

Monetary policy synchronization of Korea and United States reflected in the statements

Youngjae Chang^{1,a}

^aDepartment of Statistics and Data Science, Korea National Open University

Abstract

Central banks communicate with the market through a statement on the direction of monetary policy while implementing monetary policy. The rapid contraction of the global economy due to the recent Covid-19 pandemic could be compared to the crisis situation during the 2008 global financial crisis. In this paper, we analyzed the text data from the monetary policy statements of the Bank of Korea and Fed reflecting monetary policy directions focusing on how they were affected in the face of a global crisis. For analysis, we collected the text data of the two countries' monetary policy direction reports published from October 1999 to September 2020. We examined the semantic features using word cloud and word embedding, and analyzed the trend of the similarity between two countries' documents through a piecewise regression tree model. The visualization result shows that both the Bank of Korea and the US Fed have published the statements with refined words of clear meaning for transparent and effective communication with the market. The analysis of the dissimilarity trend of documents in both countries also shows that there exists a sense of synchronization between them as the rapid changes in the global economic environment affect monetary policy.

Keywords: FOMC statement, monetary policy decision, piecewise regression tree, text mining

1. 서론

최근 한국은행이 정책금리 동결을 결정하면서 0 퍼센트 대의 저금리 상황이 지속되고 있다. 중앙은행은 통화정책을 운용하면서 통화정책 방향에 관한 보고서를 통해 시장과 소통하고 있다. 경제환경의 변화와 통화정책에 관한 뉴스는 시장에 상당한 영향을 미치는 요인이다. 최근의 Covid-19 팬데믹은 세계 경제의 급격한 위축을 초래하였다. 2008년 글로벌 금융위기 시와 비교해 보더라도 불확실성이 적지 않은 상황이다. 그 파급효과가 전 세계적으로 확산되면서 경기침체의 장기화에 관한 우려도 증폭되고 있다. 이러한 상황 하에서 경제주체는 정부와 중앙은행의 정책 시그널에 더욱 민감하게 반응하게 된다. Woo와 Chang (2016)은 미국 연준의 통화정책 의결문(statement)을 분석하여, 동결이나 인상, 인하 등 각 금리 정책 결정 시 통화당국이 의결문을 통해 시장과 꾸준하게 소통하였음을 보인 바 있다. 본 논문에서는 미 연준과 한국은행의 통화정책을 담고 있는 통화정책방향 결정문 또는 의결문의 특징을 살펴보고 정책 결정 문서가 세계적인 경제 위기에 어떠한 영향을 받았는지 분석하였다. 선행연구에서는 정책금리의 변동과 그 경제적 영향을 계량적으로 측정하고 의사소통의 주요 특징을 종합하는 데 그쳤으나, 본 연구는 양 국가의 통화정책방향 보고서를 비교하여 시장과의 소통 행태의 추이 및 차이점까지 분석하였다. 양국 중앙은행의 주요 소통 채널인 한국은행의 통화정책방향 결정문과 미 연준의 통화정책 의결문 보도자료를 분석하여 중앙은행이 정책적 시그널이 시장에 효과적으로

¹ Department of Statistics and Data Science, Korea National Open University, 86 Daehak-Ro Jongno-Gu Seoul 03087, Korea. E-mail: yjchang@knou.ac.kr

전달되었는지 살펴보았으며, 최근의 세계적인 경제 위기가 초래한 정책 동조화 현상이 결정문 자료에도 반영되어 있는지 분석하였다. 장기간 지속적으로 금리 동결을 결정하더라도 보고서 내의 단어의 조합이나 구조 등에 따라 뉘앙스가 변하므로 경제 주체들이 받아들이는 의미는 달라질 수 있다. 계량모형에 나타나지 않는 통화정책의 의미적 변화를 텍스트 분석을 통해 효과적으로 포착할 수 있을 것이라 판단하였다. 본 논문의 구성은 다음과 같다. 2장에서는 선행연구에 나타난 접근방법 등을 정리하고 본 연구와의 차이점을 살펴보았으며 3장에서는 실증분석 시 사용된 데이터에 대해 설명하고 분석에 사용된 모형의 특징을 요약하였다. 4장에서는 분석결과를 정리하여 통화정책 보고서 및 의결문의 특징과 상호 관계를 살펴보았으며 5장에서는 본 연구의 시사점을 간략하게 정리하였다.

