

# テキストマイニングを用いた管理会計・原 価計算100年の研究トピック発見の試み

——LDA (Latent Dirichlet Allocation) を適用した  
トピックモデリングによる分析——

## A Study on the Discovery of Research Topics in Management Accounting and Cost Accounting over 100 Years using Text Mining

——An Analysis using LDA (Latent Dirichlet Allocation) Topic Modeling——

大西 智之

ONISHI Tomoyuki

# テキストマイニングを用いた管理会計・原価計算100年の研究トピック発見の試み

——LDA (Latent Dirichlet Allocation) を適用したトピックモデリングによる分析——

## A Study on the Discovery of Research Topics in Management Accounting and Cost Accounting over 100 Years using Text Mining

——An Analysis using LDA (Latent Dirichlet Allocation) Topic Modeling——

大西 智之

ONISHI Tomoyuki

**要旨：**本研究ノートは、管理会計および原価計算分野における長期にわたる研究について、テキストマイニングを用いてトピックを明らかにすることを目的とする。具体的には、CiNiiに登録されている1907年から2018年までの8,878本の論文タイトルを対象に、Latent Dirichlet Allocation (LDA) によるトピックモデリングを実施した。LDAの結果、管理会計および原価計算に関する20個の主要なトピックが抽出され、各トピックの内容およびトピック間の関係性を考察した。これにより、理論と実践の融合、コスト管理としての原価計算、業績評価と意思決定、産業別管理会計の応用、教育と普及活動といった管理会計・原価計算の主要な特徴が浮かび上がった。また、トピック間の親和性に基づいて、直接原価計算・標準原価計算・原価計算基準が近い論点である、管理会計の実施と展望が論じられやすい、医療・中小企業は戦略マネジメントが研究され易い、やや独立した文脈で利益計画と物流は研究されているなどが明らかになった。本研究の成果は、管理会計および原価計算の研究のこれまでの進展および実践への示唆を提供するものであり、今後の研究の方向性を示す指針となりうるだろう。

**キーワード：**管理会計、原価計算、テキストマイニング、LDA、トピックモデリング

## 1. はじめに

会計史の研究は、主に日本会計史学会が刊行してきた学術誌『会計史学会年報』にて蓄積されてきた。従来の書誌学的研究から、近年はテキストマイニングによる分析も行われている（澤登, 2015）。また、管理会計・原価計算の研究に目を向けると、わが国だけでも、研究対象としてSDGs（小泉, 2024）、管理会計（手嶋・金川, 2022）、原価企画（趙, 2016）、サプライチェーン・マネジメント（原, 2014）などの研究がある。しかしながら、管理会計・原価計算の研究は、全期間に渡る包括的な歴史研究は、十分とは言えない。

そこで、本稿では「管理会計」「原価計算」でCiNiiのタイトル検索をした結果、最初にヒットした1907年から2018年12月21日時点までのおよそ100年間のタイトル（重複排除後8,878本）を、テキストマイニングしたい。全体を分析することによって、管理会計・原価計算が何を研究してきたのかというトピックを明らかにする。

本稿では、トピックモデリングを採用している。具体的には、Latent Dirichlet Allocation（LDA）を用いて、文献の中から隠れたトピックを抽出する。LDAは、各文献が複数のトピックから生成されていると仮定し、その生成過程を逆算することでトピックを明らかにできる。LDAの結果、20個のトピックを同定し、各トピックの内容を考察したあと、トピック間の関係性についても論じた。第2節で先行研究をレビューし、第3節で研究方法を述べ、第4節で分析結果を提示する。第5節で考察を行い、第6節で結びとしたい。

## 2. 先行研究のレビュー

近年テキストマイニングを用いた会計史研究は、2000年代終盤ごろから盛んに研究されてきた（たとえば、中野ほか, 2009; 澤登, 2015など）。管理会計や原価計算でも、テキストマイニングを用いた研究が散見されるようになってきた。たとえば、小泉（2024）は、管理会計・原価計算領域の研究におけるSDGsの取り扱いをテキストマイニングしている。その結果、5つのテーマを同定した。これにより、今後の研究の潮流が推測できる。趙（2016）は、原価企画が生まれてから現在までの151にのぼる「定義」に着目して、原価計算の「基本モデル」を考察した。それにより、基本モデルに加えた「追加要素」により原価企画の定義がなされていることを発見した。原（2014）は、サプライチェーンマネジメント研究を組織間管理会計の視座から検討し、両研究

領域のテキストマイニングをしている。その結果、従来の管理会計研究ではあまり論じられてこなかった論点が発見された。手嶋・金川（2022）は、過去30年にわたる日本の管理会計論文のタイトルに着目して、テキストマイニングを行っている。

本稿が先行研究と異なる点は、3つある。第一に、手法である。テキストマイニングの手法として、トピックモデリングを採用している。本稿ではLDA（Latent Dirichlet Allocation）を適用する。LDAは、文章に隠されたトピックを抽出可能な手法であり用途が広い。たとえば、口コミやレビューの分析などにも用いられている（荒木ほか, 2022など）。会計学の分野でも、例としてKuroki and Hirose（2022, 2023）は、40年にわたる公共サービス会計研究をLDAで分析している。他にも、Federsel et al.（2023）は、2009年～2019年の欧米誌にLDAを適用し、会計学における理論と実務のギャップの発見を試みている。しかしながら、わが国会計学を対象としたLDAは見受けられない<sup>1</sup>。

