УДК 796.431.1

**Р.Ф. Ахметов** – кандидат педагогічних наук, професор, декан факультету фізичного виховання і спорту, завідувач кафедри теорії і методики фізичного виховання Житомирського державного університету імені Івана Франка, заслужений працівник фізичної культури і спорту України

# ПРОГНОЗУВАННЯ РЕЗУЛЬТАТИВНОСТІ СТРИБУНІВ У ВИСОТУ

*Роботу виконано на кафедрі теорії і методики фізичного виховання Житомирського державного університету імені Івана Франка*

Розглядається простий алгоритм підвищення точності раннього прогнозу результативності спортсменів, оснований на розширенні кількості інформативних спортивних параметрів.

**Ключові слова:** апроксимація, регресійна матриця, лінійна регресія.

**Akhmetow R.F. The possibility of high jumping results prognosis.** The item features an elementary algorithm of stepping-up the accuracy of predicting thlets' performance efficiency, based on the increase of informative athletic parameters.

**Key words:** approximation, regression matrix, linear regression.

**Постановка проблеми.** В останні роки українським стрибунам у висоту не вдається перемагати на великих міжнародних змаганнях. Цей факт стимулює фахівців не тільки підвищувати ефективність тренувального процесу, але і продовжувати розробку точності прогнозу результативності стрибунів у висоту, що значною мірою буде сприяти якісному відбору в цьому виді спорту. У зв'язку з цим досить актуальною є розробка програми прогнозу результативності на базі деякої сукупності параметрів спортсменів.

**Аналіз останніх досліджень.** Прогнозування ґрунтується на використанні методу екстраполяції, який припускає поширення висновків, отриманих зі спостереження над одною частиною якого-небудь явища, на інші його частини [3; 6; 7]. В умовах спорту екстраполяція дозволяє здійснити прогноз зростання результативності на основі вивчення відповідних закономірностей у попередні роки. Завдання прогнозу результативності спортсменів можна вирішити на базі факторного аналізу й динаміки розвитку фізичних параметрів і результатів на деякому обмеженому інтервалі часу (наприклад, 10-13 років) [1; 2; 7]. Для цього проводиться лінійна інтерполяція результатів і фізичних параметрів спортсменів між річними атестаційними періодами на менші часові періоди – піврічні та квартальні. Тоді в завданні синтезу лінійної багатомірної регресії результативності представляється можливим використовувати більше число інформативних параметрів.

**Мета дослідження** – розробити програму раннього прогнозу результативності стрибунів у висоту за результатами аналізу вікових груп до 13 років. Для цього спортивні результати та значення усереднених фізичних параметрів спортсменів лінійно інтерполювалися на піврічні чи квартальні періоди.

**Результати досліджень.** У цій роботі дається продовження загального підходу [2] до прикладного завдання прогнозування результативності стрибунів у висоту. Оскільки результати та фізичні параметри спортсменів у групі мають випадковий розкид (дисперсію) [1], то, говорячи про завдання прогнозування результативності, має сенс розглядати прогноз середньої результативності $\overline{\mathrm{H}}(t)$, як функції середніх по групі фізичних параметрів $\vec{X}_P$, що будемо представляти у вигляді матриці стовпця:

$$\vec{\mathrm{X}}_P = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_P \end{pmatrix}, \ \mathrm{P}=1,2,\ldots,\mathrm{N_p}\text{-}1;\ \mathrm{N_p} \geq 3,$$

де $N_p$ – повне число спортивних параметрів, включаючи сам результат (H). Повна множина P-мірних групувань з ($N_p$-1) по P дорівнює числу сполучень з ($N_p$-1) по P:

$$\vec{X}_P \in U_{\vec{X}_P} = \{\vec{X}_P^{\alpha}, \alpha = 1,2,\ldots,C_{N_p-1}^P\} \quad , (1)$$

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

Так, для стрибунів у висоту виділяється наступна повна сукупність спортивних параметрів [2]:

**РОЗШИРЕНИЙ ПЕРЕЛІК З 21 ПАРАМЕТРА СТРИБУНІВ У ВИСОТУ**

**1. Спортивний результат (висота) – Цільова функція.**

**АНТРОПОМЕТРИЧНІ ПАРАМЕТРИ (2-7)**

2. Зріст.
3. Довжина гомілки.
4. Довжина стегна.
5. Окружність стегна.
6. Окружність гомілкового м’язу.
7. Вага.

