

# Факторный анализ

# Пример факторного анализа

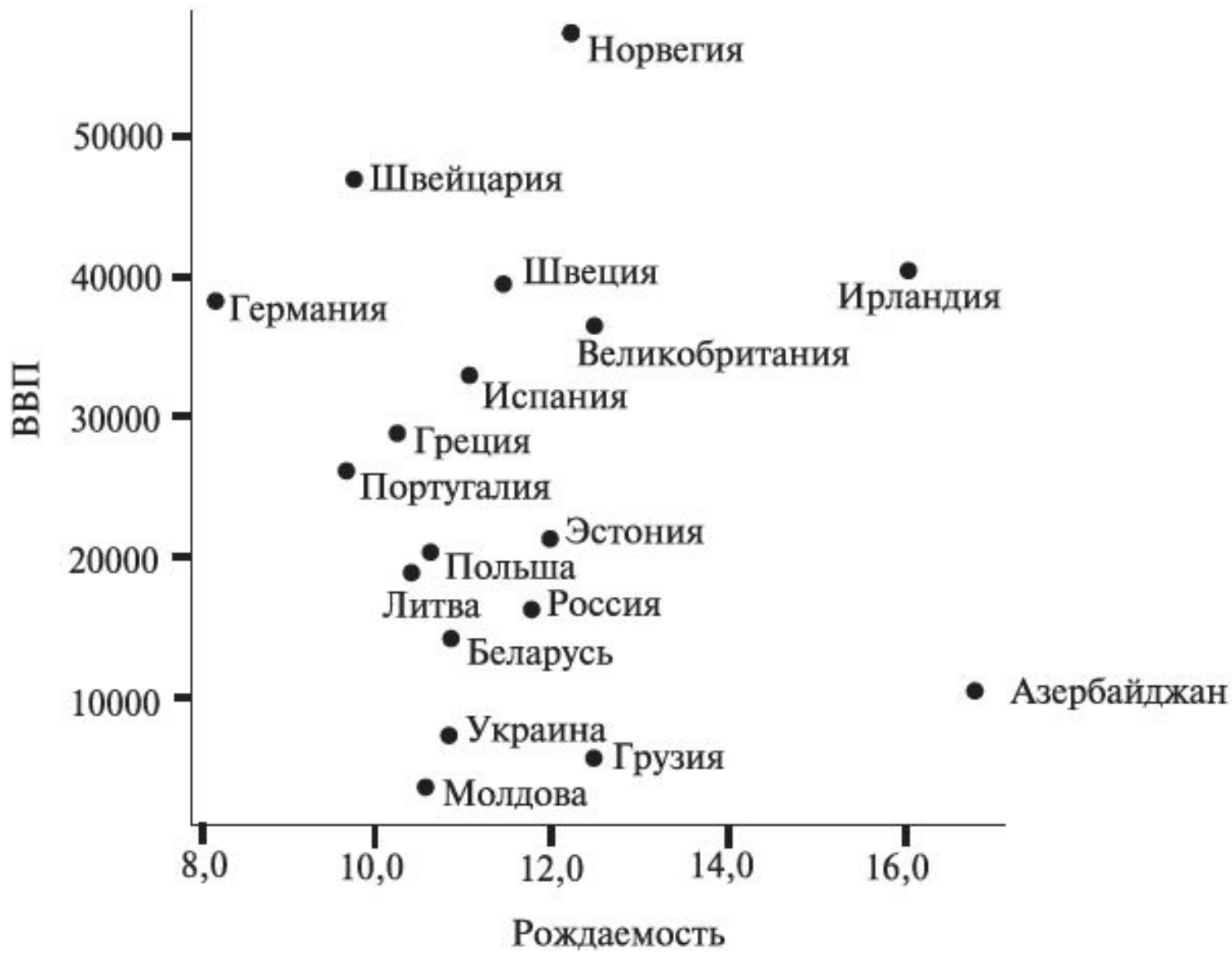
**Пример 1.1. Влияние внутреннего валового продукта (ВВП) на социально-демографические показатели (европейские страны, 2008 г.)**

