ПРОГНОЗ РЕЗУЛЬТАТИВНОСТИ СПОРТСМЕНОВ НА БАЗЕ СТАТИСТИЧЕСКОГО ФАКТОРНОГО АНАЛИЗА И ЭКСПЕРТНОГО РАНЖИРОВАНИЯ ПОЛНОЙ СОВОКУПНОСТИ АНТРОПОМЕТРИЧЕСКИХ, ТЕХНИЧЕСКИХ И СПЕЦИАЛИЗИРОВАННЫХ ПАРАМЕТРОВ

Ахметов Р.Ф.

Житомирский государственный университет имени Ивана Франка

Аннотация. Ставится и решается основная задача прогноза результативности прыгунов в высоту на базе статистического факторного анализа и экспертного ранжирования полной совокупности антропометрических, технических и специализированных параметров.

Ключевые слова: линейная регрессия, аппроксимация, весовой вектор.

Анотація. Ахметов Р.Ф. Прогноз результативності спортсменів на базі статистичного факторного аналізу й експертного ранжування повної сукупності антропометричних, технічних і спеціалізованих параметрів. Ставиться та вирішується основне завдання прогнозу результативності стрибунів у висоту на базі статистичного факторного аналізу й експертного ранжування повної сукупності антропометричних, технічних і спеціалізованих параметрів.

Ключові слова: лінійна регресія, апроксимація, ваговий вектор.

Annotation. Akhmetov R.F. Prognosticating athletes' performance on the basis of statistical and factor analysis as well as expert ranging of the comprehensive unity of anthropometric, technical and specialized parameters. The paper provides ways to solve the major problem of prognosticating the performance of high-jumping based upon statistic and factor analysis as well as expert ranging of the comprehensive unity of anthropometric, technical and specialized parameters.

Key words: linear regression, approximation, weight vector.

Постановка проблемы. В системе спортивной многолетней подготовки большая роль отводится прогнозу результативности как отдельных спортсменов, так и спортивных групп [2; 5]. В связи с этим весьма актуальным является разработка программы прогноза на базе некоторой совокупности информативных параметров спортсменов.

Анализ последних исследований и публикаций. Прогнозирование основывается на использовании метода экстраполяции, предполагающего распространение выводов, полученных из наблюдения над одной частью какого-либо явления, на другие его части [2; 6; 7]. В условиях спорта экстраполяция позволяет осуществить прогнозы роста результативности на основе изучения соответствующих закономерностей в предшествующие годы. Задачу прогноза результативности спортсменов можно решить на базе факторного анализа и динамики развития физических параметров и результатов на некотором ограниченном интервале времени [1; 6]. Особое внимание уделяется «раннему» прогнозу на период до 17 лет по данным начального периода (10-13 лет).

Целью настоящего исследования была разработка программы прогнозирования результативности мастеров спорта международного класса по прыжкам в высоту по данным их специальных физических параметров и полученной функции линейной регрессии с учетом ее дисперсии.

Результаты исследования. Поскольку результаты и физические параметры спортсменов в группе имеют случайный разброс (дисперсию), то, говоря о задаче прогноза результативности, имеет смысл рассматривать прогноз средней результативности $\overline{\mathrm{H}}(t)$, как функции средних по группе физических параметров $\vec{X}_P$, которые будем представлять в виде матрицы столбца:

$$\vec{\mathrm{X}}_P = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_P \end{pmatrix}, \ \mathrm{P}=1,2,\dots,\mathrm{N}-1;\ \mathrm{N}\geq 2,$$

где N – полное число спортивных параметров, включая сам результат (H). Полное множество P-мерных группировок из (N-1) по P равно числу сочетаний из (N-1) по P :

$$\vec{\mathrm{X}}_P \in U_{\vec{X}_P} = \{\vec{X}_P^{\alpha},\ \alpha = 1,2,\ldots,C_{N-1}^{P}\}, \tag{1}$$

$$C_{N-1}^{P} = \frac{(N-1)!}{P!(N-1-P)!}.$$

Информативность различных P-мерных группировок $\vec{\mathrm{X}}_P$ в задачах прогноза результативности будет также различной. Вопрос о выборе оптимальной совокупности наиболее информативных параметров из множества (1) при различных P требует самостоятельных глубоких исследований в рамках отдельной НИР. В данной работе предлагается один из альтернативных вариантов решения задачи, который вполне приемлем с точки зрения точности прогноза. В первом приближении рассматривается задача линейного прогноза в рамках классической теории линейной регрессии (интерполяции) в математической статистике [4; 7-9]. Речь идет о нахождении аппроксимации

$$\overline{\mathrm{H}} \cong \mathrm{H}_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_P X_P, \tag{2}$$

