

# システム設計における論点の共分散構造分析

植田 和則 田中 克明 赤石 美奈 堀 浩一  
 東京大学

## 1 はじめに

システムアプローチに基づく設計プロジェクトは、設計対象をサブシステムからなるシステムと見做すことで、設計過程を分割する。それによって困難さは分割されるが、設計者が複数となり個々の関与するサブシステムが限定される条件下では、様々な視点からのシステムのモデルが並存する事となる。

ここで問題となるのは、設計者個人にとって全体が把握出来ない事ではなく、個人が感知していない領域からの影響が把握出来ない場合である。こうした情報の不完備が、効率を損ね、さらに大きな問題の温床となる。すなわち困難さを分割する過程で、新たな困難さが生み出されているのである。

この新たに生じる問題は、人間の認識あるいは知識の在り方に依拠するものだ。「個人的知識が単に主観的であることを逃れさせるのは、その全体的構造への自己投出の行為である」[Polanyi 1958]と言われるように、この問題を克服するためには、個人の知識が組織の知へと表出される必要がある。

一方で、設計者個人が持つ知識の一部は、プロジェクトの過程で文書として記録され蓄積されており、それらはシステムアプローチに基づくプロジェクトの性質として、複数領域に亘る膨大な量の情報となっているため、全てを組織の知とするにはまた新たな困難がある。

そのため本研究は、設計者が個人で、あるいは組織として持っている知識の中でも、特に相関に対する知識<sup>1</sup>を対象とし、設計プロジェクトの過程で蓄積される文書情報から、設計過程においてもつ設計要素の相関を分析する手法を提案する。

## 2 システムの概要

本システムは、組織が蓄積した情報として、プロジェクトの過程で作成される文書情報を用い、相関に対する知識としては、単語をノードとする無向グラフによって記述される事前知識が、利用者の動機に基づいて与えられるものとする。

<sup>1</sup>設計対象に関して設計者が認識している概念のトポロジー構造

これらの情報に対して本システムは、図1に示した矢印に対応する、次に挙げる二つの分析からなる。

- 1 設計過程でドキュメントとして残される情報から事前知識にかかわる議論を計量
- 2 論点の変遷がもつ相関を推定
- 3 推定結果を認識が容易な形式で提示

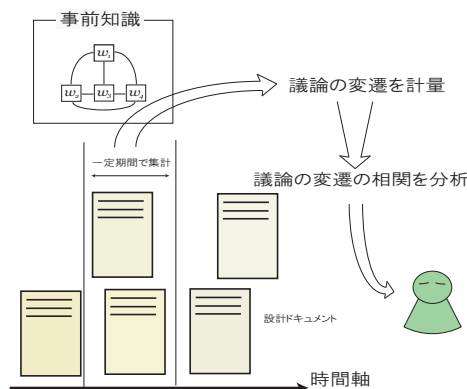


図1: システムの概要

このように、事前知識として与えられた、設計対象に関する概念のトポロジー構造に対して、過去に蓄積されたドキュメントから、それらの概念がどのような文脈で相互に関わり合いを持っていたかを定量して提示する事で、初めて目的を果たす事が出来る。

これらの分析に関する詳細を以下にそれぞれ示す。

### 事前知識に基づく論点の概観

個々の文書における単語の共起によって、文書の論点が構成されているという前提に基づいて分析する。このような単語の共起に基づいて文書の主張を分析する方法は、文書単位での有効性が確認されている [大澤 1999]。

事前知識のノード  $w_i (i = 1, \dots, n_w)^2$  が、一つの文書においてそれぞれ出現する数を要素とするベクトル  $w$  から、積率行列  $\mathbf{W} (= w w^t)$  を個々の文書

<sup>2</sup>シラブルを定義して複数の単語を一つのノードとする事もある。

における  $w_i$  の相関を表す指標として得る。この指標は  $w_i$  が相互にどの程度関連付けられて議論されたか、すなわち論点  $w_i - w_j$  に対する言及の程度を表す。