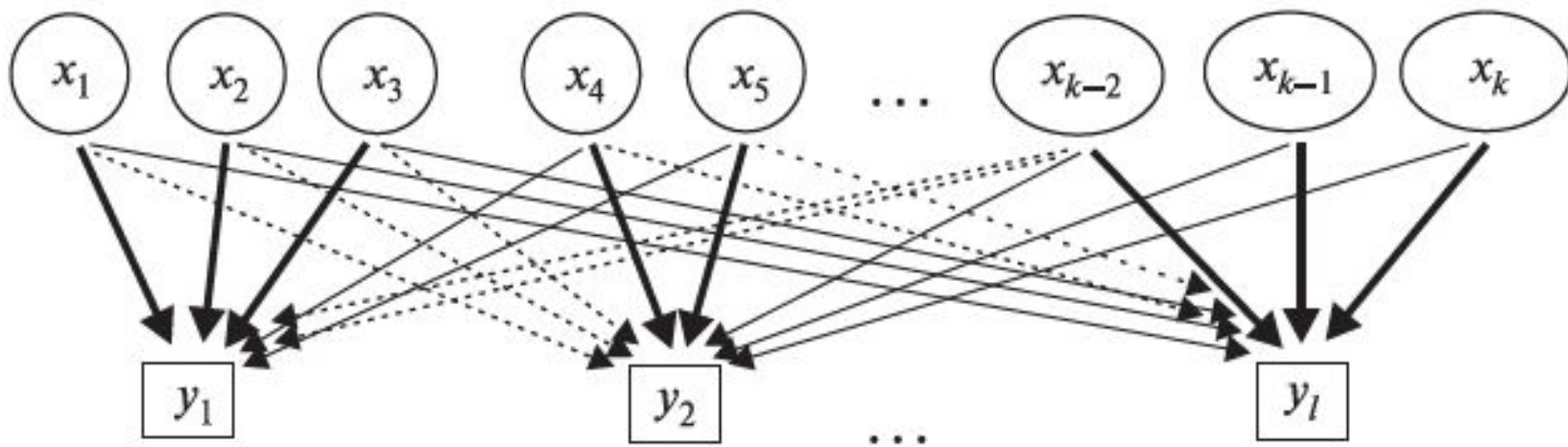
*Переменные:*

1. Медианный возраст населения страны.
2. Рождаемость (число родившихся на 1000 жителей).
3. Смертность (число умерших на 1000 жителей).
4. Естественный прирост (разность рождаемости и смертности на 1000 жителей).
5. Детская смертность (до 1 года на 1000 живорожденных).
6. Ожидаемая продолжительность жизни мужчин при рождении.
7. Ожидаемая продолжительность жизни женщин при рождении.
8. Валовой внутренний продукт на душу населения.

1. **Испания**
2. **Италия**
3. **Латвия**
4. **Литва**
5. **Македония**
6. **Молдова**
7. **Нидерланды**
8. **Норвегия**
9. **Польша**
10. **Португалия**
11. **Россия**
12. **Румыния**
13. **Сербия**
14. **Словакия**
15. **Словения**
16. **Украина**
17. **Финляндия**
18. **Франция**
19. **Хорватия**
20. **Чехия**
21. **Швейцария**
22. **Швеция**
23. **Эстония**







*Рис. 2.6.* Модель метода главных компонент:

— значительный вклад; — вклад средней величины;  
 --- незначительный вклад

## Собственные значения матрицы корреляций

Компонента	Объясненная дисперсия		
	$\lambda_i$	% дисперсии	накопленный %
$y_1$	3,82	54,60	54,60
$y_2$	2,32	33,12	87,72
$y_3$	0,45	6,37	94,09
$y_4$	0,19	2,70	96,79
$y_5$	0,15	2,15	98,94
$y_6$	0,06	0,83	99,77
$y_7$	0,02	0,23	100,00
Сумма	7,00	100,00	