где $\mathrm{H}_0, \alpha_1, \ldots, \alpha_P$ – неизвестные параметры регрессии, которые требуется оценить по данным некоторого количества возрастных групп. В более точной постановке приближенная линейная регрессия (2) представляется в виде:

$$\overline{\mathrm{H}}(t) = \mathrm{H}_0 + \alpha_1 X_1(t) + \alpha_2 X_2(t) + \ldots + \alpha_P X_P(t) + \xi(t),\ t \in T = (a,b), \tag{3}$$

где $\xi(t)$ – ошибка прогноза с нулевым средним ($M\xi(t) = 0$) и неизвестной дисперсией $\sigma_\xi^2 = M\xi^2$ (M – оператор математического ожидания – среднего). Если в результате решения задачи линейной регрессии на интервале времени T получены оценки неизвестных параметров регрессии:

$$\mathrm{H}_0 = \mathrm{H}_0^{\wedge}(T);\ \alpha_n = \alpha_n^{\wedge}(T), \quad n = 1,2,\ldots,P,$$

то прогнозное значение средней результативности вне этого интервала представляется в виде:

$$\overline{\mathrm{H}}^{\wedge}(t_0) = \mathrm{H}_0^{\wedge}(T) + \sum_{n=1}^{P} \alpha_n^{\wedge}(T) X_n(t_0), \quad t_0 > b, \tag{4}$$

где набор физических параметров $\{ X_n(t_0),\quad n = 1,2,..., P \}$ – задается на прогнозируемый момент времени t_0. При этом среднеквадратическая ошибка (СКО) прогноза оценивается величиной $\sigma_\xi(T)$. Насколько «удачно» получена оценка (4), – зависит от многих факторов и последнее слово здесь за практикой (экспериментальной апробации). Проведенная в данной работе апробация модели (4) показывает, что она практически вполне приемлема. СКО при этом не превышает 3-х сантиметров, а прогнозируемый рекордный результат составляет 250 см.

Матричное решение задачи линейной регрессии результативности по заданной совокупности наиболее информативных параметров

Для оценки параметров регрессии $\mathrm{H}_0, \alpha_1, ..., \alpha_P$ составляется следующая система линейных алгебраических уравнений:

$$\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_1) = \overline{\mathrm{H}}(t_1)$$

$$\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_2) = \overline{\mathrm{H}}(t_2) \tag{5}$$

$$\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_N) = \overline{\mathrm{H}}(t_N),$$

где в данном разделе, следуя стандартным обозначениям, N – число возрастных групп (в данной работе N<9). Система (5) представляется в матричном виде:

$$\mathrm{H}_0 \vec{1}_N + \sum_{m=1}^{P} \alpha_m \vec{X}_N^m = \vec{\overline{\mathrm{H}}}_N \Rightarrow \tag{6}$$

$$\vec{1}_N = \begin{pmatrix} 1 \\ 1 \\ .. \\ 1 \end{pmatrix}_N, \quad \vec{X}_N{}^m = \begin{pmatrix} X_m(t_1) \\ X_m(t_2) \\ \dots\dots\dots \\ X_m(t_N) \end{pmatrix}, \quad \vec{\overline{\mathrm{H}}}_N = \begin{pmatrix} \overline{\mathrm{H}}(t_1) \\ \overline{\mathrm{H}}(t_2) \\ \dots\dots\dots \\ \overline{\mathrm{H}}(t_N) \end{pmatrix}.$$

Вводя т.н. «сигнальный» регрессионный вектор (СРВ):

$$\vec{s}_M = \begin{pmatrix} \mathrm{H}_0 \\ \alpha_1 \\ \dots \\ \alpha_P \end{pmatrix}_M = \begin{pmatrix} s_1 \\ s_2 \\ \dots \\ s_M \end{pmatrix}, \quad M = P + 1, \tag{7}$$

$$s_1 = \mathrm{H}_0\ , s_2 = \alpha_1, s_3 = \alpha_2, \dots, s_M = \alpha_P,$$

матричную систему (6) представляем также в стандартном виде:

$$\sum_{m=1}^{M} s_m \vec{Y}_N^m = \vec{\bar{\mathrm{H}}}_N \quad \Rightarrow \quad Y_{NM}\vec{s}_M = \vec{\bar{\mathrm{H}}}_N, \tag{8}$$

$$\vec{Y}_N^1 = \vec{1}_N, \vec{Y}_N^2 = \vec{X}_N^1, \dots, \quad \vec{Y}_N^M = \vec{X}_N^P, \quad Y_{NM} = (\vec{Y}_N^1 \vec{Y}_N^2 .. \vec{Y}_N^P),$$