第二に期間である。日本の原価計算・管理会計研究において、1907年から2018年までの100年を超えた史的研究はない。管理会計のタイトルを対象とした手嶋・金川（2022）であっても、30年間である。むろん、テキストマイニングに限定しなければ、伊藤（2008）などの研究は存在しているし、日本に限定しなければさらに多数の研究がある（上総, 1989; 廣瀬, 1993など）。

第三に、量である。文献数は9,754本ヒットし、重複の排除後でも8,878本に上った。参考として、Kuroki and Hirose（2022, 2023）は823本、Federsel et al.（2023）は2,251本だった。ただし、両文献とも全文を対象としている。LDAは長文にこそ有利に働くが、大量のタイトルに適用することで、一定の傾向が推定できるはずである<sup>2</sup>。

以上より、日本の原価計算・管理会計研究のテキストマイニングを行う本稿は、既存の先行研究と、手法・期間・量の3点から異なる特徴を持っている。

### 3. 研究方法

#### 3.1 概要

CiNiのタイトル検索で「管理会計」「原価計算」でそれぞれ2018年12月21日に検索した結果、1907年を起源とし合計で9,754本がヒットした。重複文献を排除すると8,878本になる。本稿では、Pythonを利用し、8,878本のタイトルを対象に、GensimのLDAモデルによるトピックモデリングを行った。本稿では、本来該当する研究が漏れてしまう可能性を排除することを優先したために、

既存研究で行われているような掲載誌の絞り込みは実施していない<sup>3</sup>。

形態要素解析エンジンとして、京都大学情報科学科と日本電信電話株式会社コミュニケーション科学基礎研究所の共同で開発されたMeCabを利用する。ストップワードの除去にstopwords-ja.json<sup>4,5</sup>を利用した。結果にゆれが生じないようにSEED値を42と設定し固定している。トピックモデルによる潜在意味解析の対象は、名詞、動詞、形容詞とする。試行錯誤した結果、20個のトピックを設定した<sup>6</sup>。

要するに、MeCabを用いた形態素解析とGensimのLDAモデルによるトピックモデリングを実施し、論文タイトルのトピック分析を行った。再現性を確保するためにランダムシードを固定し、安定した結果を得ることができた。

### 3. 2 実際のコードの作成と手順

#### 3. 2. 1 環境設定とライブラリのインストール

ここからは、実際に使ったコードを記述していく。無論、至らない点は数多あるはずであるが、今後の研究と正確に比較するためにも、研究ノートとしてご批判を受けたく、あえて記載させていただいた。まず、LDAを使ったテキストマイニングを行うための環境設定を行う。Google Colabを使用し、必要なライブラリ（MeCab、pandas、gensim、pyLDAvisなど）をインストールする。また、再現性を確保するためにシード値を設定する。シード値は42で固定している。

```
# システムのアップデートとMeCabのインストール
lapt-get update
lapt-get install -y mecab libmecab-dev mecab-ipadic-utf8
!pip install mecab-python3 pandas gensim pyLDAvis ipadic

# ライブラリのインポート
import MeCab
import ipadic
import numpy as np
import random
```

```
import pandas as pd
from google.colab import files
from gensim import corpora
import gensim
import pyLDAvis.gensim_models
import pyLDAvis

# シード値の設定
SEED = 42
np.random.seed(SEED)
random.seed(SEED)

# MeCabの初期化
CHASEN_ARGS = r'-F "%m\t%f[7]\t%f[6]\t%F-[0,1,2,3]\t%f[4]\t%f[5]\n"'
CHASEN_ARGS += r'-U "%m\t%m\t%m\t%F-[0,1,2,3]\t\t\n"'
tagger = MeCab.Tagger(ipadic.MECAB_ARGS + CHASEN_ARGS)
```

### 3.2.2 データのアップロードとストップワードの読み込み

次に、分析対象となるデータとストップワードのリストをアップロードして読み込む。アップロードされたファイルは、Pythonのpandasライブラリを使って読み込み、データフレームとして処理する。

```
# データファイルのアップロード
uploaded = files.upload()
filename = list(uploaded.keys())[0]
data = pd.read_excel(filename)

# ストップワードリストのアップロード
```

```
uploaded = files.upload()
stopwords_filename = list(uploaded.keys())[0]

import json
with open(stopwords_filename, 'r', encoding='utf-8') as f:
    japanese_stop_words = json.load(f)
```

### 3.2.3 テキストのトークン化と前処理

テキストデータをトークン化し、ストップワードを除去する。名詞、動詞、形容詞のみを対象とし、それ以外の単語やストップワードは除去した。これにより、LDAのトピックモデリングに適したクリーンなデータを作成する。

```
# テキストをトークン化する関数の定義
def tokenize(text):
    if not isinstance(text, str):
        text = str(text)
    node = tagger.parseToNode(text)
    words = []
    while node:
        pos = node.feature.split(",")[0]
        if pos in ["名詞", "動詞", "形容詞"] and node.surface not in japanese_stop_words:
            words.append(node.surface)
        node = node.next
    return words

# データの前処理
data['processed'] = data['タイトル'].apply(tokenize)
data = data[data['processed'].apply(lambda x: len(x) > 0)]
```

```
# トークン化されたテキストのリストを作成
texts = data['processed'].tolist()
```

### 3.2.4 辞書とコーパスの作成

トークン化されたテキストを基に、gensimライブラリを用いて辞書とコーパスを作成した。辞書には、全てのユニークな単語が含まれ、コーパスに各文書内の単語の出現頻度が記録される。

```
# 辞書とコーパスの作成
dictionary = corpora.Dictionary(texts)
corpus = [dictionary.doc2bow(text) for text in texts]

print("辞書:")
print(dictionary)
print("コーパス:")
```

