ВИДІЛЕННЯ НАЙБІЛЬШ ІНФОРМАТИВНИХ ПАРАМЕТРІВ СТРИБУНІВ У ВИСОТУ В ЗАДАЧАХ ПРОГНОЗУ ЇХ РЕЗУЛЬТАТИВНОСТІ

Ахметов Рустам Фагимович

Анотація. У статті розглядаються актуальні питання виділення найбільш інформативних узагальнених параметрів стрибунів у висоту з деякої повної сукупності спортивних параметрів для вирішення в подальшому важливого завдання прогнозу результативності.

Ключові слова: вектор спортивних параметрів, кореляційний еліпсоїд, цільова функція.

Постановка проблеми. У системі спортивної підготовки велика роль приділяється прогнозу зростанню тактико-технічних і функціональних можливостей окремих спортсменів і спортивних груп [1; 5]. Проблема полягає в тому, що використання всієї повної сукупності спортивних параметрів для рішення задач прогнозу результативності відомими математичними методами поки не представляється можливим через досить велику їхню кількість ($P_{макс}>20$) у той час, як число вікових груп (S), використаних для прогнозу, обмежується звичайно величиною $S_{макс}=7$ (10-16 років). Тому виникає задача вибору серед великої сукупності спортивних параметрів найбільш інформативної сукупності меншої розмірності $P<S_{макс}$. Більше того, можна показати, що в задачах лінійного прогнозу повинна виконуватися необхідна умова $P<S-1$ $(S>2)$.

Аналіз останніх досліджень і публікацій. Прогнозування ґрунтується на використанні методу екстраполяції, що припускає поширення висновків, отриманих зі спостереження над однією частиною якого-небудь явища, на інші його частини [1; 4; 5]. В умовах спорту екстраполяція дозволяє здійснити прогнози зростання результативності на основі вивчення відповідних закономірностей у попередні роки. Проведений у деяких роботах [1; 5] детермінований аналіз повної сукупності параметрів спортсменів (антропометричних, технічних і спеціалізованих) розкриває їхній фізичний зміст і показує, що всі вони є важливими характеристиками, які в сукупності й визначають, у кінцевому рахунку, спортивний результат (цільову функцію). Однак, детермінований аналіз не відповідає на дуже істотне питання: а яким чином оцінювати кількісно ступінь впливу на результат окремих параметрів чи деякої групи параметрів? Заздалегідь ясно, що цей вплив різний для різних параметрів і різних груп параметрів, і він залежить, звичайно, від віку групи. Оскільки конкретні значення (реалізації) параметрів залежать випадковим образом від конкретного спортсмена, остільки вони завжди мають деякий випадковий розкид, який можна описати методами математичної статистики [3; 6]. При цьому особливе значення має факторний аналіз [6-8], тому що основною метою факторного аналізу є виділення найбільш інформативних і значимих параметрів з деякої безлічі випадкових параметрів. У зв’язку з цим виникає, насамперед, важлива задача статистичного аналізу повної сукупності спортивних параметрів для кожної вікової групи та наступного упорядкування цієї сукупності за ступенем інформативності. Така постановка вирішення поставленої проблеми нам не зустрічалася у спортивній літературі. Дана робота є введенням у цикл робіт, виконуваних згідно Об’єднаного Плану НДР Держкомспорту України з питань фізичної культури і спорту на 2001-2005 рр. по темі “Теоретико-методичні основи раціональної підготовки спортсменів на різних етапах їх багаторічного вдосконалення”.

Ціль статті. Стаття носить постановочний і програмний характер. Її основна мета – позначити перелік основних питань, що вимагають вирішення в рамках НДР, пов’язаних із розробкою сучасної, досить науково-обґрунтованої методики виділення найбільш інформативної сукупності інформативних параметрів стрибунів у висоту в задачах прогнозу їх результативності. При цьому центральне місце займають питання факторного аналізу й оцінки максимального числа інформативних параметрів спортсменів, який можна використовувати в задачах прогнозу результативності.

Результати досліджень

1. Векторні і матричні статистичні характеристики для групи спортсменів

Повна сукупність параметрів, включаючи і спортивний результат (Н), представляється у вигляді деякого N-мірного вектора $\vec{x}_N$ (матриці-стовпця):

$$\vec{x}_N^T = (x_1, x_2, ..., x_N),$$

де „Т" – операція матричного транспонування, $\vec{x}_N^T$ – рядок, $\vec{x}_N$ – стовпець. У цій роботі дослідження обмежується випадком N=21: x_1 = Н – спортивний результат (висота стрибка; називається також цільовою функцією (ЦФ)).