где Y_{NM} – измеримая матрица наблюдений (ИМН); $\vec{\bar{\mathrm{H}}}_N$ – измеримый вектор средних результатов (ВСР).

Согласно общей теории линейной регрессии система (8) может быть решена, если она полностью определена или переопределена:

$$N \geq M + 1 = P + 2 \Rightarrow \quad Rank Y_{NM} = M\ . \tag{9}$$

Отметим, что величина (M+1) обусловлена тем, что в число неизвестных помимо M=P+1 неизвестных параметров регрессии необходимо включить также и неизвестное СКО σ_ξ. При выполнения условия (9) статистическое решение задачи линейной регрессии представляется в виде:

$$\vec{s}_M^{\wedge} = Y_{NM}^{-} \vec{\bar{\mathrm{H}}}_N\ , Y_{NM}^{-} = (Y_{NM}^T Y_{NM})^{-1} Y_{NM}^T\ , \tag{10}$$

$$(\sigma_\xi^2)^{\wedge} = \frac{1}{N-M} // \vec{\bar{\mathrm{H}}}_N^{\wedge} - \vec{\bar{\mathrm{H}}} //^2 = \frac{// \Lambda_{NN}^{M\perp} \vec{\bar{\mathrm{H}}}_N //^2}{N-M}\ , \tag{11}$$

$$\vec{\bar{\mathrm{H}}}_N^{\wedge} = Y_{NM}\vec{s}_M^{\wedge} = \Lambda_{NN}^M, \quad \Lambda_{NN}^M = Y_{NM} Y_{NM}^{-}, \quad \Lambda_{NN}^{M\perp} = \mathrm{I}_{NN} - \Lambda_{NN}^M,$$

$$Rank \Lambda_{NN}^M = M, \quad Rank \Lambda_{NN}^{M\perp} = N - M,$$

где Y_{NM}^{-} – псевдообратная матрица [5]; Λ_{NN}^M – проектор в линейную оболочку из базисных векторов $\{\vec{Y}_N^m, m = 1,2,\dots,M\}$; $\Lambda_{NN}^{M\perp}$ – ортогональный вектор.

В данной работе наиболее точное решение получено в случае P=3 при различных N с учетом необходимого условия разрешения (9):

$$5 \leq N \leq 8\ . \tag{12}$$

Разработана специализированная программа corrS1m.com в среде Turbo Pascal. Результаты расчетов приведены в приложении. Специфической математической особенностью задачи регрессии спортивного результата является то, что в силу довольно однородного состава групп столбцовые вектора ИМН Y_{NM} оказываются хотя и случайными, но с малым угловым расхождением относительно «единичного» вектора $\vec{1}_N$. Последнее обстоятельство требует жесткого контроля точности обращения матрицы Грама $(Y^T_{NM}Y_{NM})_{MM}$, т.к. в случае высокой угловой корреляции («схожести») векторов $\vec{Y}^m_N$ матрица Грама оказывается часто плохо обусловленной [3] с большим динамическим диапазоном собственных чисел в области малых величин. При этом точность обращения матрицы Грама с ростом размерности P>3 (числа учитываемых информативных параметров) начинает резко падать и дальнейшее увеличение размерности P не представляется возможным.

Отметим также, что в данной работе максимальное число возрастных групп матеріальної крапки 8.

