УДК 669.018.14: 669.15 – 194: 621.78: 620.197

МОДЕЛЬ РЕДУКЦИИ ПРОСТРАНСТВА ОПИСАНИЯ СТАЛЕЙ

к.т.н., доц. О.П. Юшкевич, к.т.н., доц. В.И. Погорелый, студ. П.О. Юшкевич

Национальная металлургическая академия Украины

1. Постановка проблемы и её связь с важнейшими научными и практическими задачами. Общую характеристику сталей и сплавов определяют наиболее важные свойства. Такие свойства мало изменяются в пределах одной плавки, марки стали или при переходе от одного плавочного состава к другому. Они имеют наименьшую дисперсию в границах внутриплавочных, межплавочных выборок или в пределах марок сталей, которая также мало изменяется при переходе от стали одного производителя к другому. При этом дисперсия по всей совокупности близких по составу марок сталей может быть велика. Таким образом, важными являются те свойства, которым соответствуют наименьшие диагональные элементы в диагональной ковариационной матрице K_{mpt}, определяющей связи между характеристиками, и равные дисперсии их измерений:

$$k_{ii} = D\left(\sigma_{mpt}^{(i)}\right), \tag{1}$$

где k_{ii} – диагональный коэффициент ковариации свойства $\sigma_{mpt}^{(i)}$. При этом недиагональные $k_{ii'} = 0$. Корреляционная матрица, в этом случае, не позволяет оценить уровень важности свойств сталей. Так как, если она для выбранной плавки или марки стали диагональная, то является единичной. То есть, все элементы главной диагонали $r_{ii}^{mpt} = 1$.

Однако свойства с большой дисперсией дают представления об отклонениях в технологии производства сталей. Их дисперсии являются сигналами к выявлению факторов влияющих на качество металлоизделий. Поэтому свойства обладающие высокой дисперсией являются технологически чувствительными и могут быть наиболее информативными.

2. Анализ последних исследований и публикаций, в которых начато решение поставленной задачи. Судить о значимости физических, механических и химических свойств можно по величине недиагональных членов $r_{ii'}^{mpt}$ их корреляционной матрицы:

$$R_{mpt} = \left\| \begin{matrix} r_{11}^{mpt} & \cdots & r_{1i}^{mpt} & \cdots & r_{1\rho}^{mpt} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{i1}^{mpt} & \cdots & r_{ii}^{mpt} & \cdots & r_{i\rho}^{mpt} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{\rho 1}^{mpt} & \cdots & r_{\rho i}^{mpt} & \cdots & r_{\rho\rho}^{mpt} \end{matrix} \right\|, \tag{2}$$

которая образуется из корреляционной таблицы 1, составленной из свойств сталей $\sigma_{mpt}^{(i)}$, взятых из таблицы экспериментальных данных (ТЭД) [1] и их коэффициентов корреляции r_{ii}^{mpt}[2]. При этом значимыми являются взаимно независимые или мало зависимые друг от друга свойства, которым соответствуют наименьшие недиагональные элементы равные или близкие 0.

Таблица 1.

Корреляционная таблица свойств стали марки M_{pt}

Свойства Стали	$\sigma_{mpt}^{(1)}$	$\sigma_{mpt}^{(2)}$	…	$\sigma_{mpt}^{(i)}$	…	$\sigma_{mpt}^{(\rho)}$
$\sigma_{mpt}^{(1)}$	r_{11}^{mpt}	r_{12}^{mpt}	…	r_{1i}^{mpt}	…	$r_{1\rho}^{mpt}$
$\sigma_{mpt}^{(2)}$	r_{21}^{mpt}	r_{22}^{mpt}	…	r_{2i}^{mpt}	…	$r_{2\rho}^{mpt}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\sigma_{mpt}^{(i)}$	r_{i1}^{mpt}	r_{i2}^{mpt}	…	r_{ii}^{mpt}	…	$r_{i\rho}^{mpt}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\sigma_{mpt}^{(\rho)}$	$r_{\rho 1}^{mpt}$	$r_{\rho 2}^{mpt}$	…	$r_{\rho i}^{mpt}$	…	$r_{\rho\rho}^{mpt}$

Однако, если в результате обработки ТЭД установлено, что все измеряемые свойства $\sigma^{(1)}, \dots, \sigma^{(i)}, \dots, \sigma^{(p)}$ сильно коррелированны между собой, то есть они являются статистически зависимыми друг от друга (могут быть вычислены одни из других с помощью аппроксимирующих функций или регрессионных уравнений), то их корреляционная матрица недиагональная т.е. значения недиагональных элементов сильно отличаются от 0. Такие свойства не дают репрезентативного представления о качестве стали и часть этих характеристик может быть исключена из рассмотрения и анализа [2]. При этом разброс каждого из свойств мало влияет на положение центров облаков информационных объектов (точек) и, следовательно, на величину расстояния между математическими ожиданиями образов сталей или плавочных выборок [3] признакового пространства.

Поэтому судить о важности того или иного свойства по диагональным элементам недиагональной ковариационной матрицы, им соответствующим, нельзя.

3. Постановка задачи. Для четкой трактовки понятий важные, значимые и технологически чувствительные свойства сталей, представляется целесообразным ввести определения.

<u>Определение 1.</u> Важными называются те свойства сталей, которые имеют минимальную дисперсию.

Определение 2. Значимыми называют свойства, которые имеют наименьшие величины коэффициентов корреляции по отношению к иным характеристикам сталей.