Антропометричні параметри ($x_2 ... x_7$): x_2 – зріст; x_3 – довжина гомілки; x_4 – довжина стегна; x_5 – окружність стегна; x_6 – окружність литкового м'яза; x_7 – вага.

Технічні параметри ($x_8 ... x_{14}$): x_8 – швидкість розбігу перед відштовхуванням; x_9 – швидкість вильоту ЗЦТ (у момент відриву); x_{10} – кут вильоту ЗЦТ; x_{11} – тривалість фази відштовхування; x_{12} – висота вильоту ЗЦТ; x_{13} – імпульс сили відштовхування.

Спеціалізовані параметри ($x_{14} ... x_{21}$): x_{14} – ступінь використання силових можливостей при відштовхуванні (%); x_{15} – біг на 30 м з високого старту (час, секунди); x_{16} – швидкість спринтерського бігу (10 м з ходу); x_{17} – стрибок угору з двох ніг із місця; x_{18} – стрибок у довжину з місця; x_{19} – потрійний стрибок із місця; x_{20} – стрибок угору з штовхової ноги (махом іншої); x_{21} – стрибок угору з трьох кроків.

Векторний параметр спортсмена (ВПС) $\vec{x}_N$ залежить від конкретного спортсмена m=1,2,…,M у групі з M спортсменів (у даній роботі M=12). Залежність ВПС від спортсмена (його номера) і від часу (віку) представляється у вигляді :

$$\vec{x}_N = \vec{x}_N^m(t),\ t = t_1, t_2, ..., t_L, t_0;\ \text{m=1,2,…,M},$$

$$t_n = 10 + (n-1), \quad n = 1,2,...,8,$$

де L – число вікових груп (у даній роботі L=8); t_0 – умовний вік провідних спортсменів. Для простоти залежність ВПС від часу поки-що опускається і вікова група цілком характеризується M-мірним набором N-мірних ВПС спортсменів і подається у вигляді прямокутної матриці X_{NM}, яка називається далі груповою параметричною матрицею (ГПМ):

$$X_{NM} = (\vec{X}_N^1 \vec{X}_N^2 ... \vec{X}_N^M) = (x_{nm})_{NM},\ x_{nm} = \vec{X}_N^m[n],$$

$$X_{NM} = \begin{pmatrix} x_{11}x_{12}\dots x_{1M} \\ x_{21}x_{22}\dots x_{2M} \\ \dots\dots\dots\dots\dots \\ x_{N1}x_{N2}\dots x_{NM} \end{pmatrix}, \tag{1}$$

де N – число рядків; M – число стовпців матриці; x_{nm} – елементи матриці (n-а компонента (координата) вектора $\vec{X}_N^m$).

Виділяючи окремо спортивний результат $x_1 = H$, ВПС $\vec{x}_N$ можна представити також у блоковому вигляді:

$$\vec{x}_N = \begin{pmatrix} H \\ \vec{y}_{N-1} \end{pmatrix}, \quad \vec{y}_{N-1} = \begin{pmatrix} x_2 \\ x_3 \\ . \\ . \\ x_N \end{pmatrix}, \tag{2}$$

де $\vec{y}_{N-1}$ – (N-1) – мірний вектор фізичних параметрів (ВФП) спортсмена.

У рамках статистичної термінології будемо вважати, що кожний із параметрів x_n (для кожної вікової групи) є деякою випадковою величиною, а ВПС $\vec{x}_N$ – випадковим вектором. Статистичні характеристики ВПС визначаються шляхом арифметичного усереднення:

$$\vec{a}_N = \bar{\bar{x}}_N = \frac{1}{M}\sum_{m=1}^{M} \vec{X}_N^m, \quad a_n = \bar{x}_n = \frac{1}{M}\sum_{m=1}^{M} x_{nm}, \tag{3}$$

$$D(x_n) = \sigma_n^2 = \overline{\Delta x_n^2} = \overline{x_n^2} - \overline{x_n}^2, \quad \Delta x_n = x_n - \overline{x_n}, \tag{4}$$

$$\Phi_{nk} = \overline{x_n x_k} = \frac{1}{M}\sum_{m=1}^{M} x_{nm}x_{km}\,, \tag{5}$$

$$\Psi_{nk} = \overline{\Delta x_n \Delta x_k} = \Phi_{nk} - \overline{x_n}\,\overline{x_k}\,, \tag{6}$$

$$\Psi_{nk} = \sigma_n \sigma_k \rho_{nk}\,, \quad \rho_{nk} = \frac{\Psi_{nk}}{\sigma_n \sigma_k}\,, \tag{7}$$