Поэтому в силу условия (9) предельное число наиболее информативных параметров ограничивается величиной 6:

$$P \le N - 2 \le N_{\max} - 2 = 8 - 2 = 6. \qquad (13)$$

ЭКСПЕРТНОЕ РАНЖИРОВАНИЕ И ВЫБОР НАИБОЛЕЕ ИНФОРМАТИВНОЙ СОВОКУПНОСТИ ФИЗИЧЕСКИХ ПАРАМЕТРОВ СПОРТСМЕНОВ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПРОГНОЗА РЕЗУЛЬТАТИВНОСТИ

Как уже отмечалось, размерность вектора информативных физических параметров спортсменов $\vec{X}_P$ принципиально ограничивается числом возрастных групп и в данной работе оказывается не более шести (13). Прямой перебор различных возможных комбинаций информативных параметров из числа сочетаний (1) оказывается достаточно большим и весьма трудоемким. Так для p=3 и N=8 требуется перебрать

$$C^3_5 + C^3_6 + C^3_7 + C^3_8 = 10 + 15 + 21 + 15 = 102 \qquad (14)$$

комбинации 3-х мерных совокупностей $\vec{X}_p$.

Непосредственное использование результатов факторного анализа также наталкивается на трудности, обусловленные сменой первых наиболее информативных комбинаций параметров от одной возрастной группы к другой. В связи с этим представляет интерес альтернативные методы выделения наиболее информативной совокупности физических параметров для задачи прогноза средней результативности. В данной работе впервые предпринимается попытка совместить хорошо известные и признанные методы современного факторного анализа и пока еще мало известной в педагогических кругах методах фундаментальной теории т.н. доминантных иерархических систем, развиваемой крупным американским ученым Т.Л. Саати [8]. Теория Саати Т.Л., которая получила мировое признание, освещена уже во многих монографиях в области экспертного оценивания и в последние 10 лет находит приложение в самых различных областях науки и техники. Поэтому в данной работе общие положения теории Саати Т.Л. опускаются. Основное внимание уделяется вопросам непосредственного приложения этой теории к задаче экспертного оценивания и ранжирования полной 20-ти мерной совокуности атропометрических, технических и специализированных параметров спортсменов. Отметим, что существенное отличие и оригинальность теории Саати Т.Л. заключается в том, что она позволяет проводить достаточно корректно научно-обоснованные экспертные оценки не только количественных параметров (к которым относятся все измеримые физические параметры), но также и нередко используемые в спортивно-педагогической работе т.н. лингвистические показатели качества (ЛПК), которые можно описать только словесно без обращения к традиционным количественным показателям, – например, к ЛПК относится такой показатель, как «способность выслушивать спортсменом от тренера критические замечания и мобилизовать соответствующие физические резервы».

Одним из центральных положений теории доминантных иерархических систем Саати является введение 9-ти бальной шкалы степени важности (приоритетности) параметров при их попарном сравнении (таблица С).

Таблица С

Степень важности	**Определение**	**Объяснение**
1	Одинаковая важность	Два действия носят одинаковый вклад в достижение цели
3	Некоторое преобладание значимости одного над другим (слабая значимость)	Существуют соображения в пользу предпочтения одного. Однако, эти соображения недостаточно убедительны
5	Существенная значимость или сильная значимость	Имеются надежные данные и логические соображения для того, чтобы показать предпочтение одного перед другим
7	Очевидная значимость	Убедительное свидетельство в пользу предпочтения одного перед другим
9	Абсолютная значимость	Свидетельства в пользу предпочтения одного другому в высшей степени убедительны
2.4.6.8	Промежуточные значения между двумя соседними суждениями	Ситуации, когда необходимо компромисное решение
Обратные величины приведенных выше чисел	Если действию **i** при сравнении с действием **j** приписывается одно из определенных выше ненулевых чисел, то действию **j** при сравнении с действием **i** приписывается обратное значение.	Если согласованность была постулирована при получении n числовых значений для образования матрицы попарных сравнений
Рациональные значения	Отношения, возникающие для заданной шкалы	То же

Используя указанную таблицу Саати, в данной работе была сформирована т.н. экспертная матрица приоритетности (ЭМПР) – квадратная несимметричная матрица попарных сравнений Саати Т.Л. размером N×N (N=20) с положительными элементами $\mathbf{A_{ij}}$ и с обратной симметрией (таблица А).