Следствие из определения 2. Свойство стали S_{mpt} по отношению к одним характеристикам может быть значимым, а по соотношению с другими нет.

Определение 3. Технологически чувствительными признаками называются свойства сталей, которые имеют высокую дисперсию.

Таким образом, максимальной информативностью о технологическом процессе производства металлоизделий обладают характеристики, имеющие высокую значимость и технологическую чувствительность. При этом в совокупности приёмо-сдаточных свойств они могут быть не основными и находиться на последних местах в ряду признаков, выстроенном в порядке возрастания важности.

Цель работы. Разработать модель преобразований и перехода к новым признакам сталей, т.е. переменным y_{mpt}, которые образуют совокупность новых координат $(y_{mpt}^{(1)}, \ldots, y_{mpt}^{(j)}, \ldots, y_{mpt}^{\rho'})$, образующих вектор $\overline{Y_{mpt}}$, определяющий сталь S_{mpt} в новом признаковом пространстве $Y = (Y^{(1)}, \ldots, Y^{(j)}, \ldots, Y^{(\rho')})$, в котором можно получить новые характеристики, обладающие максимально возможной значимостью и технологической чувствительностью.

Описание пространства главных компонент марок сталей. В общем виде новые признаки можно представить в виде функций от первоначально выбранных свойств $\sigma_{mpt}^{(i)}$ металлоизделия из стали марки - *m*, имеющего конкретную форму - *p*, геометрические размеры- *t* :

$$\begin{aligned} y_{mpt}^{(1)} &= f_1^{mpt}(\sigma_{mpt}^{(1)}, \ldots, \sigma_{mpt}^{(i)}, \ldots, \sigma_{mpt}^{(\rho)}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{mpt}^{(j)} &= f_j^{mpt}(\sigma_{mpt}^{(1)}, \ldots, \sigma_{mpt}^{(i)}, \ldots, \sigma_{mpt}^{(\rho)}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_{mpt}^{(\rho')} &= f_{\rho'}^{mpt}(\sigma_{mpt}^{(1)}, \ldots, \sigma_{mpt}^{(i)}, \ldots, \sigma_{mpt}^{(\rho)}) \end{aligned} \quad (3)$$

Представляя функциональные связи $f_1^{mpt}, \ldots, f_j^{mpt}, \ldots, f_{\rho'}^{mpt}$, как линейные и применяя линейные преобразования матриц:

$$Y_{mpt} = A^{mpt}[\sigma_{mpt}] \quad (4)$$

где $[\sigma_{mpt}]$ - матрица первоначальных свойств стали,

$A^{mpt} = \{a_{ji}^{mpt}\}$ - матрица преобразования,

можно ковариационную матрицу для новых переменных y_{mpt} привести к диагональному виду и выбрать новые важнейшие признаки.

Таким образом, новые признаки y_{mpt} будут являться линейными комбинациями первоначально выбранных свойств сталей $\sigma_{mpt}^{(i)}$. При этом они могут оказаться более удобными для описания марок сталей, идентификации и распознавания их образов.

Если значения некоторых новых признаков сталей равны первоначальным свойствам с низким уровнем значимости или малой технологической чувствительностью: $y_{mpt}^{(j)} = \sigma_{mpt}^{(i)}$, то их можно убрать из совокупности переменных, образующих информационный вектор стали $\left(y_{mpt}^{(1)} \dots, y_{mpt}^{(j)}, \dots, y_{mpt}^{\rho'}\right) = \overrightarrow{Y_{mpt}}$ в новом пространстве её описания. Таким образом, можно получить первое сокращение размерности нового признакового пространства описания марочного сортамента сталей металлоизделий.

Если система уравнений (3) определяется как линейная комбинация (4) исходных свойств, то она может быть переписана в виде:

$$
\begin{gathered}
y_{mpt}^{(1)} = \sum_{i=1}^{\rho} a_{1i}^{mpt} \sigma_{mpt}^{(i)} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
y_{mpt}^{(j)} = \sum_{i=1}^{\rho} a_{ji}^{mpt} \sigma_{mpt}^{(i)} \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
y_{mpt}^{(\rho')} = \sum_{i=1}^{\rho} a_{\rho' i}^{mpt} \sigma_{mpt}^{(i)}
\end{gathered}
\qquad (5)
$$

где $a_{1i}, \dots, a_{ji}, \dots, a_{\rho'\rho}$ – коэффициенты преобразования системы координат или факторные нагрузки на новые признаковые оси $Y^{(1)}, \dots, Y^{(j)}, \dots, Y^{(\rho')}$, которые образуют матрицу:

$$A_{mpt} = \left\{a_{ji}^{mpt}\right\} \qquad (6)$$

преобразования первоначальных свойств в новые признаки:

$$
A_{mpt} = \begin{Vmatrix}
a_{11}^{mpt} & a_{12}^{mpt} & \dots & a_{1i}^{mpt} & \dots & a_{1\rho}^{mpt} \\
a_{21}^{mpt} & a_{22}^{mpt} & \dots & a_{2i}^{mpt} & \dots & a_{2\rho}^{mpt} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
a_{j1}^{mpt} & a_{j2}^{mpt} & \dots & a_{ji}^{mpt} & \dots & a_{j\rho}^{mpt} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
a_{\rho'1}^{mpt} & a_{\rho'2}^{mpt} & \dots & a_{\rho'i}^{mpt} & \dots & a_{\rho'\rho}^{mpt}
\end{Vmatrix} \qquad (7)
$$