де a_n, σ_n^2 – середні значення і дисперсії параметрів x_n ($\sigma_n = \sqrt{D(x_n)}$ – СКО); Δx_n – флуктуації параметрів щодо середніх значень; Φ_{nk}, Ψ_{nk} – взаємні кореляції та коваріації параметрів x_n, x_k; ρ_{nk} – взаємні коефіцієнти кореляції ($/\rho/ \le 1$).

Відповідні кореляційні та коваріаційні матриці представляються в алгебраїчному вигляді:

$$\Phi_{NN} = \frac{1}{M} X_{NM} X_{NM}{}^{T} = \overline{\vec{X}_N^m \vec{X}_N^{mT}} = \frac{1}{M}\sum_{m=1}^{M} \vec{X}_N^m \vec{X}_N^{mT}\,, \tag{8}$$

$$\Psi_{NN} = \overline{\Delta\vec{X}_N^m \Delta\vec{X}_N^{mT}}\,, \tag{9}$$

де риска зверху означає арифметичне усереднення за номером m $(m = 1,2,\ldots,M)$, тобто статистичне усереднення по спортсменах у групі з рівномірним дискретним розподілом імовірностей $p_m = 1/M$.

Відзначимо, що вихідна ГПМ X_{NM} містить інформацію не тільки про зв'язок різних параметрів x_n між собою, але і ступеня „схожості” чи параметричної близькості спортсменів між собою в групі. Для цього досить розглянути близькість чи кореляцію векторів $\vec{X}_N^m$, оцінюючи скалярні добутки векторів [2]:

$$B_{mk} = \frac{1}{N}(\vec{X}_N^m, \vec{X}_N^k) = \frac{1}{N}\sum_{n=1}^{N} \vec{X}_N^m[n]\vec{X}_N^k[n]\,. \tag{10}$$

Матрицю скалярних добутків (МСП) можна представити через ГПМ X_{NM} :

$$B_{MM} = \frac{1}{N} X_{NM}^T X_{NM}\,. \tag{11}$$

Мірою параметричної близькості спортсменів у групі може служити алгебраїчна кореляція $\vec{X}_N^m$ векторів чи так званий косинус кута між векторами:

$$R_{mk} = \cos\varphi_{mk} = \frac{(\vec{X}_N^m, \vec{X}_N^k)}{// \vec{X}_N^m //* // \vec{X}_N^k //}, \tag{12}$$

$$// \vec{X}_N // = \sqrt{(\vec{X}_N, \vec{X}_N)} = \sqrt{\sum_{n=1}^{N} x_n^2},$$

де $// \vec{X}_N //$ – норма вектора в N-мірному евклідовому просторі [2].

2. Багатомірний нормальний закон розподілу та кореляційний еліпсоїд вектора спортивних параметрів. Задача факторного аналізу

Нормальна щільність імовірності ВСП представляється в стандартному вигляді [2]:

$$W(\vec{X}_N / \bar{\vec{X}}_N, \Psi_{NN}) = \frac{1}{\sqrt{(2\pi)^N \det(\Psi_{NN})}} \exp\{-\frac{1}{2}(\Psi_{NN}^{-1}\Delta\vec{X}_N, \Delta\vec{X}_N)\},$$

де $\det(\Psi_{NN})$ – визначник коваріаційної матриці Ψ_{NN}.