2. 선행연구

텍스트 마이닝 기법은 대표적인 비정형 데이터인 텍스트 데이터로부터 유용한 정보를 추출해 내는 방법이다. 효과적인 분석을 위해 텍스트 데이터의 수집, 전처리 과정 및 정량화 과정에 이르기까지 다양한 기법들이 응용된다. 텍스트 데이터는 매우 복잡하고 문장 구조에 따른 의미 변화도 빈번하므로 전처리 절차를 거치게 된다. 단어의 출현 빈도를 바탕으로 정량화를 거치면 궁극적으로는 데이터 마이닝 등 잘 알려진 기존 방법을 적용할 수 있는 형태로 변환이 된다. 선행연구에서는 텍스트 데이터의 전처리 및 정량화 과정을 거쳐 계량 분석을 시행하였으며 문서 내에 포함된 의미를 파악하여 시사점을 도출하고 있다.

Choi 등 (2015)에서는 기업 파산 예측 모형을 구축하기 위해 일반적인 텍스트 데이터를 사용하였다. 전통적인 기업 파산 예측 방법론은 기업의 상태를 객관적으로 반영하기 위해 재무 회계 데이터에 의존하지만, 이러한 데이터의 이용은 신속한 분석에 제약이 있으므로 이를 보완하기 위해 실시간 뉴스 텍스트 데이터를 사용하였다. 뉴스 정보를 활용 한 파산 예측을 위해 기업의 파산 사건이 실제로 발생하기 6개월 전 뉴스 문안을 수집하고, 텍스트 마이닝 기법을 활용하여 데이터를 기반으로 한 파산 가능성을 예측하였다. 기존 기업 파산 예측 모형만큼 높은 수준의 예측력을 얻을 수는 없었지만 적시성 있는 파산 모형의 적합이라는 점에서 의미있는 연구라고 주장하였다. Shin과 Kim (2016)은 한글 검색에 있어서 질의어의 효과적인 확장을 위한 방법으로 워드 임베딩과 아프리오리 단계의 중요성을 강조하였다. 명사추출 및 단어빈도-역문서빈도(term frequency-inverse document frequency)를 산출하고 연관성의 측도로 코사인 유사도를 이용하여 텍스트 마이닝을 실시하였으며, 특정 도메인에서 질의어를 확장하여 효과적인 유사답변을 추출하는 모델을 제안하였다. 동 모델을 통해 해당 도메인에서의 질의어와 검색결과 등을 학습하여 연관성 높은 답변을 도출할 수 있음을 보였다. 한편, Woo와 Chang (2016)에서도 미 연준의 의결문을 텍스트 마이닝을 통해 분석한 바 있는데, 문서 내의 단어 목록을 파악하고 이들의 빈도 변화를 파악하여 통화정책 기조의 변화를 판단하였다. 특히 통화정책 의결문의 의미 변화가 실제 정책으로 이루어지는지 정책금리의 움직임을 비교 분석하였다. 동 연구에서는 의결문의 기조 변화가 시차를 두고 실제 금리 변동으로 정책으로 나타났음을 밝혔으며 이를 통해 텍스트 데이터를 통한 향후 정책방향 예측이 가능함을 보였다. Son과 Ko (2018)는 텍스트 마이닝 기법을 활용하여 수학수업 중 나타나는 교사의 감성언어를 탐색하였다. 우수 수업 동영상을 토대로 교사의 강의 내용 전체를 수집하고 이를 텍스트화하였으며 텍스트 데이터의 전처리 과정을 실행하였다. 텍스트 마이닝을 통해 분석한 결과 감성적 단어는 거의 나타나지 않는 것으로 파악되었으며 수학 교과목이라는 특성 상 수업에 사용되는 언어는 정의적 영역에 국한되어 있다고 주장하였다.