Ее транспонированная матрица имеет вид:

$$
A_{mpt}^{T} = \begin{Vmatrix}
a_{11}^{mpt} & a_{21}^{mpt} & \dots & a_{j1}^{mpt} & \dots & a_{\rho'1}^{mpt} \\
a_{12}^{mpt} & a_{22}^{mpt} & \dots & a_{j2}^{mpt} & \dots & a_{\rho'}^{mpt} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
a_{1i}^{mpt} & a_{2i}^{mpt} & \dots & a_{ji}^{mpt} & \dots & a_{\rho'i}^{mpt} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
a_{1\rho}^{mpt} & a_{2\rho}^{mpt} & \dots & a_{j\rho}^{mpt} & \dots & a_{\rho'\rho}^{mpt}
\end{Vmatrix} \qquad (8)
$$

В результате линейных преобразований, полученные новые признаки $y_{mpt}^{(1)},\dots,y_{mpt}^{(j)},\dots,y_{mpt}^{\rho'}$ сталей S_{mpt} могут стать важными и значимыми. При этом они повышают однозначность, строгость и четкость описания образов марок сталей по сравнению с первоначальными, традиционными свойствами $\sigma^{(1)},\dots,\sigma^{(i)},\dots,\sigma^{(p)}$. Однако, для этого они должны отвечать ряду критериев [4]:

1) функциональности, в частности линейности

$$y_{mpt}^{(j)} = \sum_{i=1}^{\rho} a_{ji}^{mpt}\sigma_{mpt}^{(i)}, \tag{9}$$

где $j=1,\dots,\rho'$.

Причем коэффициенты преобразования должны быть весовыми коэффициентами начальных свойств сталей и удовлетворять условию:

$$\sum_{i=1}^{\rho}(a_{ji}^{mpt})^2 = 1; \tag{10}$$

2) статистической независимости друг от друга. То есть равенству 0 коэффициента парной ковариации новых признаков $y_{mpt}^{(b)}$ и $y_{mpt}^{(l)}$:

$$k_{bl} = \mathrm{cov}\,(y_{mpt}^{(b)}, y_{mpt}^{(l)}) = 0,\; b\neq l\,; \tag{11}$$

3) неизменности обобщенной дисперсии или суммы дисперсий всех признаков при переходе от исходных свойств к новым:

$$D = \sum_{j=1}^{\rho'} D_j(y_{mpt}^{(j)}) = \sum_{i=1}^{\rho} D_i\,(\sigma_{mpt}^{(i)}), \tag{12}$$

где $D_i\left(\sigma_{mpt}^{(i)}\right)$ - дисперсия *i*-того начального свойства стали S_{mpt},

$D_j(y_{mpt}^{(j)})$ - дисперсия *j*-того нового признака стали S_{mpt}

Термин дисперсия признака аналогичен понятию энергии сигнала:

$$\xi = \overline{a_j}\overline{a_j} \sim D\,(y_{mpt}^{(j)}), \tag{13}$$

где $\overline{a_j}$ – вектор-столбец матрицы преобразования A_{mpt}.

То есть квадрат отклонения признака от его среднего $D\,(y_{mpt}^{(j)})$ пропорционален энергии его сигнала ξ. Однако необходимо отметить, что энергия сигнала-признака и его информативность не тождественные, но тесно связанные друг с другом понятия.

4) упорядоченности новых признаков по их дисперсиям. Признак стали с максимально возможной дисперсией:

$$D\left(y_{mpt}^{(1)}\right) \geq D\left(y_{mpt}^{(2)}\right) \geq \cdots \geq D(y_{mpt}^{(\rho')}) \qquad (14)$$

должен быть первым в совокупности $y_{mpt}^{(1)} \dots, y_{mpt}^{(j)} \dots, y_{mpt}^{\rho'}$. Таким образом, он является главным новым признаком и обладает наибольшей энергией сигнала. При этом дисперсии можно рассчитывать для различных выборок свойств: одной плавки стали S_{mxpt} (где x – индекс (номер) химического состава выплавленной стали), для выборки разных плавочных химических составов стали S_{mpt} или для заданной марки стали M_{pt} произведенной на разных предприятиях.

Если все перечисленные критерии выполняются, то множество исследуемых новых признаков представляет собой проекции на оси нового признакового пространства образуемого осями $Y^{(1)}, \dots, Y^{(j)}, \dots, Y^{(\rho')}$. Новую систему координатных осей получают, выполняя поворот исходных осей $\Sigma^{(1)}, \dots, \Sigma^{(l)}, \dots, \Sigma^{(\rho)}$, первоначальных свойств $\sigma^{(1)}, \dots, \sigma^{(l)}, \dots, \sigma^{(\rho)}$, с таким расчетом, чтобы проекции, образа марки стали, произведенной на различных предприятиях, информационных выборок, из объединённых разных плавок стали S_{mpt}, изготовленной на одном предприятии, или объектов-точек заданных плавок, на первую ось обладали наиболее возможно высокой для этих множеств дисперсией. Так как, согласно второму критерию, значения признаков на 1-й и 2-й осях должны быть некоррелированными между собой, то есть являться линейно-независимыми друг от друга, то вторая ось должна быть перпендикулярна первой. Проекция выборки на вторую ось должна иметь значительную дисперсию, но меньшую, чем для первой оси. Следующие оси также должны быть не коррелированными друг с другом и упорядоченными в порядке убывания дисперсий новых признаков сталей различных плавок.