Перетин щільності ймовірності визначає у просторі ВСП так званий кореляційний еліпсоїд:

$$W(\vec{X}_N /.) = const \quad \Rightarrow \quad (\Psi_{NN}^{-1}\Delta\vec{X}_N, \Delta\vec{X}_N) = const'. \tag{13}$$

Зокрема, у випадку незалежних параметрів x_n рівняння кореляційного еліпсоїда представляється у вигляді:

$$\left(\frac{x_1 - \bar{x}_1}{\sigma_1}\right)^2 + \left(\frac{x_2 - \bar{x}_2}{\sigma_2}\right)^2 + \ldots + \left(\frac{x_N - \bar{x}_N}{\sigma_N}\right)^2 = const'.$$

Відзначимо, що при належному виборі постійної $const'$ ВСП $\vec{X}_N$ знаходиться з високою ймовірністю всередині свого кореляційного еліпсоїда. У загальному випадку багатомірний кореляційний еліпсоїд характеризується своїми розмірами й орієнтацією, що визначаються в результаті рішення задачі про приведення квадратичної форми (13) до канонічного вигляду. На малюнку 1 приведено кореляційний еліпсоїди ВСП на площині двох параметрів:

На відміну від більшості відомих робіт, у даній роботі факторний аналіз розглядається з позицій аналізу орієнтації та розмірів багатомірного кореляційного еліпсоїда повного вектора спортивних параметрів (ВСП). При цьому виділяється так званий принцип локалізації ВСП в обмежених підпросторах меншої розмірності, коли розміри кореляційного еліпсоїда в деяких головних напрямках стають нехтувано малими величинами. Потрібно, однак, підкреслити одну специфічну особливість статистичної обробки параметрів у малій групі спортсменів. Це принципова обмеженість числа спортсменів у групі ($M = 12$), що може призвести до великих відносних погрішностей середньоарифметичних оцінок невідомих статистичних середніх (при $M = 12$ вони складають 30-47% [6]). У зв'язку з цим необхідно, насамперед, уточнити, а з якою основною метою оцінюються групові статистичні параметри? І який узагалі мають сенс „арифметичні" статистичні характеристики? У цій роботі основною метою є вирішення завдання прогнозу результативності за деякою сукупністю інформативних параметрів спортсменів у залежності від методики тренування. Тому на першому етапі досліджень питання впливу погрішностей арифметичних оцінок самих статистичних характеристик у цій роботі поки опускаються, а арифметичне усереднення розглядається просто, як аналог і окремий випадок статистичного усереднення (з рівномірним розподілом імовірності) для вирішення питань локалізації та факторного аналізу ВСП. Обґрунтуванням і критерієм корисності такого підходу є досить прийнятне для практики вирішення кінцевого завдання прогнозу результативності.

3. Сингулярні числа ГПМ і максимальне число найбільш інформативних параметрів спортсменів

У більшості випадків число аналізованих фізичних параметрів перевищує кількість спортсменів у групі: $N > M$.

У цьому випадку ранги симетричних матриць Φ_{NN} і B_{MM} збігаються і рівні M :

$$Rank\Phi_{NN} = Rank\mathrm{B}_{MM} = M \,. \tag{3.18}$$

Це випливає з того, що строкові та стовпцеві ранги довільних матриць збігаються [5]. Більше того, можна показати, що ненульові власні числа матриць $(X_{NM} X_{NM}^T)_{NN}$ і $(X_{NM}^T X_{NM})_{MM}$ збігаються та дорівнюють квадратам сингулярних чисел ГМП X_{NM} [2].

Таким чином, у випадку N=21 і M=12 серед двадцяти фізичних параметрів можна методами математичної статистики виділити для задач прогнозу не більш дванадцяти інформативних параметрів. У наступних науково-дослідних роботах представляється доцільним формувати об'єднані групи спортсменів з декількох автономних груп для забезпечення нерівності M>N. Тоді для задач прогнозу результативності можна використовувати всі N параметрів.