Таблица А

i/j	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1	5	5	9	9	7	1/5	1/5	1/3	1/3
2	9	1	1	5	7	1/3	1/3	1/5	1/3	1/5
3	1	1/5	1	7	7	1/3	1/3	1/5	1/3	1/5
4	5	1	1/5	1	1	1/3	1/5	1/7	1/5	1/7
5	7	1	5	3	1	1/5	1/5	1/7	1/5	1/7
6	3	1/5	5	1/3	1/5	1	1/5	1/7	1/3	1/5

7	1	1/3	1	1	1/5	1/5	1	1/3	3	5
8	5	1	1	3	1/3	3	5	1	5	3
9	3	7	5	5	5	1	5	1	1	3
10	3	7	7	5	7	5	7	7	7	1
i/j	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	1/3	1/3	1/5	1/3	1/3	1/3	1/3	1/3	1/3	1/5
2	1/5	1/5	1/7	1/5	1/5	1/3	1/3	1/3	1/5	1/7
3	1/5	1/5	1/7	1/5	1/5	1/3	1/3	1/3	1/5	1/7
4	1/7	1/7	1/7	1/5	1/5	1/3	1/3	1/3	1/3	1/5
5	1/7	1/7	1/7	1/5	1/5	1/3	1/3	1/3	1/3	1/5
6	1/7	1/7	1/7	1/5	1/5	1/3	1/3	1/3	1/3	1/5
7	1/5	1/5	1/3	5	5	7	7	7	7	1
8	1/3	1/3	1/5	7	7	9	9	9	5	3
9	1/5	1/5	1/5	1	1	3	3	3	3	1/5
10	1/5	1/3	1/5	1	1	3	3	3	3	1/5

i/j	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
i/j	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
11	1	7	5	9	9	7	5	5	5	1
12		1	1/7	3	3	5	5	5	5	1
13			1	5	5	7	7	7	5	3
14				1	3	5	5	5	3	1
15					1	5	5	5	3	1/3
16						1	3	3	1/3	1/5
17							1	1	1/3	1/5
18								1	1/3	1/5
19									1	1/7
20										1

$$\mathbf{A=(A_{ij}),\ i,j=1,2,\dots,N;\ A_{ji}=1/A_{ij}\,;}$$

$$\mathbf{A_{ii}=1,\ A_{ij}\in\{1,2,3,4,5,6,7,8,9,1/2,1/3,1/4,1/5,1,6,1/7,1/8,1/9\},}$$

где $\mathbf{A_{ij}}$ – приоритет параметра $\mathbf{A_i}$ перед $\mathbf{A_j}$ – показывает на сколько параметр $\mathbf{A_i}$ важнее (более приоритетно (при $\mathbf{A_{ij}>1}$) или менее приоритетно (при $\mathbf{A_{ij}<1}$), чем парметр $\mathbf{A_j}$.

Отметим, что согласно психологическим исследованиям один эксперт в состоянии объективно сравнивать и ранжировать одновременно не более 5-ти параметров. В данной работе мы сравниваем 20 физических параметров (!), но сравнение проводим попарно (по методике Саати Т.Л.).

После формирования ЭМПР требуется решить задачу определения весов или количественной меры степени важности каждого из 20 параметров. Отметим, что обычно делают эвристическое взвешивание параметров весьма субъективно и ориентировочно без какого-то математического анализа и соответствующего обоснования. Так, в данном случае был дан следующий эвристический весовой вектор физических параметров (номера в расширенном списке 2-21):

$$\mathbf{P_Э^T=(P_1,P_2,\ldots,P_{20}),\ \sum P_n=1,\ n=1,2,\ldots,20} \qquad (1)$$

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\mathbf{P_n}$	0.04	0.01	0.01	0.01	0.01	0.02	0.10	0.10	0.10	0.04
n	11	12	13	14	15	16	17	18	19	20
$\mathbf{P_n}$	0.20	0.06	0.10	0.02	0.03	0.02	0.01	0.01	0.03	0.08

Согласно же теории иерархических систем Саати Т.Л. задача оптимального взвешивания (ранжирования) сводится к алгебраической спектральной задаче для ЭМПР **A**, т.е. к нахождению собственных значений и собственных векторов матрицы **A**:

$$\mathbf{AH=\lambda H \Rightarrow \lambda=\lambda_m,\ m=1,\ldots,N,\ \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N,\ H=H_1,H_2,\ldots,H_N},$$

где $\mathbf{\{\lambda_m,H_m\}}$ – совокупность собственных значений и собственных векторов матрицы **A.** Оптимальный весовой вектор – это нормированный первый собственный вектор, соответствующий максимальному собственному значению $\mathbf{\lambda_{max}=\lambda_1}$:

$$\mathbf{P_n^{opt}=(H_1)_т/\Sigma\ ,\ \Sigma=\sum (H_1)_n,\ n=1,2,\ldots,N}\,.$$