본 논문에서는 앞서 살펴본 선행연구의 방법론을 참고하되 분석의 대상을 한국과 미국 2개 국가로 확대하고 양 국가의 통화정책 결정문 및 의결문의 특징을 파악하는 한편, 문서에 투영된 중앙은행의 통화정책방향의 변화를 살펴보았다. 또한 대외 환경 변화가 통화정책방향을 설정하는 결정문의 내용에 어떠한 영향을 미쳤는지 살펴보고, 이러한 환경변화로 양국 통화정책 의결문의 의미상 연관성이 어떻게 변화하였는지 분석해

보았다.

3. 분석방법 및 분석대상

3.1. 텍스트 데이터의 수집 및 가공

본 논문에서는 한국은행과 미 연준의 통화정책의 특징을 파악하고 비교하고자 각 보고서 및 의결문에 포함된 빈출단어를 분석하였다. 미 연준의 의결문이 영문인 점을 고려하여 한국은행 영문 홈페이지에 탑재되어 있는 통화정책방향 결정문의 영어 번역문을 이용하였다. 한국은행 영문 통화정책방향이 국문의 단순 번역본이라는 한계점은 있지만, 각국 중앙은행의 의결문을 참고하고 국문본의 의미를 최대한 살려 작성한 유일한 번역본이므로 유용성이 있다고 판단하였다. 통화정책방향 결정문 또는 의결문의 텍스트 데이터를 수집하기 위해서는 직접 홈페이지에 접속하고 다운로드하여 파일로 저장할 수 있으며, R 프로그램의 *rvest*와 *dplyr* 패키지를 이용할 수도 있다. 이 패키지는 텍스트 자료가 홈페이지에 적절하게 게재되어 있을 경우에는 매우 간단하게 텍스트 자료를 스크랩해 올 수 있는 크롤링 도구이다. 이와 같이 자료 수집이 완료된 이후 전처리 과정을 거치게 된다. 전처리 과정은 대소문자 변환, 문장부호 삭제, 원형복원 등의 과정으로 구성된다. 이후 시각화 도구를 이용하여 각 문서집합의 특징을 요약할 수 있는데, 본 연구에서는 양국의 의결문 별로 워드클라우드(word cloud)를 생성하고 비교하여 빈출단어의 분포를 살펴보았다. 개별 시점의 통화정책 결정문 또는 의결문의 특징을 비교하기 위해서는 단어집합을 토대로 문서단어행렬을 생성하였다. 이 문서단어행렬은 개별 문서의 특징을 나타내는 단어들로 구성되어 있으므로 단어집합의 형태를 살펴보면 특정 문서간의 동일성 또는 이질성을 손쉽게 판단할 수 있다. 다만, 단어빈도에만 근거한 문서단어행렬을 이용할 경우 그 의미 상 중요성보다도 통상적으로 많이 사용되는 단어가 강조될 가능성이 있다. 따라서, 이를 방지하기 위해서 식 (3.1)과 같은 단어빈도-역문서빈도를 사용할 수 있다. 여기서 단어빈도 $Tf_{i,j}$ 는 특정문서 d_j 에서의 단어 w_i 의 출현빈도를 나타내며 문서빈도 Df_j 는 전체 문서들 중에서 단어 w_i 를 포함한 문서의 수를 의미한다. 즉, 단어빈도에 문서빈도의 역수를 곱하여 새로운 빈도수를 정의하는 것이다. 이렇게 변형하면 자주 사용되는 단어들의 가중치는 낮추는 반면 특정 주제와 관련된 문서에 나타나는 단어들의 가중치를 높여줄 수 있게 된다.