### 3.2.5 LDAモデルの作成と可視化

LDAモデルをトレーニングし、トピックを表示および保存する。さらに、pyLDavisを用いてモデルを可視化した。トピック数は20に設定した。

```
# LDAモデルの作成
lda_model = gensim.models.LdaModel(corpus, num_topics=20, id2word=dictionary,
    passes=15, random_state=SEED)

# トピックの表示
for idx, topic in lda_model.print_topics(-1):
```

```
print(f"Topic: {idx} \nWords: {topic}\n")

# トピックの表示と結果の保存のためのデータフレームの作成
topics_data = []
for idx, topic in lda_model.print_topics(-1):
    topics_data.append([idx+1, topic])
topics_df = pd.DataFrame(topics_data, columns=['TopicID', 'Words'])
topics_df.to_excel("lda_topics_output.xlsx", index=False)

# LDAモデルの可視化
lda_display = pyLDAvis.gensim_models.prepare(lda_model, corpus, dictionary,
sort_topics=False)
pyLDAvis.display(lda_display)
```

#### 4. 分析結果

LDAでは、20個のトピックを設定し、各トピック（Topic: 0~19）が発見された。図1は、頻出単語のグラフである。

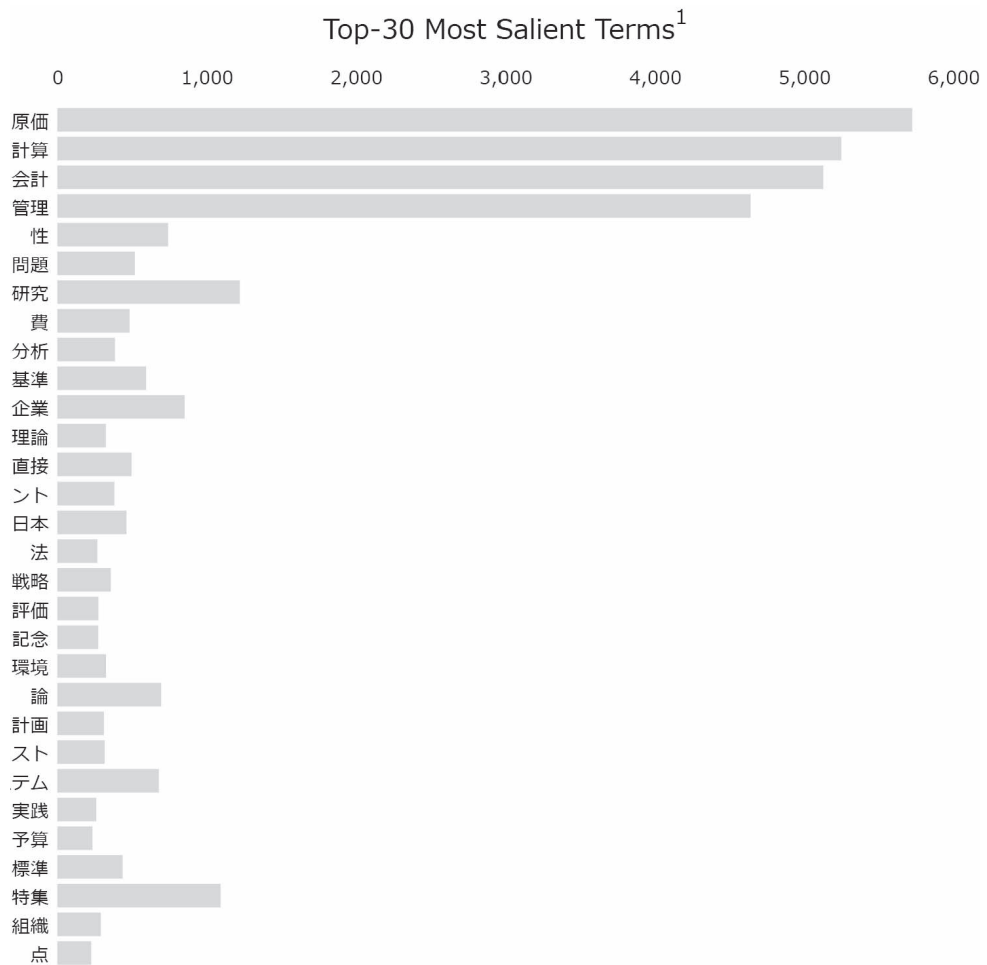
LDAの結果、発見された20個のトピック（Topic: 0~Topic:19）について、Pythonで得られた結果を記していく。

Topic: 1 Words: 0.108\*"記念" + 0.088\*"教授" + 0.085\*"方法" + 0.080\*"号" + 0.072\*"上" + 0.029\*"論" + 0.028\*"文集" + 0.027\*"7" + 0.025\*"退職" + 0.022\*"周年"

Topic: 2 Words: 0.228\*"問題" + 0.134\*"分析" + 0.100\*"点" + 0.051\*"損益" + 0.050\*"差異" + 0.031\*"書" + 0.028\*"報告" + 0.021\*"調整" + 0.020\*"めぐる" + 0.017\*"教育"

Topic: 3 Words: 0.101\*"研究" + 0.064\*"日本" + 0.063\*"企業" + 0.042\*"業" + 0.042\*"学会" + 0.041\*"実態"

図1：頻出単語のヒストグラム



" + 0.040\*"調査" + 0.039\*"回" + 0.029\*"わが国" + 0.029\*"製造"