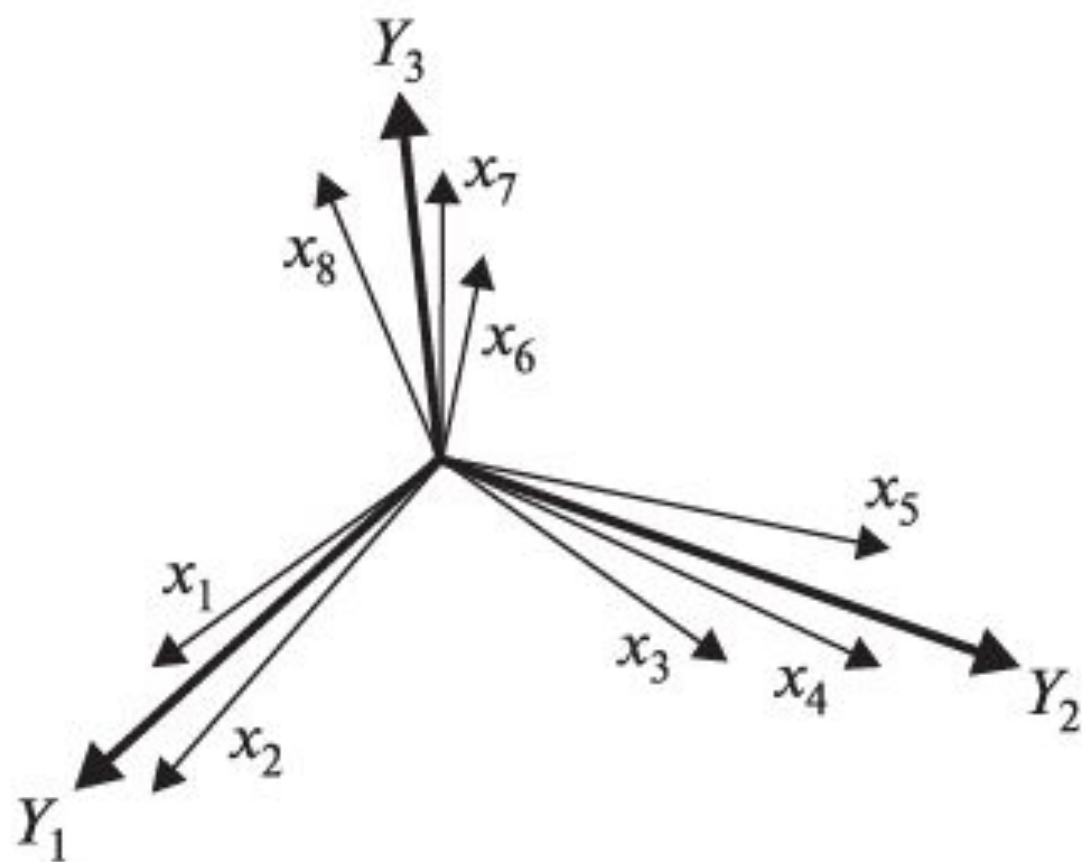


Рис. 2.3. Снижение размерности пространства переменных



# Последовательность ФА

- Выбор исходных данных
- Вычисление корреляционной матрицы
- АГК и оценка числа факторов по графику собственных значений
- Выбор и применение метода ФА
- Вращение факторов
- Интерпретация факторов, принятие решения о качестве факторной структуры
- (Вычисление факторных оценок)

# Анализ главных компонент (МГК) и ФА

МГК в матричной форме:

$$R = A \cdot A',$$

кол-во компонент = числу переменных

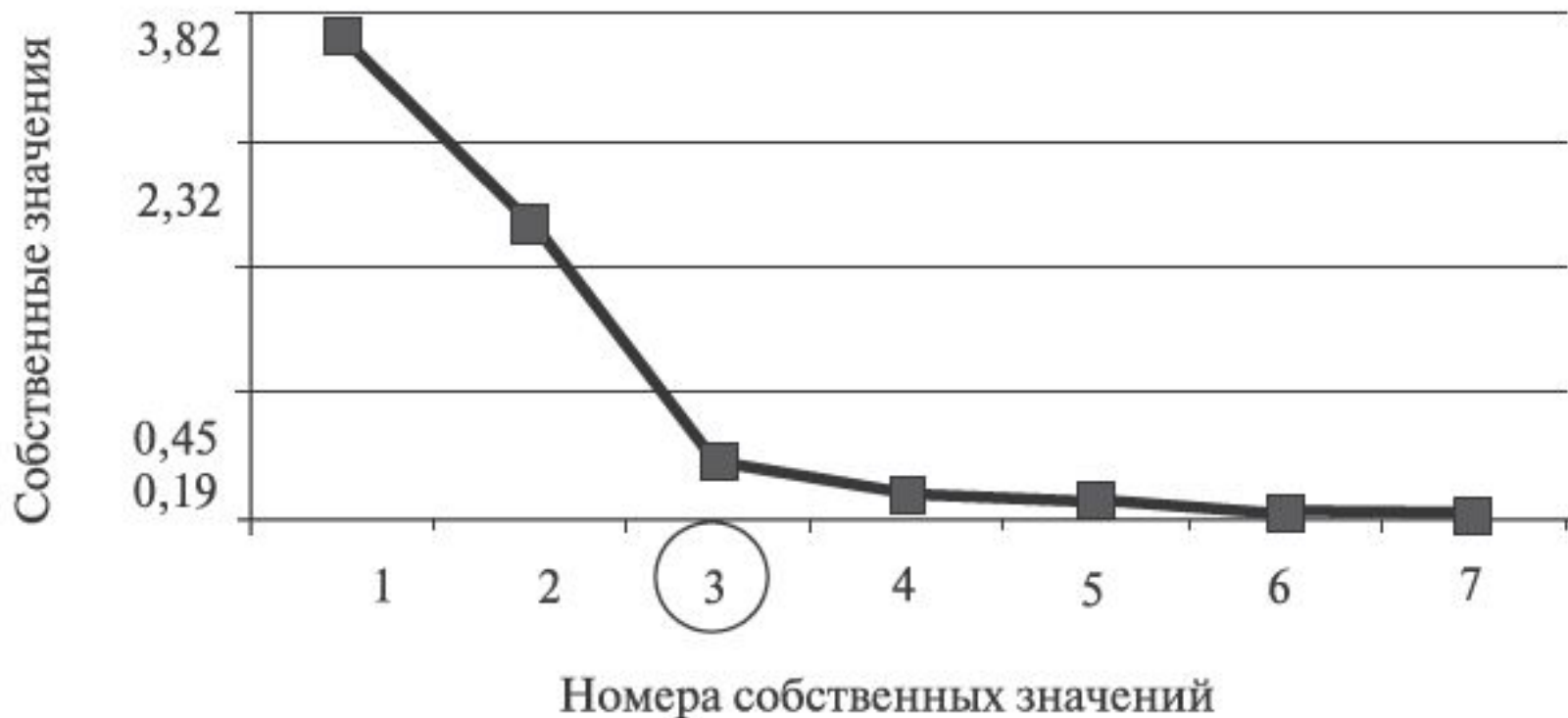
**Факторный анализ:**

$$\hat{R} = A \cdot A' \text{ при условии } \hat{R} \rightarrow R$$

кол-во факторов ( $M$ ) существенно меньше кол-ва переменных ( $P$ )

# График собственных значений

Критерий числа факторов:  
Р.Кеттелла (scree-test), «каменистой осыпи»