Полученные новые признаки и их координатные оси называют системой главных компонент (ГК). Отбрасывая малоинформативные ГК с низкой дисперсией признаков можно существенно сократить пространство описания марочного сортамента сталей металлоизделий.

При этом основная задача анализа главных компонент состоит в объяснении максимальной доли дисперсии наблюдений вдоль главной оси, которая обладает наибольшей значимостью и технологической чувствительностью.

Первая главная компонента, обладающая максимальной дисперсией, будет содержать основную информацию, необходимую для описания образа каждой марки стали или объектов плавок (разброс точек в пространстве новых признаков). Однако, если дисперсии (энергия) и информативность нового признака будут несвязанными друг с другом для рассматриваемой марки стали, то главная ось будет отсутствовать и анализ главных компонент не способствует даже минимальному сокращению (сжатию) результатов наблюдений [4].

Таким образом, применяя метод главных компонент необходимо выполнить преобразование исходной последовательности наблюдаемых переменных-свойств в иную меньшую по объёму последовательность ненаблюдаемых, гипотетических, скрытых переменных-признаков или, что тоже самое -- факторов. При этом новые факторы будут обладать более высокой значимость и значительной технологической чувствительностью.

4. Изложение основного материала исследований и обсуждение результатов. Основной математический метод преобразования исходных направлений пространства описания свойств марок сталей в главные оси нового признакового пространства основан на нахождении собственных чисел и собственных векторов корреляционных или ковариационных матриц первоначальных свойств сталей $\sigma_{mpt}^{(i)}$и их новых признаков $y_{mpt}^{(j)}$.

В матрице коэффициентов преобразования A_{mpt} строка

$$\langle a_{j1}, a_{j2}, \ldots, a_{ji}, \ldots, a_{jp} \rangle = \overline{a_j}^{T} \quad (15)$$

является собственным вектором-строкой корреляционной матрицы R_{mpt} коэффициентов корреляции исходных свойств $\sigma_{mpt}^{(i)}$ и соответствует собственному числу μ_j. Справедливо и обратное – вектор:

$$\overline{a_j} = \langle a_{j1}, a_{j2}, \ldots, a_{ji}, \ldots, a_{jp} \rangle^{T} \quad (16)$$

представляет собой собственный вектор-столбец матрицы R_{mpt}

При этом матрица линейного преобразования A_{mpt} состоит из собственных векторов корреляционной матрицы R_{mpt}:

$$A_{mpt} = \begin{pmatrix} \vec{a}_1^{T} \\ \vec{a}_2^{T} \\ \vdots \\ \vec{a}_j^{T} \\ \vdots \\ \vec{a}_{p'}^{T} \end{pmatrix}, \quad (17)$$

где $\overline{a_j}^{T} = \langle a_{j1}, a_{j2}, \ldots, a_{ji}, \ldots, a_{jp} \rangle$ – матрица-строка.

Её транспонированная матрица представляет собой матрицу строку, элементы которой являются вектор-столбцами:

$$A_{mpt}^{T} = (\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_j \ldots \vec{a}_{p'}) \quad (18)$$

Связь между старой и новой корреляционными матрицами выражается:

$$R_{mpt}\overline{a_j} = \overline{a_j}C_{mpt}, \quad (19)$$

где C_{mpt} - корреляционная матрица новых признаков:

$$C_{mpt} = \left\| \begin{matrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1p'} \\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & c_{2p'} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{j1} & c_{j2} & \cdots & c_{jj} & \cdots & c_{jp'} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{p'1} & c_{p'2} & \cdots & c_{p'j} & \cdots & c_{p'p'} \end{matrix} \right\|. \quad (20)$$

Учитывая, что в рассматриваемой модели линейное преобразование старых координат на новые ортогональное, получаем

$$\vec{a}_j^{T} = \vec{a}_j^{-1}. \quad (21)$$

Таким образом, выражение (19) может быть представлено в виде:

$$\vec{a}_j^{T} R_{mpt}\overline{a_j} = C_{mpt}. \quad (22)$$

Если новые признаки $y_{mpt}^{(j)}$ линейно-независимы, то

$$C_{mpt} = E, \quad (23)$$

и C_{mpt} является единичной матрицей. Поэтому, произведение векторов (15) и (16) даёт коэффициент корреляции:

$$\overrightarrow{r_{jl}} = \overrightarrow{a_j} \cdot \overrightarrow{a_l}^{T}, \quad (24)$$

являющийся элементом матрицы R_{mpt}.

В матричной форме, получаем:

$$C_{mpt} = A_{mpt}^{T} R_{mpt} A_{mpt}. \quad (25)$$

Если новые признаки $y_{mpt}^{(j)}$ полностью линейно независимы, то есть все $c_{jj} = 1$ и $c_{jl} = 0$, то получаем, что новая корреляционная матрица является единичной:

$$C_{mpt} = E = A_{mpt}^{T} R_{mpt} A_{mpt} \quad (26)$$

или

$$R_{mpt} = A_{mpt} A_{mpt}^{T}. \quad (27)$$

Т.е. элементы корреляционной матрицы R_{mpt}:

$$r_{jl} = a_{j1}a_{l1} + a_{j2}a_{l2} + \dots + a_{jp}a_{lp}. \quad (28)$$

Следовательно, зная весовые нагрузки на ГК, можно определить корреляционную матрицу R_{mpt} исходных свойств $\sigma_{mpt}^{(i)}$.