Можно показать, что такая довольно нетривиальная процедура формирования весового вектора совершенно не противоречит естественной эмпирической оценке, по крайней мере, в случае, когда все параметры равнозначны. Тогда, очевидно, эмпирическая оценка весового вектора представляется в виде равномерного распределения $\mathbf{P_n=1/N}$. Оказывается, что спектральный анализ ЭМПР с $\mathbf{A_{ij}=1}$ также дает равномерное распределение $\mathbf{P_n^{opt}=1/N}$. Численный спектральный анализ ЭМПР **A** (таблица А) на ПЭВМ с математическим обеспечением типа **MatLab** дает следующий оптимальный весовой вектор:

Ранжирование 20 параметров по Саати (номера из 1-21)

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
12	14	9	13	21	8	15	16	11	10	20	2	17	18	19	7	3	4	5	6
17.4	14.7	10.8	9.2	7.9	7.8	4.6	3.9	3.7	3.6	2.7	2.6	2.1	1.8	1.8	1.3	1.3	1.3	0.8	0.8

(2 строка – номер параметра; 3 строка – вес Саати в %; степень доверия – 75,3%)

РАСШИРЕННЫЙ ПЕРЕЧЕНЬ-21 ПАРАМЕТРОВ СПОРТСМЕНОВ

1. Спортивный результат (высота) – Целевая функция.

АНТРОПОМЕТРИЧЕСКИЕ ПАРАМЕТРЫ (2-7)

2. Рост.
3. Длина голени.
4. Длина бедра.
5. Окружность бедра.
6. Окружность икроножной мышцы.
7. Вес.

ТЕХНИЧЕСКИЕ ПАРАМЕТРЫ (8-14)

(Регистрируемые и расчетные показатели техподготовки)

8. Скорость разбега перед отталкиванием.
9. Скорость вылета ОЦТ (в момент отрыва).
10. Угол вылета ОЦТ.
11. Длительность фазы отталкивания.
12. Высота вылета ОЦТ.
13. Импульс силы отталкивания.
14. Степень использования силовых возможностей толчка (%).

СПЕЦИАЛИЗИРОВАННЫЕ ПАРАМЕТРЫ (15-21)

(Уровень спецфизподготовки)

15. Бег – 30 м. (с).
16. Скорость спринтерского бега (10 м. с хода).

17. Прыжок вверх в высоту с двух ног с места.
18. Прыжок в длину с места.
19. Тройной прыжок с места
20. Прыжок вверх с толчковой ноги (махом другой).
21. Прыжок вверх в высоту с трех шагов.

$$\mathbf{P_{opt}^{T} = (P_1^{opt}, P_2^{opt}, \ldots, P_{20}^{opt}),} \qquad (2)$$

n	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\mathbf{P_n^{opt}}$	0.026	0.013	0.013	0.008	0.008	0.013	0.078	0.108	0.036	0.037
n	11	12	13	14	15	16	17	18	19	20
$\mathbf{P_n^{opt}}$	0.174	0.092	0.147	0.046	0.039	0.021	0.018	0.018	0.027	0.079

$$\boldsymbol{\lambda_{max}=12{,}658,\ \gamma=1-(\lambda_{max}-N)/N=75.3\%,}$$

где параметр **γ** характеризует степень доверия к экспертам (чем ближе к 100%, тем больше степень доверия; практически приемлемая степень доверия часто полагается более 75%). Сравнение оптимального вектора (2) с эмпирическим (1) показывает их существенное различие, что еще раз подчеркивает отмеченный ранее вывод психологических исследований о недостаточной обоснованности и эффективности эмпирических оценок в случае **N>5**.

Для последующего регрессионного анализа были выбраны следующие 4 параметра: X_{12}, X_9, X_{21}, X_8 .

Апробация алгоритмов прогноза результативности прыгунов в высоту по различному числу информативных параметров

Программа РЕГРЕССИЯ (corrS1m.com) содержит следующие разделы:

1. Вызов исходных статистических данных (файл g1_21_9.dat)

2. Шифр файла: tN-M, где N – число возрастных групп, по которым проводится прогноз на будущее; M – число информативных параметров (N≥M+2)

3. Выбор M информативных параметров (из номеров 2-21).

4. Анализ ранга регрессионной матрицы $Y_{N(M+1)}$ методом Грама-Шмидта.

5. Анализ корреляции информативных параметров по годам.