ここで  $W$  は個々の文書に対して計算されるが、それぞれが持つ情報の粒度や質の違いが直接反映されるため、マクロな論点の変遷を解析するにはノイズになると考えられる。そこで、そのような影響を排除するため、期間ごとに平滑化した指標として  $W_t$  を以下のように求める。

$$W_t = \sum_{t_d \subset t} W_d$$

そして論点の変遷に対応する指標として次のように  $W_{\Delta t}$  を構成する。

$$W_{\Delta t} = W_{t+1} - W_t$$

これは対象行列であり、対角成分を除いた下三角要素は論点  $w_i - w_j$  に関する議論の増減を計量した値となっているから、それらを成分とするベクトル  $x$  を以降の分析の対象とする。

#### 共分散構造分析による相関の分析

共分散構造分析とは、多変量解析の一手法であり、観測変数と潜在変数である構成概念の相関に対して仮説を反映した線形モデルを構成し分析するものである。ここで構成概念の導入やモデルの構成は、モデル構成者が世界を解釈しようとする動機によって選択されるものであるから、本システムはその部分には踏み込まずに、観測変数間の共分散構造のみを分析する。

まず観測された  $x$  を期待値 0 とした平均偏差ベクトル  $v$  および平均 0 分散 1 に標準化した標準得点ベクトル  $z$  を、それぞれ標本数だけ列に並べて平均偏差行列  $V$ 、標準化データ行列  $Z$  とする。そこから標本共分散行列  $S$ 、標本相関行列  $R$  は次の式で計算される。ここで  $N$  は標本数を表す。

$$S = \frac{1}{N} SS^t$$

$$R = \frac{1}{N} ZZ^t$$

標準化モデルの相関構造  $\sum_r$  に対する観測変数  $z$  の構造方程式は、以下のように定義される。

$$Iz = A^*z + Be^*$$

ここで、 $I$  は単位行列、 $A^*$  は  $z_j$  から  $z_i$  への直接効果を表す  $a_{ij}$  を要素として持つ正方行列、 $e^*$  は分

散 1 に標準化された誤差ベクトル、 $B$  は  $e_i^*$  から  $z_i$  への直接効果を表す誤差係数を要素として持つ対角行列である。

$(I - A^*)$  に逆行列が存在するとき、

$$T^* = (I - A^*)^{-1}$$

という表記を用いて、観測変数  $z$  の相関構造  $\sum_r$  は

$$\begin{aligned} \sum_r &= E[zz^t] \\ &= E[T^*Be^*(T^*Be^*)^t] \\ &= T^*BE[e^*e^{*t}]B^tT^{*t} \end{aligned}$$

とモデルの母数から導かれる。同様にして観測変数  $z$  の共分散構造  $\Sigma$  も定義される。

標本相関行列  $R$  からモデルの母数  $A^*$ 、 $B$  を推定することで、論点  $w_i - w_j$  の変遷が相互にもつ効果<sup>3</sup>を定量的に解釈することができる。

すなわち、ある二つの単語の相関の増減が、それに関する議論の発生や収束を表すという仮設のもとにおいて、事前知識として与えられた単語が論点となるような議論の、過去における相関を抽出できる。

ただし共分散構造分析の性質として、与えられるモデルのグラフ構造によっては母数が不定となるため、そのようなモデルはシステムとして棄却する。

### 3 終わりに

本稿では、システムアプローチに基づく設計プロジェクトにおいて蓄積されるような、複数領域に亘る膨大な文書情報から、単語の相関と文書の期間的対応に関して分析することで、プロジェクトの過程における設計に関わる事象の、状況に依存した相関を抽出する方法を提案した。

### 参考文献

[Polanyi 1958] Polanyi, M., “Personal Knowledge: Towards a Post-Critical Philosophy”, The University of Chicago Press, 1958

[大澤 1999] 大澤幸生, Benson, N. E., 谷内田正彦, “KeyGraph: 単語共起グラフの分割・統合によるキーワード抽出”, 電子情報通信学会論文誌 D-, Vol. J82-D-, No. 2, pp. 391-400 (1999)

<sup>3</sup>論点の変遷が無向パスで結ばれている場合は標本相関係数に一致する。