Между коэффициентами корреляции и ковариации существует связь:

$$r_{il} = \frac{k_{il}}{\sqrt{k_{ii}k_{ll}}}. \quad (29)$$

Для определения собственных векторов ковариационной матрицы можно найти экстремум функции Лагранжа [3]:

$$\varphi = \vec{\alpha}_j^{T} K_{mpt} \overrightarrow{\alpha_j} - \lambda(\vec{\alpha}_j^{T}\,\overrightarrow{\alpha_j} - \overrightarrow{\alpha_j} \cdot \overrightarrow{\alpha_j}), \quad (30)$$

где $\overrightarrow{\alpha_j}$ – собственный вектор ковариационной матрицы K_{mpt};

λ – множитель Лагранжа, который в данной задаче является собственным числом матрицы K_{mpt};

$\alpha_j\alpha_j$ = ξ = const – энергетическая характеристика свойств новых признаков.

Экстремум функции Лагранжа находим, приравнивая нулю её частные производные:

$$\frac{\partial\varphi}{\partial\vec{\alpha}^{T}} = K_{mpt}\overrightarrow{\alpha_j} - \lambda\overrightarrow{\alpha_j} = \left(K_{mpt} - \lambda E\right)\overrightarrow{\alpha_j} = 0, \quad (31)$$

где E – единичная матрица.

Таким образом, получаем систему уравнений в матричной форме:

$$K_{mpt}\vec{\alpha}_j = \lambda E \vec{\alpha}_j. \quad (32)$$

В полученной системе уравнений $\overrightarrow{\alpha_j}$ и λ – неизвестные, причем λ может иметь несколько значений.

Произведение собственных множителей λ на единичную матрицу E дает диагональную матрицу, состоящую из собственных чисел:

$$\lambda E = L = \begin{Vmatrix} \lambda_1 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_j & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & \dots & \lambda_{\rho'} \end{Vmatrix}, \qquad (33)$$

где λ_j-- собственные значения симметричной матрицы K_{mpt}. При этом значение каждого собственного числа является корнем определителя системы уравнений (31). Преобразуя уравнение (32) в выражение

$$K_{mpt}\vec{\alpha}_j = L\vec{\alpha}_j, \qquad (34)$$

находим матрицу L:

$$\vec{\alpha}_j^{-1} K_{mpt}\vec{\alpha}_j = L. \qquad (35)$$

Матрица A_{mpt} состоит из собственных векторов матрицы K_{mpt}, поэтому полученное векторное уравнение можно записать в матричной форме:

$$A_{mpt}^T K_{mpt} A_{mpt} = L. \qquad (36)$$

Так как при повороте координатных осей $\Sigma^{(i)}$ новые оси $Y^{(J)}$ остаются ортогональными, то

$$A_{mpt}^T = A_{mpt}^{-1}. \qquad (37)$$

Следовательно,

$$L = A_{mpt}^{-1} K_{mpt} A_{mpt}. \qquad (38)$$

Таким образом, получаем равенство

$$K_{mpt} A_{mpt} = A_{mpt} L, \qquad (39)$$

из которого следует, что

$$\vec{\alpha}_j^T K_{mpt}\vec{\alpha}_j = L. \qquad (40)$$

При выполнении условия экстремума (31):

$$K_{mpt}\vec{\alpha}_j - \lambda E\vec{\alpha}_j = 0. \qquad (41)$$

у новых признаков могут быть наибольшие дисперсии.
Полученная, система уравнений

$$(K_{mpt} - \lambda E)\vec{\alpha}_j = 0 \qquad (42)$$

имеет решения, если

$$\mathrm{Det}\,(K_{mpt} - \lambda E) = 0. \qquad (43)$$

То есть определитель матрицы коэффициентов этих уравнений представляет собой характеристическое уравнение, которое решают относительно λ с помощью численных методов [5]. Решениями этого уравнения будут собственные числа λ_j ($j=1,\dots,\rho'$), причем $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_j \geq \cdots \geq \lambda_\rho \geq 0$.

Однако дисперсии новых признаков, являющихся линейной комбинацией прежних свойств, [6] могут быть аналогично выражены через ковариационную матрицу первоначальных свойств:

$$\sigma_{y_j}^2 = D(y^{(J)}) = \overline{\alpha}_j^T K_{mpt} \overline{\alpha}_j. \qquad (44)$$

где K_{mpt} – ковариационная матрица начальных свойств,

$\overline{\alpha}_j$- вектор коэффициентов преобразования старой ковариационной матрицы в новую Cv_{mpt}.
Преобразуем приведенное равенство (44) в выражение

$$K_{mpt}\ \overrightarrow{\alpha}_j=\overrightarrow{\alpha}_j\ D\ (y^{(J)}). \tag{45}$$

Таким образом, получаем, что дисперсии новых признаков являются собственными числами ковариационной матрицы K_{mpt} исходных свойств. Следовательно, дисперсии $D_j\ (y^{(J)})$ являются корнями характеристического уравнения:

$$\mathrm{Det}\ (K_{mpt} - D\ (y^{(J)})E) = 0. \tag{46}$$

Учитывая, что дисперсии равны диагональным элементам ковариационной матрицы новых признаков Cv_{mpt}, получаем:

$$D(y^{(J)})_j = cv_{jj}, \tag{47}$$

где cv_{jj} - диагональные коэффициенты ковариационной матрицы новых признаков $y^{(J)}$, приходим к выражению:

$$cv_{jj} = \vec{a}_j^{\,T} K_{mpt}\vec{a}_j. \tag{48}$$

Таким образом, могут быть вычислены диагональные элементы новой ковариационной матрицы.