Topic: 4 Words: 0.202\*"原価" + 0.200\*"計算" + 0.059\*"直接" + 0.036\*"論" + 0.035\*"一" + 0.033\*"考察"  
+ 0.022\*"工業" + 0.021\*"的" + 0.018\*"中心" + 0.017\*"2"

Topic: 5 Words: 0.059\*"病院" + 0.059\*"会" + 0.046\*"部門" + 0.040\*"医療" + 0.037\*"別" + 0.029\*"研究"  
+ 0.027\*"活用" + 0.026\*"中小" + 0.025\*"工場" + 0.025\*"企業"

Topic: 6 Words: 0.253\*<sup>1</sup>"会計" + 0.240\*<sup>1</sup>"管理" + 0.047\*<sup>1</sup>"特集" + 0.039\*<sup>1</sup>"的" + 0.032\*<sup>1</sup>"情報" + 0.023\*<sup>1</sup>"企業" + 0.022\*<sup>1</sup>"課題" + 0.019\*<sup>1</sup>"研究" + 0.019\*<sup>1</sup>"論" + 0.019\*<sup>1</sup>"経営"

Topic: 7 Words: 0.089\*<sup>1</sup>"計画" + 0.075\*<sup>1</sup>"概念" + 0.071\*<sup>1</sup>"利益" + 0.052\*<sup>1</sup>"3" + 0.052\*<sup>1</sup>"2" + 0.049\*<sup>1</sup>"完" + 0.023\*<sup>1</sup>"物流" + 0.019\*<sup>1</sup>"序説" + 0.018\*<sup>1</sup>"紹介" + 0.016\*<sup>1</sup>"売上"

Topic: 8 Words: 0.134\*<sup>1</sup>"費" + 0.076\*<sup>1</sup>"組織" + 0.044\*<sup>1</sup>"事業" + 0.042\*<sup>1</sup>"間接" + 0.041\*<sup>1</sup>"原則" + 0.040\*<sup>1</sup>"測定" + 0.036\*<sup>1</sup>"影響" + 0.032\*<sup>1</sup>"変動" + 0.025\*<sup>1</sup>"部" + 0.023\*<sup>1</sup>"間"

Topic: 9 Words: 0.094\*<sup>1</sup>"講座" + 0.046\*<sup>1</sup>"関連" + 0.046\*<sup>1</sup>"4" + 0.042\*<sup>1</sup>"者" + 0.041\*<sup>1</sup>"5" + 0.040\*<sup>1</sup>"入門" + 0.032\*<sup>1</sup>"特殊" + 0.027\*<sup>1</sup>"構造" + 0.026\*<sup>1</sup>"経営" + 0.022\*<sup>1</sup>"調査"

Topic: 10 Words: 0.154\*<sup>1</sup>"理論" + 0.114\*<sup>1</sup>"実践" + 0.056\*<sup>1</sup>"ドイツ" + 0.048\*<sup>1</sup>"日" + 0.046\*<sup>1</sup>"米" + 0.039\*<sup>1</sup>"品質" + 0.033\*<sup>1</sup>"9" + 0.030\*<sup>1</sup>"鉄道" + 0.021\*<sup>1</sup>"実験" + 0.018\*<sup>1</sup>"文献"

Topic: 11 Words: 0.073\*<sup>1</sup>"環境" + 0.050\*<sup>1</sup>"生産" + 0.047\*<sup>1</sup>"システム" + 0.037\*<sup>1</sup>"会計" + 0.029\*<sup>1</sup>"管理" + 0.027\*<sup>1</sup>"変化" + 0.024\*<sup>1</sup>"性格" + 0.023\*<sup>1</sup>"経営" + 0.022\*<sup>1</sup>"展開" + 0.022\*<sup>1</sup>"企業"

Topic: 12 Words: 0.087\*<sup>1</sup>"管理" + 0.082\*<sup>1</sup>"会計" + 0.029\*<sup>1</sup>"実施" + 0.025\*<sup>1</sup>"経営" + 0.025\*<sup>1</sup>"システム" + 0.018\*<sup>1</sup>"実務" + 0.018\*<sup>1</sup>"研究" + 0.017\*<sup>1</sup>"回" + 0.017\*<sup>1</sup>"特集" + 0.017\*<sup>1</sup>"展望"

Topic: 13 Words: 0.113\*<sup>1</sup>"評価" + 0.090\*<sup>1</sup>"価格" + 0.089\*<sup>1</sup>"決定" + 0.049\*<sup>1</sup>"業績" + 0.038\*<sup>1</sup>"意思" + 0.035\*<sup>1</sup>"役割" + 0.033\*<sup>1</sup>"主義" + 0.033\*<sup>1</sup>"費" + 0.028\*<sup>1</sup>"販売" + 0.024\*<sup>1</sup>"配"

Topic: 14 Words: 0.235\*<sup>1</sup>"性" + 0.049\*<sup>1</sup>"著" + 0.036\*<sup>1</sup>"可能" + 0.035\*<sup>1</sup>"ABC" + 0.028\*<sup>1</sup>"セミナー" + 0.022\*<sup>1</sup>"法人" + 0.021\*<sup>1</sup>"財団" + 0.020\*<sup>1</sup>"型" + 0.020\*<sup>1</sup>"技法" + 0.018\*<sup>1</sup>"討論"

Topic: 15 Words: 0.130\*<sup>1</sup>"法" + 0.066\*<sup>1</sup>"年" + 0.053\*<sup>1</sup>"所" + 0.032\*<sup>1</sup>"期間" + 0.029\*<sup>1</sup>"力" + 0.024\*<sup>1</sup>"必要" + 0.021\*<sup>1</sup>"連結" + 0.020\*<sup>1</sup>"構成" + 0.015\*<sup>1</sup>"要素" + 0.015\*<sup>1</sup>"方"