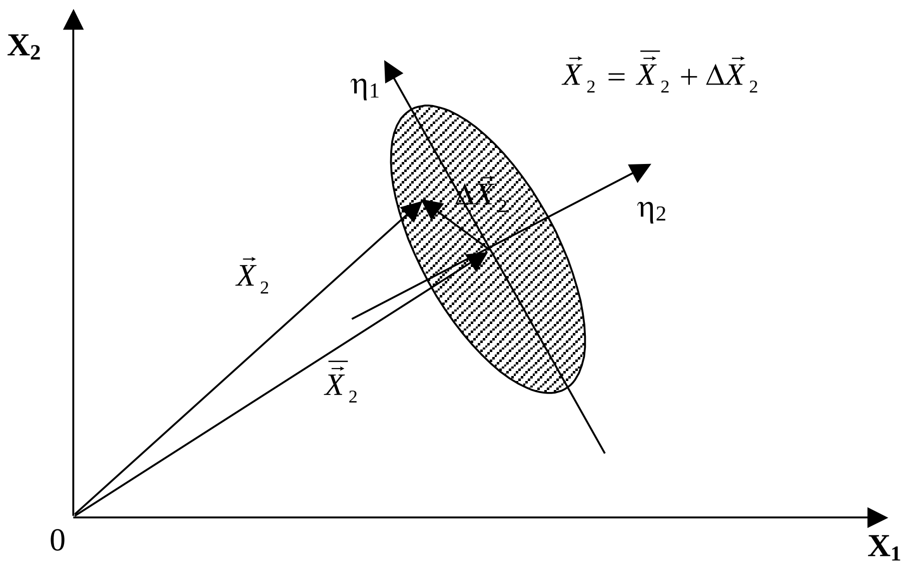

Рис. 1. Кореляційний еліпсоїд (еліпс) на площині двох (N=2) параметрів (X_1, X_2).

Висновки

Отримані результати та проведене попереднє дослідження 8 вікових груп спортсменів (включаючи групу майстрів спорту міжнародного класу) по 12 спортсменів у групі дозволяють зробити такі висновки:

1. Спостерігається дуже однорідний склад груп у розумінні параметричної близькості спортсменів у групі. Фізичні параметри виявляються майже детермінованими з малою дисперсією, – що й обумовлює їхню параметричну близькість. Остання обставина висуває підвищені вимоги до точності

спектрального алгебраїчного аналізу кореляційних матриць параметрів (eps < 10^{-12}).

2. Задача факторного аналізу про виділення найбільш інформативних параметрів спортсменів означає, власне кажучи, розкриття області локалізації вектора фізичних параметрів (ВФП) у деякому обмеженому підпросторі повного багатомірного евклідового простору параметрів. При цьому базисом підпростору є набір перших „значимих” власних векторів коваріаційної матриці ВФП, які визначають орієнтацію кореляційного еліпсоїда ВФП. Власні значення коваріаційної матриці ВФП визначають розмір кореляційного еліпсоїда, у якому локалізується ВФП.

3. Спектральний аналіз кореляційних матриць параметрів підтверджує теоретичний висновок про максимальне число інформативних параметрів, яке дорівнює числу спортсменів у групі (М = 12). При цьому спостерігається різке падіння власних чисел матриць, починаючи з номерів 4-7. Звідси випливає, що для завдань прогнозу ЦФ на першому етапі достатньо обмежитися трьома-шістьма найбільш інформативними параметрами: x_{12} (висота вильоту ЗЦТ); x_9 (швидкість вильоту ЗЦТ); x_{21} (стрибок угору з трьох кроків розбігу); x_5 (швидкість розбігу перед відштовхуванням); x_{15} (біг на 30 м з високого старту); x_{14} (ступінь використання силових можливостей при відштовхуванні).

Підводячи підсумок відмітимо, що склад найбільш інформативних комбінацій параметрів залежить від вікової групи. Тому питання про єдину сукупність найбільш інформативних параметрів для всіх груп залишається, строго говорячи, поки відкритим і тут потрібно провести ще додаткові самостійні дослідження в рамках окремих НДР. У даній роботі вибір трьох-шести мірних сукупностей зроблений з міркувань високої кореляції з ЦФ і максимальної частоти повторення в першій десятці власних векторів кореляційної матриці.

Литература

1. Баландин В.И., Блудов Ю.М., Плахтиенко В.А. Прогнозирование в спорте. – М.: Физкультура и спорт, 1986. – 193 с.
2. Гантмахер Ф.Р. Теория матриц. – 4-е изд.– М.: Наука, 1988. – 552 с.
3. Крамер Г. Математические методы статистики/ Пер.с англ. Под ред. академика А.Н. Колмогорова. – М.: Мир, 1975. – 648 с.
4. Платонов В.Н. Общая теория подготовки спортсменов в олимпийском спорте. – К.: Олимпийская литература, 1997. – 583 с.
5. Плахтиенко В.А., Мельник В.Г. Прогнозирование в спорте. – Л.: ВДКИФК, 1980. – 79 с.
6. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.
7. Harman H.H. Modern factor analysis. – University of Chicago Press, 1960. Русский перевод: Современный факторный анализ. – М.: Статистика, 1972. – 516 с.
8. 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.