$$TF-IDF_{i,j} = \frac{Tf_{i,j}}{Df_j}. \tag{3.1}$$

특히 서로 다른 두 문서를 비교할 경우에는 전체 문서 집합 중 개별 문서에 나타난 단어의 빈도 수 비중이 반영되므로 보다 객관적인 비교가 가능하다. 전체 문서 집합의 정의에 따라 소폭 변이가 있을 수 있다는 점은 유념해야 한다. 한편, 단어빈도 $Tf_{i,j}$ 대신 로그변환($\log(1 + Tf_{i,j})$)을 취하여 사용할 수도 있는데, 이렇게 변환을 통하면 출현 빈도가 높은 특정 단어의 빈도수를 상대적으로 낮추어 반영할 수 있다.

두 문서의 이질성을 판단하는 가장 기본적인 방법은 거리의 개념을 이용하는 것이다. 모든 문서들에 나타난 총 단어 수를 n 이라고 할 때, 특정 두 문서는 개별 단어 빈도를 바탕으로 n 차원 상 한 점에 대응이 가능하다. 이렇게 공간의 한 점에 대응된 각 문서 간의 이질성은 단순히 두 점 간 거리를 계산함으로써 산출이 가능하다. 앞서 살펴 본 개별 문서의 단어빈도-역문서빈도를 바탕으로 두 문서 간 거리를 산출하여 이질성을 판단할 수 있다. 단순한 유클리디안(Euclidean) 거리 개념을 이용할 수도 있지만, 식 (3.2)와 같은 코사인 비유사성(cosine dissimilarity)을 사용하기로 한다. 여기서 x_i 는 특정 문서에서 단어 w_i 의 출현 빈도를 나타내며 y_i 는 비교 대상 문서에서 단어 w_i 의 출현 빈도를 의미한다.

$$\text{Dissimilarity} = 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} = 1 - \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}}. \tag{3.2}$$

3.2. 워드임베딩(word embedding)

워드임베딩은 유사한 의미의 단어들을 유사한 벡터로 변환하여 차원을 축소하고 단어의 연관성을 벡터 연산 등으로 표현하는 방법이다. 유사한 문맥(context)에서 나타나는 단어들은 비슷한 의미를 지니는 경향이 있다는 분포가설의 전제 하에서 이를 구현한다. 앞서 살펴 본 문서단어행렬은 단어의 의미상 유사성을 파악하기 어려운 데다, 너무 많은 단어들로 이루어진 고차원 행렬이므로 다수의 0을 포함하게 되는 문제점을 지니고 있다. 워드임베딩이 이러한 문제점을 극복하는 방법으로서 주목받기 시작하였다. 자료의 차원을 낮춤으로써 유사한 문맥에서 함께 출현한 단어의 의미상 연관성을 쉽게 파악할 수 있기 때문이다. Pennington 등 (2014)은 워드임베딩 방법으로 global vector word representation (GloVe)을 제안하였었다. 이는 특정 단어를 중심으로 적정 구간(window) 내에 문맥상 유의어(context terms)가 얼마나 빈번하게 나타나는지를 파악하여 가중치를 반영한 뒤, 식 (3.3)과 같은 비유함수를 정의하여 연관성을 최대한 잘 나타내는 단어벡터 w_i 를 찾아내는 방법이다. 이와 같은 워드임베딩 절차를 거친 후에도 비유사성 행렬을 산출할 수 있다.

$$J = \sum_{i=1}^V \sum_{j=1}^V f(X_{ij}) (w_i' w_j + b_i + b_j - \log X_{ij})^2, \quad (3.3)$$

여기서 V 는 총단어수를 의미하고 b_i, b_j 는 단어에 상응하는 편의(bias), X_{ij} 는 단어 i 의 문맥상 의미로 나타나는 단어 j 의 출현횟수를 의미한다. 또한, 함수 f 는 과도하게 공통적으로 나타나는 단어들로 인한 영향을 제한하고자 추가된 가중치 함수로서 식 (3.4)와 같다.

