательность оценок параметров МДИ, получаемая с помощью ПГП, является последовательностью без последствия и представляет собой векторный Марковский процесс. При одном оцениваемом параметре для вероятностного моделирования процесса псевдоградиентного оценивания целесообразно использование математического аппарата теории Марковских цепей. Исследование выражений, позволяющих рассчитывать переходные вероятности матрицы одношаговых переходов через вероятности сноса оценок параметров, показало, что для релейных алгоритмов при одном оцениваемом параметре матрица одношаговых переходов имеет пятидиагональную структуру и при постоянном шаге изменения оценок не зависит от номера итерации, обуславливая однородность Марковской цепи. Однако для вектора параметров использование классического аппарата теории Марковских цепей становится проблематичным из-за резкого увеличения размера матрицы переходных вероятностей. Априорно выбрать размер матрицы позволяет дискретизации области определения оцениваемых параметров, но и при таком подходе использование классического аппарата оказывается целесообразным лишь при оценивании одного параметра.

С целью сокращения вычислительных затрат предложены матрицы одношаговых переходов, размерность которых определяется только дискретизацией области определения соответствующих им параметров и не зависит от их числа. Полученная матрица допускает только рекуррентный способ ее вычисления от начального приближения параметров до требуемой итерации. При этом Марковская цепь оценок теряет свойство однородности. Также теряется точная информация о распределении вероятностей в пространстве параметров. Сохраняются только проекции этого пространственного распределения. Однако этого достаточно для решения задачи нахождения ПРВ оценок параметров межкадровых деформаций при конечном числе итераций.

5. Литература

- [1] Shalev-Shwartz, S. Accelerated proximal stochastic dual coordinate ascent for regularized loss minimization / S. Shalev-Shwartz, T. Zhang // *Mathematical Programming*. – 2016. – Vol. 155(1). – P. 105-145.
- [2] Moritz, P. A linearly-convergent stochastic L-BFGS algorithm / P. Moritz, R. Nishihara, M. Jordan // *Proc. of the 19th Int. Conf. on Artificial Intelligence and Statistics*. – 2016. – P. 249-258.
- [3] Смирнов, П.В. Методика выделения области подвижного объекта на последовательности изображений / П.В. Смирнов, А.Г. Ташлинский // *Радиотехника*. – 2015. – Т. 6. – С. 5-11.
- [4] Borisova, I.V. Application of the gradient orientation for systems of automatic target detection / I.V. Borizova, V.N. Legkiy, S.A. Kravets // *Computer Optics*. – 2017. – Vol. 41(6). – P. 931-937. DOI: 10.18287/2412-6179-2017-41-6-931-937.
- [5] Tashlinskii, A.G. Computational Expenditure Reduction in Pseudo-Gradient Image Parameter Estimation / A.G. Tashlinskii // *Lecture Notes in Computer Science*. – 2003. – Vol. 2658. – P. 456-462. DOI: 10.1007/3-540-44862-4_48.
- [6] Ташлинский, А.Г. Оценивание параметров пространственных деформаций последовательностей изображений / А.Г. Ташлинский. – Ульяновск: УлГТУ, 2000. – 131 с.
- [7] Цыпкин, Я.З. Информационная теория идентификации / Я.З. Цыпкин. – М.: Наука. Физматлит, 1995. – 336 с.
- [8] Tashlinskii, A.G. Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations / A.G. Tashlinskii // *Vision Systems: Segmentation & Pattern Recognition* / G. Obinata, A. Dutta, editors. – Austria, Vienna: I-Tech, 2007. – P. 465-494.
- [9] Тихонов, В.И. Марковские процессы / В.И. Тихонов, М.А. Миронов. – М.: Советское Радио, 1977. – 488 с.
- [10] Невет, Ж. Математические основы теории вероятностей. – М.: Мир, 1969. – 309 с.
- [11] Дынкин, Е.Б. Основания теории марковских процессов. – М.: Физматгиз, 1959. – 227 с.
- [12] Ташлинский, А.Г. Методика анализа погрешности псевдоградиентного измерения параметров многомерных процессов / А.Г. Ташлинский, В.О. Тихонов // *Известия вузов: Радиоэлектроника*. – 2001. – Т. 44, № 9. – С. 75-80.

- [13] Ташлинский, А.Г. Вероятность сноса оценок параметров межкадровых геометрических деформаций изображений при псевдоградиентном измерении / А.Г. Ташлинский, И.В. Воронов // Известия Самарского научного центра Российской академии наук. – 2014. – Т. 16, № 6(2). – С. 612-615.
- [14] Ташлинский, А.Г. Исследование возможностей сокращения вычислительных затрат при моделировании псевдоградиентных алгоритмов / А.Г. Ташлинский, В.О. Тихонов // Современные проблемы создания и эксплуатации радиотехнических систем: Труды 3-й всерос. науч.-практич. конференции. – Ульяновск, 2001. – С. 77-78.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ и Правительства Ульяновской области в рамках научного проекта № 18-41-730006.

Adaptation of the mathematical apparatus of the Markov chain theory for the probabilistic analysis of recurrent estimation of image inter-frame geometric deformations

G.L. Safina¹, A.G. Tashlinskii², M.G. Tsaryov²

¹National Research Moscow State University of Civil Engineering, Yaroslavskoe Shosse 26, Moscow, Russia, 129337

²Ulyanovsk State Technical University, Severnii Venetz 32, Ulyanovsk, Russia, 432027

Abstract. The paper is devoted to the analysis of the possibilities of using Markov chains for analyzing the accuracy of stochastic gradient relay estimation of image geometric deformations. One of the ways to reduce computational costs is to discretize the domain of studied parameters. This approach allows to choose the dimension of transition probabilities matrix a priori. However, such a matrix has a rather complicated structure. It does not significantly reduce the amount of computations. A modification of the transition probabilities matrix is proposed, it's dimension does not depend on the dimension of estimated parameters vector. In this case, the obtained relations determine a recurrent algorithm for calculating the matrix at the estimation iterations. For the one-step transitions matrix, the calculated expressions for the probabilities of image deformation parameters estimates drift are given.