**ТЕХНІЧНІ ПАРАМЕТРИ (8-14)**

(Зареєстровані та розрахункові показники технічної підготовки)

8. Швидкість розбігу перед відштовхуванням.
9. Швидкість вильоту ЗЦТ (у момент відриву).
10. Кут вильоту ЗЦТ.
11. Тривалість фази відштовхування.
12. Висота вильоту ЗЦТ.
13. Імпульс сили відштовхування.
14. Ступінь використання силових можливостей поштовху (%).

**СПЕЦІАЛІЗОВАНІ ПАРАМЕТРИ (15-21)**

**(**Рівень спецфізпідготовки)

15. Біг – 30 м (с).

16. Швидкість спринтерського бігу (10 м з ходу).

17. Стрибок вгору у висоту з двох ніг з місця.

18. Стрибок у довжину з місця.

19. Потрійний стрибок з місця

20. Стрибок вгору з штовхової ноги (махом іншої).

21. Стрибок вгору у висоту з трьох кроків.

Інформативність різних Р-мірних групувань $\vec{\mathrm{X}}_P$ у завданнях прогнозування результативності буде також різною. Питання про вибір оптимальної сукупності найбільш інформативних параметрів з множини (1) при різних Р вимагає самостійних глибоких досліджень у рамках окремої НДР. У роботі [2] запропоновано один з альтернативних варіантів вирішення завдання, який цілком прийнятний з погляду точності прогнозу. У першому наближенні розглядається завдання лінійного прогнозу в рамках класичної теорії лінійної регресії (інтерполяції) у математичній статистиці [5; 8]. Мова йде про вирахування апроксимації

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

де $\mathrm{H}_0, \alpha_1, \ldots, \alpha_P$ – невідомі параметри регресії, які потрібно оцінити за даними деякої кількості вікових груп. У більш точній постановці наближена лінійна регресія (2) представляється у вигляді:

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

де $\xi(t)$ – помилка прогнозу з нульовим середнім ($M\xi(t) = 0$) і невідомою дисперсією $\sigma_\xi^2 = M\xi^2$ ($M$ – оператор математичного очікування – середнього). Якщо в результаті вирішення задачі лінійної регресії на інтервалі часу $T$ отримані оцінки невідомих параметрів регресії:

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

то прогнозне значення середньої результативності поза цим інтервалом подається у вигляді:

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

де набір фізичних параметрів $\{\, X_n(t_0), \quad n = 1,2,..., P \,\}$ – задається на прогнозований момент часу $t_0$. При цьому середньоквадратична похибка (СКП) прогнозу оцінюється величиною $\sigma_\xi(T)$. Наскільки „вдало” отримана оцінка (4), – залежить від багатьох факторів і останнє слово тут за практикою (експериментальною апробацією). Проведена в роботі [1] апробація моделі (4) показує, що вона практично цілком прийнятна. СКП при цьому не перевищує 3-х сантиметрів, а прогнозований рекордний результат становить 250 см. Залежність (4) прогнозованого значення результативності від часу (вікової групи) називається далі прогнозованою динамічною характеристикою результативності (ПДХР). Як показано в роботах [1; 2], для розрахунку ПДХР потрібно виконати необхідну умову: $N \ge P - 2$, де N – обсяг тимчасової вибірки (число аналізованих вікових груп). При цьому точність прогнозу зростає зі збільшенням числа Р використовуваних інформативних спортивних параметрів. Отже, для одержання задовільної точності прогнозу результативності потрібно мати досить великий обсяг тимчасової вибірки N вікових результатів і усереднених (по групі спортсменів) фізичних спортивних параметрів. До цього часу найбільш поширеною є річна реєстрація результатів і фізичних параметрів спортсменів (як правило, після змагань) у віці від 10 до 17 років. У цьому випадку обсяг тимчасової вибірки обмежується величиною $N_1$=8 або $N_2$=9 (якщо реєструються ще й результати майстрів спорту міжнародного класу). У зв’язку з цим можливості раннього прогнозу результативності, наприклад, за результатами аналізу у вікових групах 10-12 (13) років виявляються досить обмеженими. Як показано в роботах [1; 2], досить задовільний прогноз результативності стрибунів у висоту за трьома важливими інформативними параметрами ($X_{12}$, $X_9$, $X_{21}$) можливий тільки при N=5 (вік від 10 до 14 років) на період до 17 років.