Topic: 16 Words: 0.085\*"マネジメント" + 0.076\*"戦略" + 0.069\*"コスト" + 0.037\*"的" + 0.030\*"コントロール" + 0.026\*"A" + 0.017\*"展開" + 0.015\*"収益" + 0.014\*"世紀" + 0.013\*"モデル"

Topic: 17 Words: 0.115\*"予算" + 0.075\*"統制" + 0.039\*"会社" + 0.038\*"政策" + 0.034\*"設定" + 0.027\*"利用" + 0.019\*"マテリアルフローコスト" + 0.017\*"方式" + 0.016\*"診療" + 0.015\*"準則"

Topic: 18 Words: 0.075\*"総合" + 0.052\*"企画" + 0.051\*"費用" + 0.048\*"固定" + 0.042\*"個別" + 0.032\*"費" + 0.029\*"特質" + 0.028\*"活動" + 0.027\*"分解" + 0.027\*"構想"

Topic: 19 Words: 0.358\*"原価" + 0.318\*"計算" + 0.053\*"基準" + 0.039\*"標準" + 0.020\*"特集" + 0.019\*"制度" + 0.011\*"2" + 0.011\*"1" + 0.010\*"研究" + 0.010\*"活動"

Topic: 20 Words: 0.103\*"会計" + 0.046\*"管理" + 0.041\*"学" + 0.036\*"現代" + 0.027\*"価値" + 0.027\*"研究" + 0.026\*"期" + 0.024\*"明治" + 0.022\*"造船" + 0.019\*"建設"

以上がPythonで得られたLDAの結果である。次節にて、各トピックの考察を行いたい。

## 5. 分析結果の考察

### 5.1 トピックの考察

本節にて、各トピック（Topic: 1～Topic: 20）の考察をおこなっていく。以下、各トピックのタイトル付けと、その詳細の考察である。トピック個々の考察をしたのち、Intertopic Distance Map (IDM) を基にトピック間の考察を行いたい。

#### Topic 1: 記念と教授

このトピックは「記念」、「教授」、「方法」、「号」、「上」などの単語で構成されている。これらの単語から、記念号や教授の退職記念文集など、特定の出来事や人物を記念する文脈での使用が推測される。これは、管理会計における特定の教授の退職記念号や記念文集のような特集記事に関連するものであり、研究者のキャリアや業績を讃える文化が存在することを示している。

## Topic 2: 損益分岐点と差異分析

「問題」、「分析」、「点」、「損益」、「差異」などの単語が含まれており、主に損益分岐点や差異分析に関連するトピックである。教育に関連する単語も見られることから、教育現場に頻出する管理会計の論点である可能性がある。このトピックは、損益分岐点や細分性について、教育現場での実践的な問題解決に役立つ研究が進められていることを示している。

## Topic 3: 日本企業の調査と実態

「研究」、「日本」、「企業」、「業」、「学会」、「実態」、「調査」などの単語が見られる。日本企業に関する調査研究や実態分析がこのトピックの中心であると考えられる。これは、日本の企業文化や経営実態に基づいた管理会計の適用方法やその効果を検証する研究が多く行われていることを反映していると推測される。

## Topic 4: (直接) 原価計算

「原価」、「計算」、「直接」、「論」、「考察」、「工業」などの単語から、原価計算やその考察に関するトピックである。直接原価計算の方法論に関する考察が多く、これらの技法が企業のコスト管理にどのように役立つかを論じている。

## Topic 5: 医療と中小企業

「病院」、「会」、「部門」、「医療」、「中小」、「企業」などの単語が含まれ、医療機関や中小企業における原価計算・管理会計の活用に関するトピックである。特に、医療機関でのコスト管理や中小企業での部門別管理の重要性が示唆されており、異なる業界での応用可能性を示している。

## Topic 6: 管理会計

「会計」、「管理」、「特集」、「情報」、「企業」などの単語が多く含まれており、管理会計全般に関する特集記事や課題について論じられていると考えられる。管理会計の特集が頻繁に組まれ、企業における情報管理や経営課題への対応策が検討されている。

## Topic 7: 利益計画と物流

「計画」、「概念」、「利益」、「物流」、「売上」などの単語が見られ、計画や利益、物流に関連するトピックである。特に利益計画や物流管理の概念に焦点を当てている可能性がある。利益計画や物流管理の概念が中心となっており、これらの分野における管理会計の役割が重要視されていることが示されている。

#### Topic 8: 組織と間接費

「費」、「組織」、「事業」、「間接」、「原則」、「測定」、「影響」などの単語が含まれており、組織内での費用管理や間接費の測定、影響についてのトピックである。組織構造と費用管理の関係性や間接費の効果的な管理方法が議論されている。

#### Topic 9: 入門講座

「講座」、「関連」、「入門」、「経営」、「調査」などの単語から、管理会計に関する入門講座や基礎知識に関するトピックである。初心者向けの教育資料やセミナーに関する内容が含まれ、管理会計・原価計算の教育普及活動が活発であることが伺える。

#### Topic 10: 管理会計・原価計算の理論と実践

「理論」、「実践」、「ドイツ」、「米」、「品質」、「鉄道」などの単語が含まれ、管理会計・原価計算の理論とその実践、特に国際的な視点からの考察が示唆される。ドイツや米国の事例が取り上げられ、品質管理や鉄道業界における実践事例が論じられているようである。