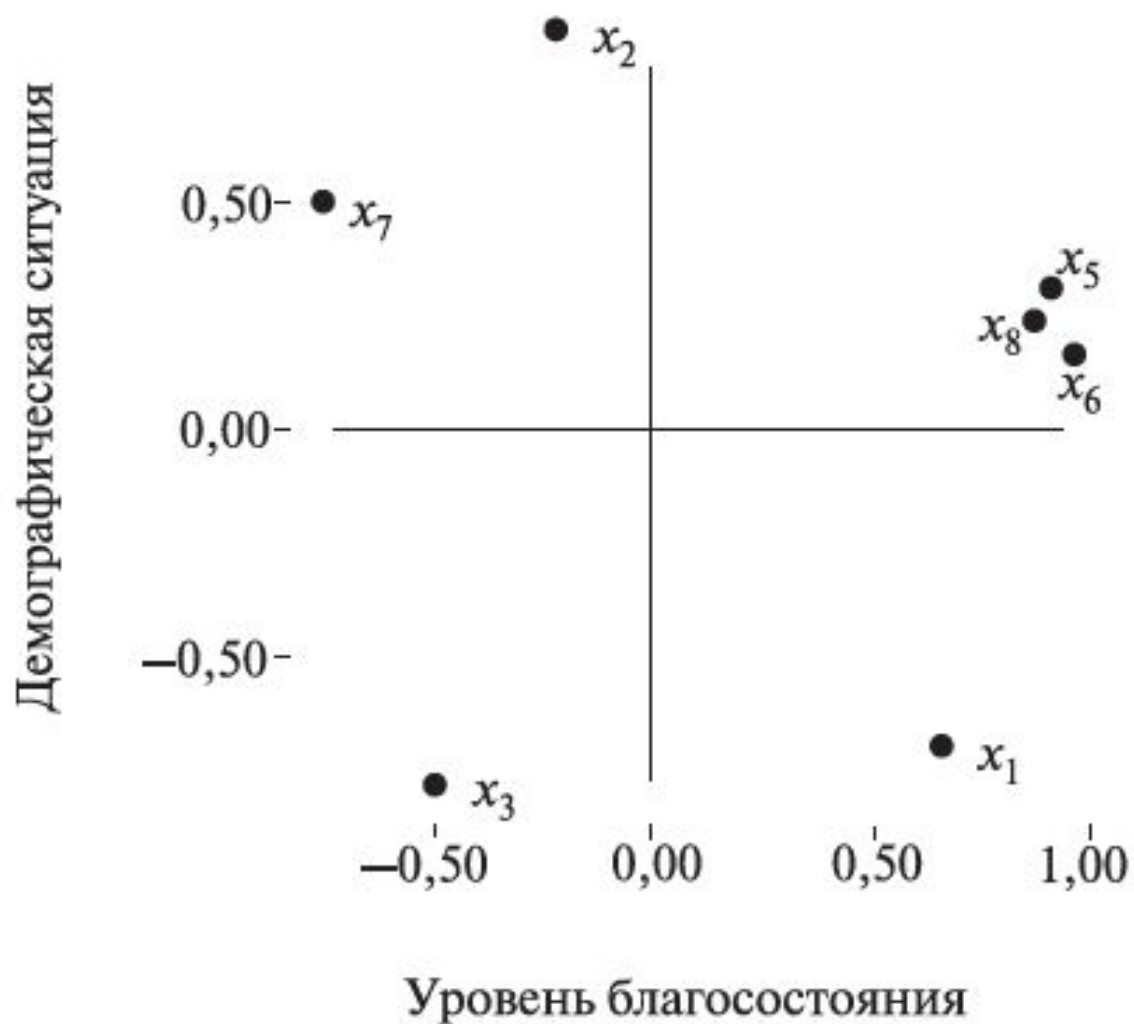


Рис. 2.10. График факторных нагрузок

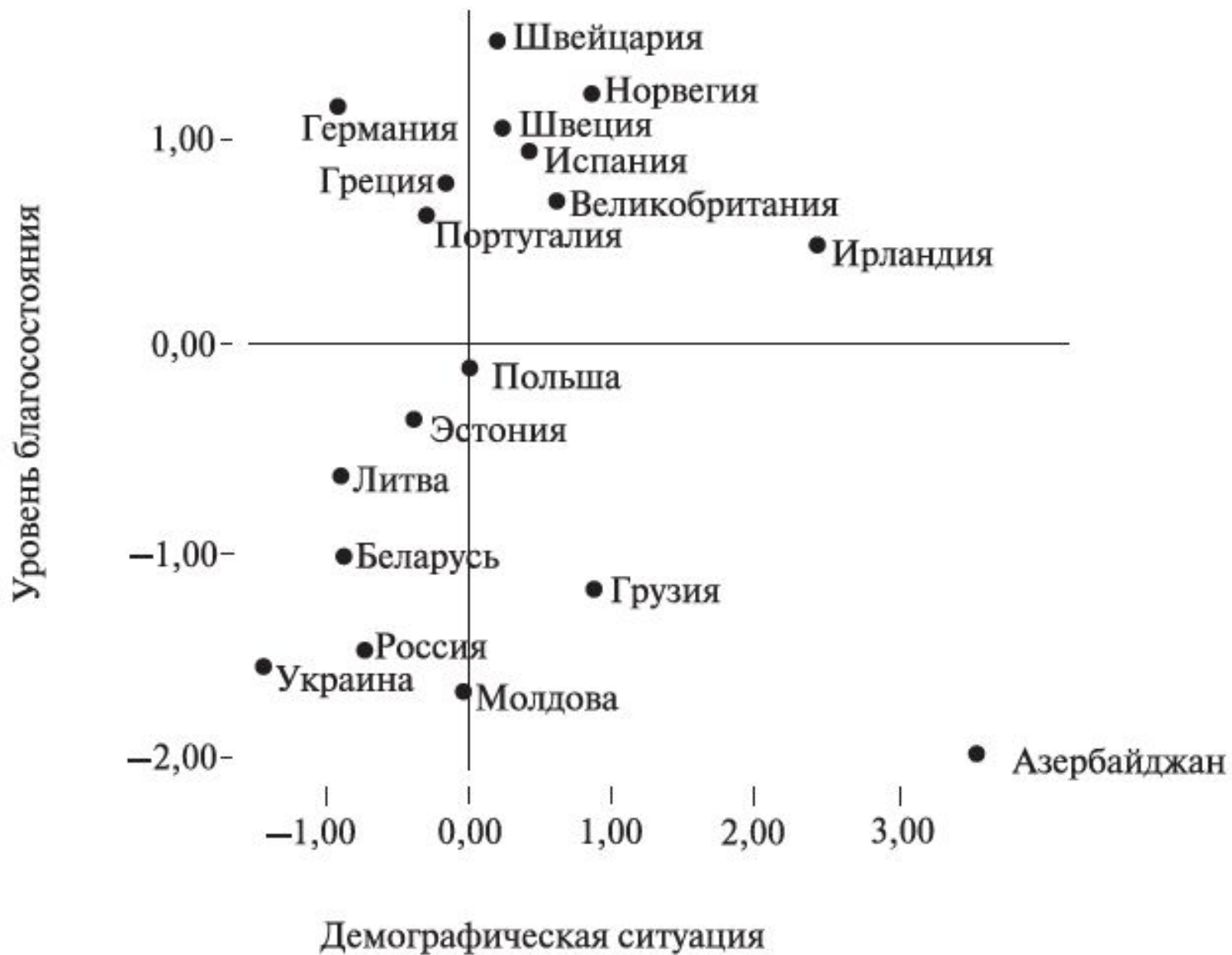


Рис. 2.11. Европейские страны в пространстве главных компонент

# Проблема оценки значений факторов

Факторные оценки – новые переменные, значения факторов для объектов (случаев). Переход от  $P$  исходных переменных к  $M$  новым переменным (факторам) с минимальными потерями исходной

$$f_{ik} = \sum_{j=1}^P \beta_{i,j} x_{jk} = \beta_{1i} x_{1k} + \beta_{2i} x_{2k} + \dots + \beta_{Pi} x_{Pk}$$



# Классические проблемы ФА

- Проблема числа факторов
- Проблема общности и метода ФА
- Проблема вращения и интерпретации
- Проблема факторных оценок

# Пример интерпретации факторов

**Фактор 1** имеет наибольший вес или наибольшую информативность (27%). Его положительный полюс определяется положительными полюсами переменных 4 (независимый), 8 (спокойный), 10 (невозмутимый) и отрицательным полюсом переменной 7 (расслабленный). Отрицательный полюс этого фактора определяется про-