Подставляя в выражение (45) найденные значения корней – дисперсий получим систему уравнений:

$$K_{mpt}\vec{a}_j - D(y^{(J)})_j\vec{a}_j = 0. \tag{49}$$

Совокупность решений α_{ji} этой системы образует вектора нагрузок на ГК:

$$\vec{a}_j^{\,*} = (\alpha_{j1}^{*}, \dots, \alpha_{ji}^{*}, \dots, \alpha_{jp}^{*}), \tag{50}$$

где α_{ji}^{*} – решения, т.е. корни системы уравнений.
Причем, суммы квадратов корней

$$\sum_{i=1}^{p} \alpha_{ji}^{*2} = D(y^{(J)}), \tag{51}$$

где $i = 1, \dots, p$ – номера исходных свойств,
представляют собой дисперсии. Поэтому вектор решения для 1-й ГК будет иметь максимальную сумму квадратов нагрузок

$$S_1 = \sum_{i=1}^{p} \alpha_{1i}^{*2} = max, \tag{52}$$

т.е. k_{jl}=$\alpha_{j1} \cdot \alpha_{l1}$. Сумма квадратов нагрузок 2-й ГК должна составлять максимум оставшейся дисперсии.

Аналогично можно найти собственные числа корреляционной матрицы R_{mpt}. Для этого матричное уравнение

$$R_{mpt}\vec{a}_j = \mu E\vec{a}_j \tag{53}$$

необходимо представить в виде системы ρ´ однородных уравнений:

$$(1-\mu)a_{11} + r_{12}a_{21} + \cdots + r_{1\rho'}a_{\rho'1} = 0$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$r_{21}a_{11} + (1-\mu)a_{21} + \cdots + r_{2\rho'}a_{\rho'1} = 0 \quad (54)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$r_{\rho'1}a_{11} + r_{\rho'2}a_{21} + \cdots + (1-\mu)a_{\rho'1} = 0,$$

Если корни μ_j характеристического уравнения

$$\begin{vmatrix} (1-\mu) & \cdots & r_{12} & \cdots & r_{1\rho'} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{21} & \cdots & (1-\mu) & \cdots & r_{2\rho'} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{\rho'1} & \cdots & r_{\rho'2} & \cdots & (1-\mu) \end{vmatrix} = 0, \quad (55)$$

найдены, то подставляя их в (54), получаем уравнение относительно $\vec{a}_j$:

$$R_{mpt}\vec{a}_j - \mu_j\vec{a}_j = 0 \quad (56)$$

Эту систему уравнений решаем относительно неизвестных коэффициентов преобразования a_{ji} и получаем их значения, которые являются корнями этих уравнений. Совокупность решений образует вектора оптимальных нагрузок на ГК:

$$\vec{a}_j^* = \left(a_{j1}^*, \ldots, a_{ji}^*, \ldots, a_{j\rho}^*\right), \quad (57)$$

где a_{ji}– решения, т.е. корни системы уравнений.

Таким образом, компоненты собственных векторов $\vec{a}_j^*$ матрицы R_{mpt} являются решениями системы

$$(R_{mpt} - \mu_j \mathrm{E})\vec{a}_j = 0. \quad (58)$$

Значения координат векторов решений $\vec{a}_j^*$ и соответствующие им собственные числа μ_j, определяют максимум функции Лагранжа, которому соответствуют наибольшие дисперсии в направлении ГК.

После подстановки $\vec{a}_j^*$ и μ_j в уравнения (53), получаем систему равенств:

$$R_{mpt}\vec{a}_j^* = \mu_j\vec{a}_j^*. \quad (59)$$

Учитывая, что в условиях рассматриваемой задачи

$$\mu_j \sim \lambda_j = \mathrm{D}\,(y^{(J)}), \quad (60)$$

получаем:

$$\vec{a}_j^{*T} R_{mpt}\vec{a}_j^* \sim D(y^{(J)}). \quad (61)$$

При этом учитывая, что матрицы R_{mpt} и K_{mpt} диагональные, то есть $r_{il} = 0$ и $k_{il} = 0$, можем для их строк записать:

$$a_{ji}^* r_{ii} a_{ij}^* = \mu_j. \quad (62)$$

Так как в условиях рассматриваемой задачи $r_{ii} = 1$, то

$$a_{ji}^* r_{ii} a_{ij}^* = \mu_j, \quad (63)$$

и

$$\propto_{ji}^* k_{ii} \propto_{ij}^* = D(y^{(J)})_j. \quad (64)$$

Учитывая, что матрицы A_{mpt} и $\mathbf{A}_{mpt}$ симметричные, то есть $a_{ji}^{*} = a_{ij}^{*}$ и $\alpha_{ji}^{*}=\alpha_{ij}^{*}$, получаем

$$a_{ji}^{*2} = \mu_j, \tag{65}$$

$$\alpha_{ji}^{*2}\, k_{ii} = D(y^{(J)})_j = \lambda_j. \tag{66}$$

Подставляя полученные уравнения в формулу связи коэффициентов корреляции и ковариации (29) находим выражение:

$$\frac{a_{ij}^{*2}}{\mu_j} = \frac{\alpha_{ij}^{*2}}{D(y^{J})_j}\sqrt{D\left(\sigma_{mpt}^{(i)}\right)^2}, \tag{67}$$

где $D\left(\sigma_{mpt}^{(i)}\right) = k_{ii} = \sigma_{ii}^2 = D\left(\sigma_{mpt}^{(l)}\right) = k_{ll} = \sigma_{ll}^2$ – дисперсии начальных свойств;

σ_{ii}^2 – квадрат среднеквадратичного отклонения.