Нами розроблена спеціальна модифікована програма cor2din.pas у оболонці Turbo Pascal, яка дозволяє розрахунковим способом лінійної інтерполяції збільшити обсяг тимчасової вибірки до $N_d$=17 (17=9*2-1). При цьому у випадку піврічної інтерполяції для 3-х мірної сукупності фізичних параметрів ($X_{12}$, $X_9$, $X_{21}$) величина N=5 відповідає „граничному" вікові 12 років. Однак, як показали розрахунки, задовільну точність прогнозу вдається одержати не при N=5, а починаючи з N=6, що відповідає граничному вікові 12,5 років.

**Матричне вирішення задачі лінійної регресії результативності за заданою сукупністю найбільш інформативних параметрів**

Для оцінки параметрів регресії $\mathrm{H}_0, \alpha_1, \dots, \alpha_P$ складається така система лінійних алгебраїчних рівнянь:

$$
\begin{gathered}
\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_1) = \overline{\mathrm{H}}(t_1) \\
\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_2) = \overline{\mathrm{H}}(t_2) \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
\mathrm{H}_0 + \sum_{m=1}^{P} \alpha_m X_m(t_N) = \overline{\mathrm{H}}(t_N)
\end{gathered} \tag{5}
$$

,

де N – число вікових груп (у цій роботі N<18). Система (5) подається в матричному вигляді:

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

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

Вводячи так званий „сигнальний” регресійний вектор (СРВ):

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

$$s_1 = \mathrm{H}_0 , s_2 = \alpha_1, s_3 = \alpha_2, \dots, s_M = \alpha_P \quad , \text{(7)}$$

матричну систему (6) подаємо також у стандартному вигляді:

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

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

де $Y_{NM}$ – вимірна матриця спостережень (ВМС); $\vec{\bar{\mathrm{H}}}_N$ – вимірний вектор середніх результатів (ВСР).

Відповідно до загальної теорії лінійної регресії система (8) може бути вирішена, якщо вона цілком визначена чи перевизначена:

$$N \ge M + 1 = P + 2 \Rightarrow \quad Rank Y_{NM} = M \quad . \text{(9)}$$

Відзначимо, що величина (M+1) обумовлена тим, що в число невідомих крім M=P+1 невідомих параметрів регресії необхідно включити також і невідому СКП $\sigma_\xi$. При виконання умови (9) статистичне вирішення задачі лінійної регресії подається у вигляді:

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

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

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

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

де $Y_{NM}^{-}$ – псевдозворотна матриця [4]; $\Lambda_{NN}^M$ – вектор у лінійну оболонку з базисних векторів $\{ \vec{Y}_N^m, m = 1,2,\dots,M \}$; $\Lambda_{NN}^{M\perp}$ – ортогональний вектор.

У даній роботі найбільш точне вирішення отримане у випадку P=3 при різних N з урахуванням необхідної умови припущення (9):

$$5 \le N \le 8. \quad (12)$$

Специфічною математичною особливістю задачі регресії спортивного результату є те, що в силу досить однорідного складу груп стовпцеві вектори ВМН $Y_{NM}$ є хоч і випадковими, але з малою кутовою розбіжністю відносно „одиничного" вектора $\vec{1}_N$. Ця обставина вимагає чіткого контролю точності перетворення матриці Грама $(Y_{NM}^T Y_{NM})_{MM}$, тому що у випадку високої кутової кореляції („схожості") векторів $\vec{Y}_N^m$ матриця Грама є часто погано зумовленою [4] з великим динамічним діапазоном власних чисел в області малих величин. При цьому точність перетворення матриці Грама зі зростанням розмірності P>3 (числа інформативних параметрів, які враховуються) починає різко падати й подальше збільшення розмірності P не є можливим.

Відзначимо також, що в цій роботі максимальне число вікових груп з піврічним періодом $N_{max}$=17.

Тому в силу умови (9) граничне число найбільш інформативних параметрів обмежується величиною 15 (у роботі [2] вона була 6).

**Апробація алгоритмів прогнозу результативності стрибунів у висоту за різною кількістю піврічних вікових груп**

У програмі РЕГРЕСІЯ (cor2din.com) є такі розділи:

1. Виклик вихідних статистичних даних (файл g1_21_9.dat).
2. Шифр файлу: TN-M (x1, x2, …, xM) для річних періодів і TNd-M (x1, x2, …, xM) для піврічних періодів, де N – число вікових груп (річних або піврічних), за якими робиться прогноз на майбутнє; M – число інформативних параметрів (N≥ M+2).
3. Інтерполяція значень фізичних параметрів на піврічні чи квартальні періоди.
4. Вибір M інформативних параметрів (з номерів 2-21).
5. Аналіз рангу регресійної матриці $Y_{N(M+1)}$ методом Грама-Шмідта.
6. Аналіз кореляції інформативних параметрів за роками.