$$f(x) = \begin{cases} \frac{x}{x_{\max}}, & \text{if } x < x_{\max}, \\ 1, & \text{otherwise.} \end{cases} \quad (3.4)$$

3.3. 조각별 회귀나무(piecewise regression tree) 모형

회귀나무 모형은 다양한 입력변수가 주어졌을 때, 이러한 입력변수 값에 대해 출력변수 값을 예측하는 모형이다. 일반적으로 출력변수가 범주형인 의사결정나무(decision trees)가 분류나무(classification trees)인 반면, 출력변수의 값이 연속형일 때의 나무모형이 회귀나무 모형이다. Woo와 Chang (2016)에서는 나무모형의 주요 특징을 분기의 기준, 가지치기의 원리 등으로 나누어 살펴본 바 있다. Breiman 등 (1984)에서 제안한 classification and regression trees (CART) 모형은 분류 및 회귀분석의 두 가지 목적을 달성하기 위해 교차타당화(cross-validation) 및 가지치기(pruning) 등의 효과적인 방법을 적용한 나무모형 알고리즘이다. 다만 Loh (2002)에서 보였듯이, CART 모형은 다범주 변수를 설명변수로 사용할 때 발생하는 선택 편의(selection bias)의 문제로 예측력이 저하될 수 있다는 단점이 있다. 이는 CART에서 최적 분기점을 탐색하는 방법으로 사용되는 전역탐색(exhaustive search) 기법의 태생적인 문제점이다. 따라서, 이를 보정하기 위해 직관적이고 유연한 GUIDE (Loh, 2002, 2006)와 같은 모형이 제안되었다. 동 모형은 붓스트랩(bootstrap) 방법으로 선택 편의를 제어했을 뿐만 아니라 다양한 선형모형을 나무모형의 각 노드에서 적합시킬 수 있도록 허용함으로써 종래의 분류나무에 국한되었던 나무구조 모형의 확장을 이루었다는 점에서 의의가 있다. 각 노드에 적합된 다수 선형모형의 조각(pieces)들을 결합함으로써 모형 전반의 예측력 향상을 이루었고 이러한 구조 때문에 조각별 회귀나무라고 불린다.

$$GI(t) = 1 - \sum_{j=1}^J p(j|t)^2, \quad (3.5)$$

$$IMP(t) = \frac{1}{N_t} \sum_{y_i \in t} (y_i - \bar{y}_t)^2. \quad (3.6)$$

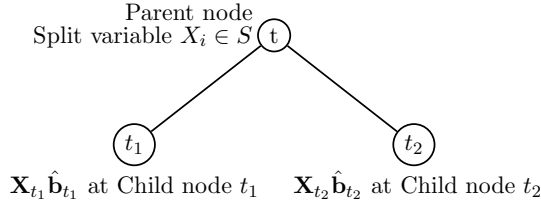


Figure 1: Piecewise regression tree.

여기서 \bar{y}_t 는 노드 t 에 속한 관측치들의 출력변수(혹은 종속변수) 값의 평균치이므로 식 (3.6)은 노드 t 에서의 표본분산을 의미한다. 분류모형에서 불순도를 나타내는 식 (3.5)와 같은 Gini 지수를 표본분산의 형태로 변환하고 이 기준에 따라 분기 변수 및 분기점을 찾아 나무를 구현하면 회귀나무 형태의 나무모형이 나타나게 된다. 여기서 각 노드에서 적합된 모형은 $y = \beta_0$ 형태의 상수항(constant) 모형이다. Chang과 Kim (2011)에서도 지적했듯이 상수항 모형으로 비선형 모형을 적합할 경우에는 다음과 같은 문제점이 발생할 수 있다. 비선형 구조를 형성하기 위해 너무 많은 분기점을 지닌 모형이 선택됨에 따라 과다적합의 가능성이 있다. 혹은 이와 반대로 전혀 분기를 이루지 않은 뿌리노드 나무구조가 적합되어 비선형성을 포착하지 못할 수도 있다. 따라서, 각 노드에 선형모형을 적합하여 상수항 모형의 제약을 완화함으로써 유연성을 확보하도록 한 것이 조각별 회귀나무 모형이다. Chang (2020)에서도 다차항들의 선형 결합으로 비선형성을 포착할 수 있는 투영모형의 원리를 소개한 바 있는데, 조각별 회귀나무 모형은 각 분기가 일어나는 공간에서 개별 선형모형을 고려함으로써 이보다 더 유연한 모형적합이 가능하다고 할 수 있다.