Для ковариационной матрицы сумма собственных чисел λ_j равна числу переменных. Поэтому, разделив *j*-е собственное число на *m* (число переменных), можем получить долю дисперсии, соответствующую данному направлению или компоненте [7], аналогичное верно и для корреляционной матрицы:

$$D = \frac{\lambda_j}{m} = \frac{\mu_j}{n} = \frac{D(y^{(J)})}{\sum_{j=1}^{\rho'} D(y^{(J)})} \ . \tag{68}$$

где *n* – сумма собственных чисел μ_j.

Отсюда следует связь для собственных чисел ковариационной и корреляционной матриц

$$\mu_j = \lambda_j \frac{m}{n}. \tag{69}$$

Таким образом,

$$a_{ji}^{*2} = \lambda_j \frac{m}{n}. \tag{70}$$

Следовательно,

$$a_{ji}^{*2} = \alpha_{ji}^{*2}\, k_{ii} \frac{m}{n} \tag{71}$$

или

$$a_{ji}^{*2} = \alpha_{ij}^{*2} D(\sigma^{(i)}) \frac{m}{n}, \tag{72}$$

то есть находим уравнение связи коэффициентов a_{ji}^{*} и α_{ji}^{*}:

$$a_{ji}^{*} = \alpha_{ji}^{*} \sqrt{D(\sigma^{(i)}) \frac{m}{n}}. \tag{73}$$

Задача определения новых признаков сводится к нахождению элементов матрицы нагрузок ГК a_{ji}. Для их определения воспользуемся равенством

$$\sum_{i=1}^{\rho} \alpha_{ji}^{*2} = D(y^{(J)}). \tag{74}$$

Преобразуя полученное уравнение (73), находим значение коэффициентов преобразования матрицы A_{mpt}:

$$a_{ji}^{*} = \frac{\alpha_{ji}^{*}\sqrt{\lambda_j}}{\sqrt{\alpha_{j1}^{*2}+\alpha_{j2}^{*2}+\cdots+\alpha_{j\rho}^{*2}}}\sqrt{D\left(\sigma_{mpt}^{(i)}\right)}, \tag{75}$$

которые вычисляются по значениям координат α_{ji}^{*} собственных векторов $\vec{\alpha}_j$. Причём

$$\sum_{i=1}^{\rho}\alpha_{1i}^{*2} > \sum_{i=1}^{\rho}\alpha_{2i}^{*2} > \cdots > \sum_{i=1}^{\rho}\alpha_{ji}^{*2} > \cdots > \sum_{i=1}^{\rho}\alpha_{\rho i}^{*2}. \quad (76)$$

Если исходные свойства $\sigma_{mpt}^{(i)}$ предварительно стандартизированы или нормированы:

$$\overline{\overline{\sigma_{mpt}^{(i)}}} = \frac{\sigma_{mpt}^{(i)} - \overline{\sigma_{mpt}^{(i)}}}{D\left(\sigma_{mpt}^{(i)}\right)}, \quad (77)$$

где $\overline{\sigma_{mpt}^{(i)}}$- среднеарифметическая величина в выборке i-го свойства,

то есть приведены к безразмерным величинам, то коэффициенты корреляции и ковариации равны

$$r_{il} = k_{il}. \quad (78)$$

Следовательно, равны корреляционная и ковариационная матрицы

$$R_{mpt} = K_{mpt}. \quad (79)$$

При этом функция Лагранжа после выполненных преобразований с учётом (22), (24) может быть записана в упрощенном виде:

$$\varphi = Cv_{mpt} - \mu(\overline{r_{jj}} - \zeta). \quad (80)$$

где $\zeta = \vec{a}_j\vec{a}_j$.

При этом искомое значение нормированных коэффициентов преобразования матрицы A_{mpt}:

$$a_{ji}^{*} = \frac{\alpha_{ji}^{*}\sqrt{\lambda_j}}{\sqrt{\alpha_{j1}^{*2} + \alpha_{j2}^{*2} + \cdots + \alpha_{j\rho}^{*2}}}. \quad (81)$$

Подставляя вычисленные коэффициенты a_{ji}^{*} и соответствующие им значения старых признаков $\sigma_{mpt}^{(i)}$ в уравнения (7) рассчитываем новые признаки $y_{mpt}^{(j)}$. Для каждого, из которых можно определить величину доли общей дисперсии (68).

Снижение размерности признакового пространства при помощи ГК представляет собой отбрасывание тех признаковых осей, на которые падает минимум дисперсии (энергии), то есть наименее информативных.

Необходимо отметить, что различные выборки свойств марки стали полученные из совокупности плавок или на разных предприятиях, имеющие одинаковые ковариационные матрицы исходных характеристик сталей, имеют одни и те же ГК.