7. Спектральний аналіз матриці Грама $Y^T Y$ розміром (M+1)*(M+1).

8. Оцінка точності перетворення матриці Грама.

9. Оцінка статистичних характеристик інформативних параметрів (середні, СКП, кореляційна матриця).

10. Вирішення задачі лінійної регресії.

11. Оцінка дисперсії шуму (СКП=s).

12. Прогнозування за межі обраних вікових груп, включаючи прогнозування рекордних результатів.

**Висновки**

1. Задача прогнозу результативності спортсменів є задачею інтерполяції середньої (по віковій групі) результативності ($\overline{\mathrm{H}}$) у вигляді лінійної комбінації середніх значень найбільш інформативних фізичних параметрів спортсменів $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_P)$ із зазначенням точності (СКП) прогнозу:

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

де $\mathrm{H}_0, \alpha_1, \alpha_2, ..., \alpha_P$ – параметри регресії; $\sigma_\xi$ – СКП прогнозу.

В умовах апріорної невизначеності про СКП прогнозу необхідною умовою вирішення задачі прогнозу є перевищення числа використовуваних вікових груп ($N_{ВГ}$) над числом використовуваних інформативних фізичних параметрів (Р), як мінімум на дві одиниці:

$$\mathrm{NВГ} \geq \mathrm{P}+2 .$$

Так, при кількості інформативних фізичних параметрів Р=3 потрібні середні значення більш, ніж по 5-ти річних вікових групах (10, 11, 12, 13 і 14 років) або більш, ніж по 6-ти піврічних вікових групах (10; 10,5; 11; 11,5; 12 і 12,5 років). При цьому можна зробити прогноз результативності не тільки на будь-який „внутрішній" момент часу $t_0$ $(10 \leq t_0 \leq 14)$ або $(10 \leq t_0 \leq 12.5)$, але й на майбутні моменти часу $t_0$>14 або $t_0$>12,5, включаючи прогноз рекордних результатів. Для цього достатньо в отриману формулу регресії підставити значення прогнозних середніх значень фізичних параметрів $\{\, \bar{x}_n(t_0),\ n = 1,2,...,\ P \,\}$:

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

Зокрема, при прогнозуванні за трьома параметрами ($x_{12}$, $x_9$, $x_{21}$) по 5-ти річних вікових групах (10, 11, 12, 13 і 14 років) отримана така регресійна функція:

$$H = 0.478 + 0.657\,x_{12} + 0.058\,x_9 + 0.806\,x_{21}\,, \quad \sigma = s = 0.9\,см.$$

де $x_{12}$ – висота вильоту ЗЦТ; $x_9$ – швидкість вильоту ЗЦТ; $x_{21}$ – стрибок вгору з трьох кроків. При цьому прогнозоване значення результату для провідних майстрів спорту складає 2,36 см, що відрізняється від їхнього середнього результату (2,33 см) усього на 3 см.

При прогнозуванні за трьома параметрами ($x_{12}$, $x_9$, $x_{21}$) по 6-ти піврічних вікових групах (10; 10,5; 11; 11,5; 12 і 12,5 років) отримана така регресійна функція:

$$H = 0.381 + 0.474\,x_{12} + 0.027\,x_9 + 1.369\,x_{21}\,, \quad \sigma = s = 0.2\,см.$$

## *Література*

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. Ахметов Р.Ф. Прогноз результативности спортсменов на базе статистического факторного анализа и экспертного ранжирования полной совокупности антропометрических, технических и специализированных параметров // Педагогіка, психологія та медико-біологічні проблеми фізичного виховання і спорту. – 2004. – № 7. – С. 82-95.
3. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорте. – М.: Физкультура и спорт, 1986. – 193 с.
4. Гантмахер Ф.Р. Теория матриц. – 4-е изд.– М.: Наука, 1988. – 552 с.

5. Крамер Г. Математические методы статистики: Пер. с англ. / Под ред. академика А.Н. Колмогорова. – М.: Мир, 1975. – 648 с.
6. Платонов В.Н. Общая теория подготовки спортсменов в олимпийском спорте. – К.: Олимпийская литература, 1997. – 583 с.
7. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте. – Л.: ВДКИФК, 1980. – 79 с.
8. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.