6. Спектральный анализ матрицы Грама Y^TY размером (M+1)*(M+1).

7. Оценка точности обращения матрицы Грама.

8. Оценка статистических характеристик информативных параметров (средние, СКО, корреляционная матрица).

9. Решение задачи линейной регрессии.

10. Оценка дисперсии шума (СКО=s).

11. Прогнозирование за пределы выбранных возрастных групп, включая прогноз рекордных результатов (приводятся соответствующие графики).

Результаты расчетов приводятся в приложении.

Выводы

Задача прогноза результативности спортсменов является задачей интерполяции средней (по возрастной группе) результативности ($\overline{\text{H}}$) в виде линейной комбинации средних значений наиболее информативных физических параметров спортсменов ($\bar{x}_1, \bar{x}_2, ..., \bar{x}_P$) с указанием точности (СКО) прогноза:

$$\overline{\text{H}} = \text{H}_0 + \alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 + \ldots + \alpha_P \bar{x}_P + \xi, \quad \overline{\xi^2} = \sigma_\xi^2,$$

где $\text{H}_0, \alpha_1, \alpha_2, ..., \alpha_P$ – параметры регрессии; σ_ξ – СКО прогноза.

В условиях априорной неопределенности об СКО прогноза необходимым условием решения задачи прогноза является превышение числа используемых возрастных групп ($N_{ВГ}$) над числом используемых информативных физических параметров (P), как минимум на две единицы:

$$N_{ВГ} \geq P+2 .$$

Так, при числе информативных физических параметров P=3 требуются средние значения по 5-ти возрастным группам (10, 11, 12, 13 и 14 лет). При этом можно дать прогноз результативности не только на любой «внутренний» момент времени t_0 ($10 \leq t_0 \leq 13$), но и на будущие моменты времени t_0>13, включая прогноз рекордных результатов. Для этого достаточно в полученную формулу регрессии подставить значения прогнозных средних значений физических параметров $\{ \bar{x}_n(t_0), n = 1,2,..., P \}$:

$$\overline{\text{H}}(t_0) \cong \text{H}_0 + \alpha_1 \bar{x}_1(t_0) + \alpha_2 \bar{x}_2(t_0) + \ldots + \alpha_P \bar{x}_P(t_0) \quad (\pm\sigma_\xi).$$

В частности, при прогнозе по трем параметрам (x_{12},x_9,x_{21}) по 5-ти возрастным группам (10, 11, 12, 13 и 14 лет) получена следующая регрессионная функция:

$$\hat{H} = 0.514 + 0.715 x_{12} + 0.031 x_9 + 0.819 x_{21}\,, \quad \hat{\sigma} = s = 0.9\tilde{n}\grave{i}\ .$$

где x_{12} – высота вылета ОЦТ; x_9 – скорость вылета ОЦТ; x_{21} – прыжок вверх с трёх шагов. При этом прогнозное значение результата для мастеров спорта международного класса составляет 2,36 см, что отличается от их среднего результата (2.33 см) всего на 3 см.

Литература

1. Ахметов Р.Ф. Групповые статистические характеристики и факторный анализ многомерной совокупности параметров спортсменов в задачах прогноза результативности // Педагогіка, психологія та медико-біологічні проблеми фізичного виховання і спорту. – 2004. – № 6. – С. 91-104.
2. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорте. – М.: Физкультура и спорт, 1986. – 193 с.
3. Гантмахер Ф.Р. Теория матриц. – 4-е изд.– М.: Наука, 1988. – 552 с.
4. Крамер Г. Математические методы статистики: Пер. с англ. / Под ред. академика А.Н. Колмогорова. – М.: Мир, 1975. – 648 с.
5. Платонов В.Н. Общая теория подготовки спортсменов в олимпийском спорте. – К.: Олимпийская литература, 1997. – 583 с.
6. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте. – Л.: ВДКИФК, 1980. – 79 с.
7. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.
8. Саати Т.Л. Взаимодействие в иерархических системах // Техническая кибернетика. – 1979. – № 1. – С. 68-84.
9. Harman H.H. Modern factor analysis. – University of Chicago Press, 1960. Русский перевод: Современный факторный анализ. – М.: Статистика, 1972. – 516 с.

10. Lawley D.N., Maxwell A.E. Factor analysis as a statistical method. – Butterworths. – London, 1963. Русский перевод: Факторный анализ как статистический метод. – М.: Мир, 1967. – 413 с.