$$IMP(t) = \frac{1}{N_t - p} (\mathbf{Y}_t - \mathbf{X}_t \hat{\mathbf{b}}_t)' (\mathbf{Y}_t - \mathbf{X}_t \hat{\mathbf{b}}_t). \tag{3.7}$$

조각별 회귀나무 모형의 적합을 위해서는 불순도 함수를 식 (3.7)과 같이 각 노드에서의 평균제곱오차(mean squared error)로 바꾸어주어야 한다. 여기서 \mathbf{Y}_t 는 노드 t 에서의 출력변수 벡터, $\mathbf{X}_t \hat{\mathbf{b}}_t$ 는 입력변수 행렬과 회귀계수 추정 벡터와의 곱으로서 해당 노드 t 에서의 예측치이다. 한편, 이러한 조각별 회귀나무 모형은 노드의 제약조건 하에서 적합된 조건부 회귀모형으로 볼 수 있으며 로지스틱 회귀모형(logistic regression)이나 분위수 회귀모형(quantile regression) 등으로의 확장이 가능하다 (Loh, 2014).

Figure 1은 뿌리 노드이자 부모 노드인 노드 t 와 분기변수 X_i 의 값에 따라 자식 노드 t_1 및 t_2 로 이루어진 스템프(stump) 구조에 조각별 회귀나무 모형을 구현한 사례이다. 결국 그림의 회귀나무는 식 (3.8)과 같이 조건부 기댓값을 추정하는 모형이다. 충분한 크기로 자란 실제 회귀나무 모형은 많은 분기변수들로 이루어진 다차원 공간의 각각의 부 공간들로 이루어진 영역에서 이와 같이 적합된 조각별 모형의 집합이라 할 수 있다.

$$E(\mathbf{Y}|t_1, t_2, i) = \mathbf{X}_{t_1} \mathbf{b}_{t_1} I(X_i \in S) + \mathbf{X}_{t_2} \mathbf{b}_{t_2} I(X_i \notin S). \tag{3.8}$$

Loh (2014) 등에서는 실제 데이터를 바탕으로 다양한 알고리즘의 성능을 비교 평가하였으며, 기존의 방법론에 비해 조각별 회귀나무 모형 접근법 및 가지치기를 하지 않은 조각별 회귀나무모형의 앙상블 방법이 예측력 향상에 도움을 준다는 점을 강조하고 있다.

3.4. 분석대상

본 논문에서는 한국은행과 미 연준이 공표한 통화정책방향 결정문 및 FOMC 의결문의 텍스트 데이터를 분석하였다. 실제 통화정책의 방향과 문서에 담긴 정책방향의 비교를 위해서 기준금리 추이를 살펴보았으며 정책금리의 변동과 텍스트 데이터 분석을 통해 얻은 의결문 및 결정문 등의 의미 변화를 비교하였다. 양국

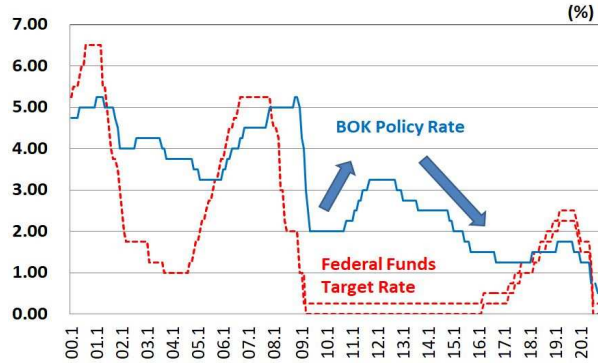


Figure 2: Policy rate trends.