#### Topic 11: 環境管理会計と生産システム

「環境」、「生産」、「システム」、「会計」、「管理」などの単語が多く、環境管理会計や生産システムに関するトピックである。環境変化への対応や持続可能な生産システムの構築が管理会計の新たな課題として取り上げられている。

#### Topic 12: 管理会計の実施と展望

「管理」、「会計」、「実施」、「経営」、「システム」、「特集」、「展望」などの単語が見られ、管理会計の実施と将来展望に関するトピックである。実務的なアプローチやシステム導入の事例が多く、管理会計の実践的応用が強調されていると思われる。

Topic 13: 業績評価と意思決定（価格決定）

「評価」、「価格」、「決定」、「業績」、「意思」、「役割」などの単語が含まれ、業績評価や意思決定、価格決定に関するトピックである。企業の意思決定における管理会計の役割やその評価方法が論じられている。

Topic 14: ABC(活動基準原価計算)

「性」、「ABC」、「セミナー」、「法人」、「財団」などの単語が見られ、ABC(Activity-Based Costing) や管理会計の技法に関するトピックである。セミナー等で技法の具体的な適用事例やその効果について議論されている。

Topic 15: 法と連結

「法」、「年」、「所」、「期間」、「連結」、「構成」などの単語から、法制度や連結会計に関するトピックである。法的規制や連結財務諸表の作成に関する課題が取り上げられている。

Topic 16: 戦略とマネジメント・コントロール

「マネジメント」、「戦略」、「コスト」、「コントロール」、「モデル」などの単語が含まれ、戦略的マネジメントやマネジメント・コントロール、コストコントロールに関するトピックである。特にマネジメント・コントロールにおける管理会計の役割が強調されていると思われる。

Topic 17: 予算と統制

「予算」、「統制」、「会社」、「政策」、「設定」などの単語が見られ、予算に関するトピックである。予算統制を主とした方法論が多く議論されているのだろう。

Topic 18: 固定費と特集企画

「総合」、「企画」、「費用」、「固定」、「個別」、「活動」などの単語が含まれ、固定費管理に関するトピックである。固定費の研究は企画や特集が組まれることが多いのかもしれない。

Topic 19: 原価計算基準と標準原価計算

「原価」、「計算」、「基準」、「標準」、「制度」などの単語から、原価計算基準や標準原価計算に関するトピックである。原価計算基準や標準原価計算に関する議論がされている。

## Topic 20: 現代管理会計と価値

「会計」、「管理」、「学」、「現代」、「価値」などの単語が見られ、現代管理会計や価値に関するトピックであろう。

以上のように、LDAモデルによって抽出されたトピックは、管理会計・原価計算の多岐にわたる側面を網羅しており、各トピックが特定のテーマに焦点を当てている。この分析により、管理会計・原価計算に関する文献がどのようなテーマに集中しているかを明確にすることができ、今後の研究や実践において重要な指針となるかもしれない。以下の5点が特に注目されよう。

### (1) 理論と実務の融合

Topic10(管理会計・原価計算の理論と実践)、Topic12(管理会計の実施と展望)、は、ドイツや米国を含む原価計算・管理会計の理論とその実践に関する議論を含んでいる。また、その将来の展望も記されている。これらのTopicは、理論的枠組みの確立と実務への適用可能性に焦点を当てている。

### (2) 原価計算とコスト管理

Topic4(直接原価計算)、Topic8(組織と間接費)、Topic14(ABC(活動基準原価計算))、Topic18(固定費と特集企画)、Topic19(原価計算基準と標準原価計算)は、原価計算やコスト管理に関するTopicである。これらのTopicは、直接原価計算や間接費や固定費の管理、活動基準原価計算(ABC)など、企業のコスト管理手法に関する議論を含んでいる。特に、これらの手法が企業の正確な原価計算やコスト削減にどのように貢献するかが重要な論点となっている。

### (3) 企業の業績管理と意思決定

Topic2(損益分岐点と差異分析)、Topic13(業績評価と意思決定)、Topic17(予算と統制)は、企業の業績評価や意思決定に関するTopicである。損益分岐点分析や差異分析、予算管理、業績評価の手法が議論されており、企業の業績管理プロセスにおける管理会計の重要性が示されている。Topic16(戦略とマネジメント・コントロール)も併せて考えれば、管理会計が企業の経営戦略の策定と実行を支援する役割を持っていることを示唆している。

#### (4) 産業別管理会計の応用

Topic5(医療と中小企業)、Topic7(利益計画と物流)、Topic11(環境管理会計と生産システム)は、異なる産業における管理会計の応用に関するTopicである。医療機関や中小企業、物流業界、環境管理会計など、各業界特有の管理会計の課題とその解決策が議論されている。これらの研究は、管理会計の技法が特定の産業にどのように適用され、どのような成果をもたらすかを示している。

#### (5) 教育と普及活動

Topic1(記念と教授)、Topic9(入門講座)は、管理会計・原価計算の教育と普及活動に関するTopicである。これらのTopicは、その基礎知識や教育資料の提供、特定の教授の業績を讃える活動など、教育と普及に関連する活動が含まれている。教育機関や実務家向けの研究や講座も提供されている。

管理会計の分野においては、理論と実践の橋渡しが重要な課題であり、各業界特有の課題に対応するための技法の適用が求められている。また、企業的意思決定と業績管理における管理会計の役割が強調されている。効率的なコスト管理と業績評価の手法が企業経営の成功に必要なことも示唆されている。原価計算はテーマがはっきりしており、コスト管理の手法としての側面も強いようである。さらに、理論と教育の普及活動は、将来の研究者や実務家の育成に欠かせない要素となっており、これらを通じて管理会計・原価計算の知識が広く浸透することが期待される。

