

名古屋市の人口データに関する主成分分析

氏名：高田恒輝 学生番号：34714086

2024年5月16日

1 6変数による主成分分析

元のデータの説明変数群を x 、主成分得点の変数群を z と置くと主成分得点は以下のようになる。

$$z_1 = -0.490447 x_1 - 0.460220 x_2 - 0.490588 x_3 - 0.469303 x_4 - 0.254028 x_5 - 0.149019 x_6$$

$$z_2 = -0.120833 x_1 - 0.239175 x_2 - 0.092701 x_3 - 0.110593 x_4 + 0.631692 x_5 + 0.712977 x_6$$
 また、得

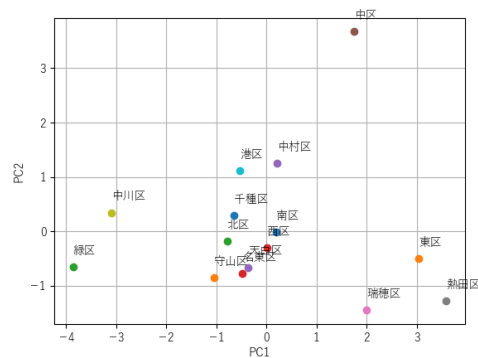


図1 6変数による主成分得点の散布図

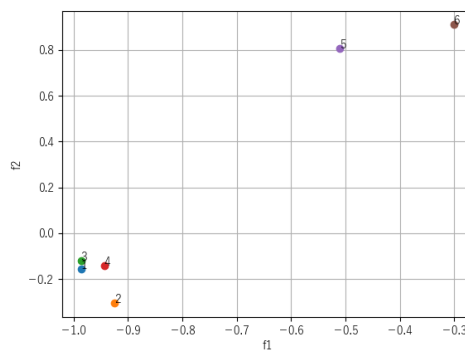


図2 6変数による因子負荷量の散布図

られた主成分得点から PC1 を横軸、PC2 を縦軸にとった散布図と因子負荷量をプロットした図を上記に示す。因子負荷量の図 2 から PC2 は元データの外国人人口と昼間人口と正の相関があり、PC1 は元データの総人口、15 歳以下の人口、65 歳以上の人口、15 から 65 歳までの人口と負の相関があることが確認できた。

私は以上のことから主成分得点 PC1 を”総人口の少なさ”、PC2 を”外国人を含む労働者人口の多さ”と言い換えられると考えた。

2 5 変数による主成分分析

元のデータの説明変数群を x 、主成分得点の変数群を z と置くと主成分得点は以下ようになる。

$$z_1 = -0.425493 x_1 + 0.460805 x_2 - 0.177588 x_3 + 0.520638 x_4 + 0.551374 x_5$$

$$z_2 = 0.488681 x_1 + 0.399237 x_2 - 0.761623 x_3 - 0.110808 x_4 - 0.097220 x_5$$

また 6 変数の時同様に、

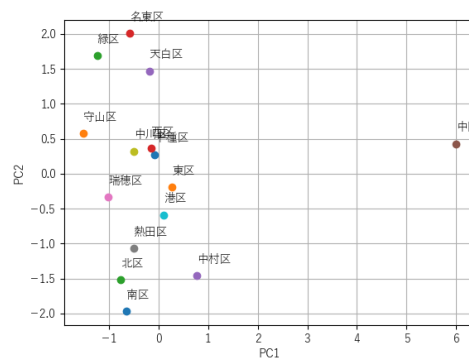


図 3 5 変数による主成分得点の散布図

得られた主成分得点から PC1 を横軸、PC2 を縦軸にとった散布図と因子負荷量をプロットした図を上記に示す。因子負荷量の図 2 から PC1 は 15 歳から 65 歳までの人口の割合、外国人人口の割合、昼間人口の割合と強い正の相関があり、15 歳以下の人口とは強い負の相関があることがわか

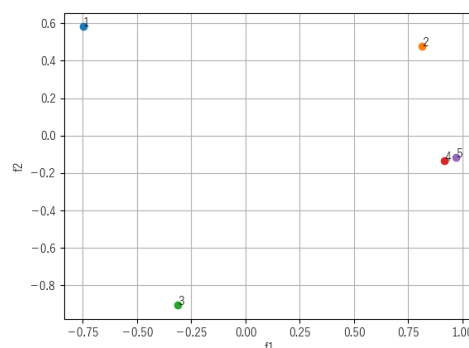


図 4 5 変数による因子負荷量の散布図

る。PC2 は 15 歳以下の人口、65 歳以上の人口と正の相関があり、15 歳から 65 歳の人口と強い負の相関があることがわかる。

私は以上のことから主成分得点 PC1 を”労働人口が多ければ正、通学している人口が多ければ負を表す値”を示していて、PC2 を”生産人口の少なさ”を示すという風に考えました。

3 6 変数による主成分分析と 5 変数による主成分分析の比較

上記の 2 点の主成分分析の結果を見ると、5 変数による主成分分析ではどちらの主成分得点も働き手に関しての情報しか得ることができないのに対して、6 変数による主成分分析では総人口についてと労働者について複数の観点からの情報を得られているため、6 変数による主成分分析のほうが良い結果が得られるという風に考える。