При переходе из пространства исходных свойств в новое признаковое пространство меньшей размерности ρ’ < ρ величина искажений признаков сталей зависит только от последних ρ - ρ’ собственных чисел λ_j и тем меньше, чем меньше эти числа. В новом признаковом пространстве, меньшей размерности, натянутом на первые ρ’ ГК, наименее искажается сумма квадратов евклидовых расстояний между возможными парами информационных объектов. Это свойство является особенно важным при кластеризации марочного сортамента металлопроката и взаимной замене сталей в границах одного кластера.

В первую очередь могут быть отброшены ГК с нулевым вкладом в общую дисперсию. Решение об отбрасывании остальных следует применять исходя из каких-либо дополнительных критериев.

При несовпадении критериев энергии и информативности ГК, решение об отбрасывании ГК может быть принято из оценки функции плотности распределения для проекции выборки свойств сталей на каждую ГК. Тогда из анализа полученного набора одномерных функций можно сделать вывод об информативной ценности каждой ГК.

Таким образом, алгоритм, реализующий метод ГК для снижения размерности пространства описания марок сталей различных металлоизделий, может иметь следующую структуру:

1) каждому виду формы заданного типоразмера металлоизделия определенной марки стали ставят в соответствие системы индексов. Таким образом, индексируют все рассматриваемые стали. То есть они получают систему номеров. Совокупности данных о марках сталей трансформируют в образы, считая номера химических составов в выборке объектами информационного моделирования [1];

2) расчёт ковариационной и корреляционной матриц начальных свойств сталей, вычисление их собственных векторов и собственных чисел [4];

3) упорядочивание собственных векторов в порядке убывания соответствующих собственных чисел [4] по критериям их максимума или минимума;

4) вычисление матрицы нагрузок на ГК путем нормирования собственных векторов (81);

5) выбор из матрицы линейного преобразования A_{mpt} двух строк весовых коэффициентов, соответствующих наибольшим значениям обобщенной дисперсии в процентах;

6) определение исходных свойств, для которых в матрице линейного преобразования или в уравнении перехода от старых свойств к новым коэффициенты преобразования > ± 0,5;

7) вычисление процента обобщенной дисперсии (68), приходящейся на каждую ГК [4];

8) предварительное вычисление матрицы ковариаций новых признаков;

9) расчет новых признаков каждой марки стали, множества информационных объекта каждому, из которых соответствует плавочный химический состав стали или для определённого плавочного химического состава, которые задают соответствующими индексами *mpt* на основе полученных уравнений 1-ой и 2-ой главных компонент;

10) построение проекции образов марок сталей, плавочных выборок или заданных объектов, которым соответствуют определённые составы сталей на плоскости 1-ой и 2-ой ГК в новом признаковом пространстве. Формирование групп и кластеров сталей. Каждая точка проекции обозначается индексами *mpt* объекта кластеризации.

Выводы. Разработанная модель представления показателей качества сталей позволяет определить главную ось заданную линией, для которой сумма средних квадратов расстояний вдоль нормали от неё до всевозможных точек признакового пространства минимальна. Таким образом при нахождении главной компоненты мы минимизируем расстояния от точки центра образа стали или совокупности информационных объектов до оси (по перпендикуляру к главной оси). Этот принцип заложен в методе наименьших квадратов ортогональной регрессии [8], что и составляет отличие от нахождения линии регрессии методом наименьших квадратов расстояний измеряемых по линии вдоль одной из исходных осей, на которой отложены значения определенного свойства стали.

5. Перспективы дальнейших исследований. Определив в пространстве свойств сталей среднюю сумму квадратов отклонений от точек – информационных объектов, образующих образ марки сталей, или от центров облаков информационных объектов до некоторой прямой линии, являющейся при минимуме суммы квадратов - главной компонентой и направленной в сторону положительного изменения качества сталей, можно произвести ранжирование сталей и сплавов или их химических составов по степени близости друг к другу. Таким образом, в пространстве 1-х ГК можно определить комплексные показатели качества металлопроката.

ИСПОЛЬЗОВАНННАЯ ЛИТЕРАТУРА

1. Юшкевич О.П. Разработка структуры данных марочного сортамента металлопроката//Строительство, материаловедение, машиностроение/ Сб. научн. трудов. Вып.22, ч.2, - Дн-ск, ПГАСА.- 2003. – 320 с. (в обл.).- С.43-49.

2. Юшкевич О.П. Анализ независимости признаков вектора описания марочного сортамента металлопроката//Металознавство та термічна обробка металів/Наук. та інформ. журнал, ПГАСА. - Днепропетросвк, 2006. - № 3.- С.58-66.

3.Пересада В.П. Автоматическое распознавание образов//Библиотека по автоматике. Вып.392.-Л.: Энергия, 1970.- 92 с.- с рис.

4. Александров В.В., Горский Н.Д. Алгоритмы и программы структурного метода обработки данных . Л.: Наука ,1983. - с. 208.

5. Иберла К. Факторный анализ/Пер.с нем. В.М. Ивановой; Предисл. А.М. Дуброва.- М.: Статистика, 1980.- 398 с., - ил. - (Математико-статистические методы за рубежом).

6.Андерсон Т. Введение в многомерный статистический анализ. Физматгиз, 1963.

7. Факторный, дискриминантный и кластерный анализ: Пер.с анг./ Дж.-О. Ким, Ч.У. Мюллер, У.Р. Кленка и др., Под ред. И.С. Енюкова.- М.: Финансы и статистика, 1989.- 215 с.: ил.

8. Айвазян С.А., Бежаев З.И., Староверов О.В. Классификация многомерных наблюдений. М.: Статистика, 1974.