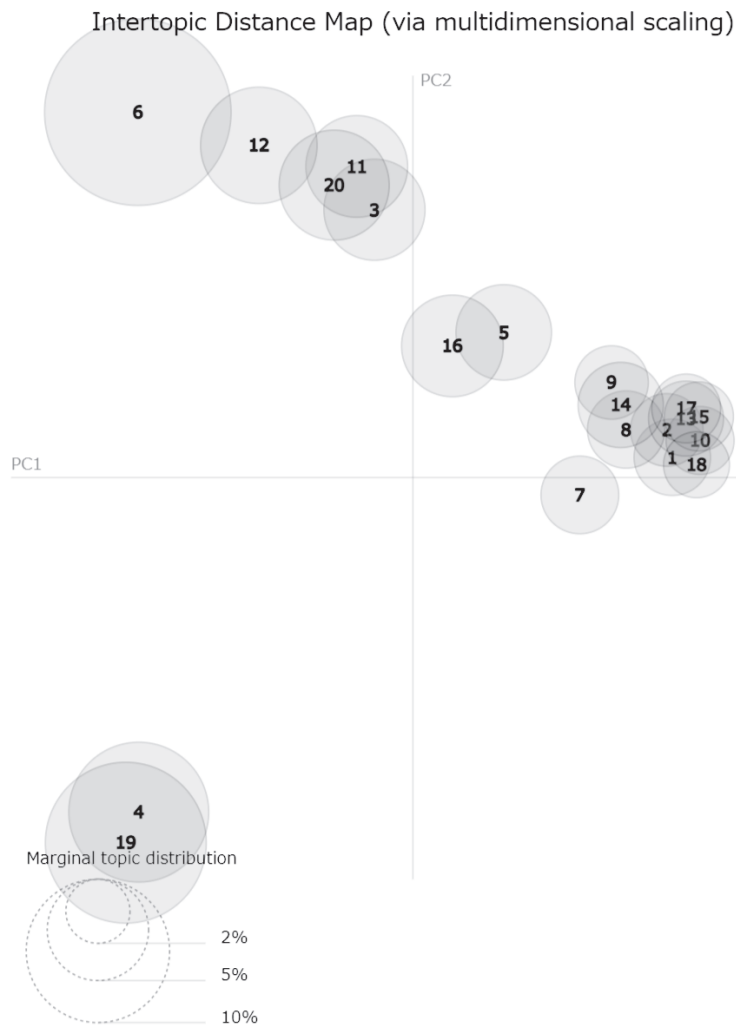
### 5.2 トピック間 (IDM) の考察

Intertopic Distance Map(IDM, 図2)はトピック間の距離をマップ化したものであり、近いトピックほど、内容が共通しているトピックであると推測される。第二の分析として、IDMを考察していく。これを考察した結果、次の5つの発見事項が顕著な論点であると思われる。

発見事項としては、第一に、(直接)原価計算(Topic4)と原価計算基準と標準原価計算(Topic19)が親和性の高いトピックとなっている。これは、原価計算における近い論点が直接原価計算・標準原価計算・原価計算基準であることを示している。第二に、管理会計(Topic6)と管理会計の実施と展望(Topic12)が、やや近いトピックである。管理会計は実施や将来の展望とともに論じられやすいようである。第三に、日本企業の実態調査(Topic3)、環境管

理会計と生産システム（Topic11）、現代管理会計と価値（Topic20）が非常に親和性の高いトピックとなっている。現代管理会計における環境管理会計や生産システムが日本企業の実態調査の対象になっていると思われる。第四に、医療と中小企業（Topic5）と戦略とマネジメント（Topic16）が近い。医療や中小企業は、戦略マネジメントの研究対象となり易いことがわかる。最後に、利益計画と物流（Topic7）がややその他与離れたトピックである。これは、利益計画や物流の研究が、独立した文脈で行われている可能性を示唆している。

図2：Intertopic Distance Map（IDM）



## 6. 結語

本研究ノートの成果は、管理会計および原価計算の研究のこれまでの進展および実践への示唆を提供するものであり、今後の研究の方向性を示す指針となりうる。

本稿の貢献は、1907年から2018年までの約100年間の管理会計・原価計算研究の潮流を俯瞰した点である。さらに、研究手法としてトピックモデリング（LDA）を用い、8,878本におよぶ論文のタイトルを分析した。よって、期間・手法・量の点において、既存の研究と差別化できる。

各トピックの内容とトピック間の関係性についてそれぞれ考察を加えた。各トピックの内容全体を俯瞰した考察としては、①理論と実践の融合、②原価計算とコスト管理、③意思決定と業績評価、④産業別管理会計の応用、⑤教育と普及活動が、管理会計・原価計算の主要な特徴になっていると思われる。

トピック間の内容の関係性の考察としては、⑥直接原価計算・標準原価計算・原価計算基準が近い論点である、⑦管理会計は実施や将来の展望とともに論じられやすい、⑧医療や中小企業は戦略マネジメントの研究対象となり易い、⑨利益計画と物流の研究がやや独立した文脈で行われている可能性があるなどがある。

本稿には限界も多い。まず、初期化方法の違いによって、出力結果が異なってしまう。加えて、ランダムシードを固定しないと、やはり異なる結果が出力されるため非常に注意が必要である。また、トピックの設定個数も任意で行えるため、数を変化させれば出力結果も変化する。さらに、むしろ長い文章にトピックモデリングは向いているため、次回はアブストラクト、もしくは本文全体を対象としたほうが、よりよい結果がでるだろう。

また、予想に反しバランスト・スコアカードがトピックして抽出されなかった理由としては、タイトル表記が「BSC」「バランスコアカード」「バランスト・スコアカード」などゆれがあるためかと思われる。また、1つのトピックに複数の概念が入っている場合があるため、トピック数を今後は増やして分析する必要があるかもしれない。

今後は、トピックモデリングに限らず、多様なテキストマイニングを行いたい。特に、期間を区切って歴史的な変化を観察していくことが必要となろう。とりわけ、期間を区分した本データを用いた共起ネットワーク分析やクラスター分析をすることで、他の研究との比較可能性を高めながら、より詳細な管理会計・原価計算の研究動向が把握できるであろう。また、研究論文を執筆するにあたっては、より慎重な先行研究のレビューは必須である。以上を踏まえ、本研究ノートを嚆矢としたさらなる研究は別稿に譲るものとする。

## 注

1. LDAの詳細についてはOnah et al. (2022) に詳しい。
2. 論文タイトルを対象としたテキストマイニング研究としては、他分野であるが医学誌『臨床神経学』の全収録論文タイトル計7,450本を研究した佐久島ほか (2012) などがある。タイトルのみを対象としているが、一定の研究成果を得た。
3. 原 (2014) を参照されたい。
4. このstopwords-j.jsonファイルは、GitHubのstopwords-iso/stopwords-j.jsonリポジトリで公開されており、自由に使用することができる (<https://github.com/stopwords-iso/stopwords-j.json>)。
5. ["あそこ","あ","あの","あのかた","あの人","あり","あります","ある","あれ","い","いう","います","いる","う","うち","え","お","および","おり","おります","か","かつて","から","が","き","ここ","こちら","こと","この","これ","これら","さ","さらに","し","しかし","する","ず","せ","せる","そこ","そして","その","その他","その後","それ","それぞれ","それで","た","ただし","たち","ため","たり","だ","だっ","だれ","つ","て","で","でき","できる","です","では","でも","と","という","と","といった","とき","ところ","として","とともに","とも","と共に","どこ","どの","な","ない","なお","なかつ","ながら","なく","なつ","など","なに","なら","なり","なる","なん","に","において","における","について","にて","によって","により","による","に対して","に対する","に関する","の","の","のみ","は","は","へ","ほか","ほとんど","ほど","ます","また","または","まで","も","もの","ものの","や","よう","より","ら","られ","られる","れ","れる","を","ん","何","及び","彼","彼女","我々","特に","私","私達","貴方","貴方方"]をストップワードとして除去している。
6. 主にFedersel et al. (2023) を参考にした。

## 参考文献

- 荒木匠平, 小村 亜唯子, 平井 裕久 (2022) 「旅行情報サイトにおけるランキングと口コミ内容の関係性」『日本経営工学会論文誌』 73 (1), 15-26.
- 趙婷婷 (2016) 「テキストマイニングによる原価企画の構成要素に関する検討」『原価計算研究』 40 (1), 133-143.
- Federsel, F. P., Fülbiel, R. U., & Seitz, J. (2024). Research-practice gap in accounting journals? A topic modeling approach. *Journal of Accounting Literature*, 46(3), 368-400.
- 原慎之介 (2014) 「テキストマイニングによる管理会計研究とSCM研究の比較：組織間管理会計への貢献に向けて」『原価計算研究』 38 (2), 136-147.
- 廣本敏郎 (1993) 『米国管理会計発達史』 森山書店.
- 伊藤和憲 (2008) 「管理会計の40年」『専修商学論集』 (88), 13-23.

上總康行 (1989) 『アメリカ管理会計史 (上) (下)』 同文館.

小泉友香 (2024) 「原価計算・管理会計学領域におけるSDGsの取り扱いの展望：テキストマイニングを用いたSSRN掲載の最新学術研究論文の動向調査」『共立女子大学・共立女子短期大学総合文化研究所紀要』 (30), 69-87.

Kuroki, M., & Hirose, Y.(2022). Intra- and interdisciplinary accounting research in public services: topic modeling for 40 years' literature. SSRN Electronic Journal.

Kuroki, M., & Hirose, Y(2023). Topic Classification of Public Sector Accounting Research: A BERT-based Approach. Yokohama City University and Osaka Metropolitan University. Retrieved from SSRN, 27 May 2020. Last revised: 14 Jul 2024.

中野常男・橋本武久・清水泰洋 (2009) 「わが国における会計史研究の過去と現在－テキストマイニングによる一試論－」『国民経済雑誌』 第200巻第4号, 1-23頁.

Onah, D. F. O., Pang, E. L. L., & El-Haj, M(2022). A Data-driven Latent Semantic Analysis for Automatic Text Summarization using LDA Topic Modelling. In 2022 IEEE International Conference on Big Data(Big Data).

佐久嶋研, 佐々木秀直, 田代邦雄 (2012) 「テキストマイニングを用いた学会誌論文タイトルの時系列分析－日本神経学会誌『臨床神経学』の分析－」『医療情報学』 32 (6), 315-321.

澤登千恵 (2015) 「会計史研究の傾向の分析におけるテキストマイニングの活用の試み：『会計史学会年報』 1983～2012年を中心に」『大阪産業大学経営論集』 16 (2/3), 231-246.

手嶋竜二・金川一夫 (2022) 「わが国における管理会計研究の動向と課題－過去30年間の論文タイトルに対するテキストマイニングを用いて－」『九州産業大学商経論叢』 63 (3・4), 1-15.