모두 중앙은행의 홈페이지를 통해 기준금리의 변동과 결정 일자, 의결문 및 결정문 등 공식 문서를 공표하고 있는데, 텍스트 자료는 미 연준은 1999년 5월부터, 한국은행은 국문의 경우 1999년 1월부터, 영문의 경우 1999년 10월부터의 문서를 탑재해 놓고 있다. 본고의 주된 분석 내용이 각국의 문서와 개별 통화정책 방향과의 관계 보다는 국가간 텍스트 데이터의 비교라는 점을 고려하여 영문 텍스트 데이터를 이용하였다. 각 문서의 가용 기간에 다소 차이가 있으므로 분석 대상 기간 시점을 일치시켜 1999년 10월로 설정하였으며 2020년 9월 공표된 결정문까지 사용하였다. 다만, 여기서 주의할 것은 미 연준의 경우 통화정책 방향을 결정하는 FOMC 회의가 매월 개최되지는 않았다는 것이다. 따라서 매월 통화정책방향 결정을 위한 금융통화위원회 회의를 개최한 한국은행보다 다소 자료 수가 적다. 한국은행도 2017년부터 매년 8회로 횟수를 축소하되 긴급한 상황이 발생하였을 경우에는 임시회의 일시를 추가하여 통화정책을 위한 금융통화위원회를 개최하고 있다. 미 연준의 경우 1999년 10월부터 2010년 9월까지 공표된 의결문은 총 175개이며 한국은행의 경우에는 총 239개이다.

4. 통화정책 동조화 현상 분석 결과

4.1. 빈출단어 분석

Figure 2는 1999년 10월부터 2020년 9월까지의 한국은행과 미 연준의 정책금리 추이를 나타낸 것이다. 실선은 한국은행의 기준금리, 점선은 미 연준의 기준금리를 의미한다. 한국은행의 금리는 한국은행 홈페이지 (<https://www.bok.or.kr>)에서, 미 연준의 정책금리는 세인트루이스 연준의 데이터베이스(<https://fred.stlouisfed.org/>)로부터 얻을 수 있는데, 2008년 12월부터 목표치의 하한과 상한을 따로 공표하여 정책금리 운용 폭을 제시하였다.

한국은행과 미 연준의 정책금리는 운용 목표치이므로 선언적 의미가 있다. Figure 2를 보면, 수준의 차이는 있으나 양국의 정책금리는 대체로 유사하게 움직이고 있다. 다만, 2003년, 2008년 및 2011년을 전후해서 양국의 정책금리 변동 방향이 상이하게 나타나고 있으며 일부 지점에서는 약간의 시차를 두면서 동일한 방향성을 보이고 있다. 특히 2008년말 미국의 금융위기에서 촉발된 글로벌 금융위기 이후 한동안 양적완화(quantitative easing; QE) 정책을 지속하면서 미 연준의 정책금리는 변동이 없었으나 한국은행의 정책금리는 일정 기간 인상을 지속하였다가 다시 인하를 거듭하는 등 상이한 행태를 보였다. 이는 경제 환경에 대한 인식 및 선제적인 정책 집행 의지 등 양국의 서로 다른 정책 시각이 영향을 미쳤다고 볼 수 있다. 이러한 정책적 차이가 통화정책 결정문 및 FOMC 의결문에 적절하게 나타나 있는지 확인해 보도록 한다. Figure 3은 한국은행과 미 연준의

Table 1: Word pairs of similar meanings in the context

Word pairs (BOK)		Cosine similarity	Word pairs (FRB)		Cosine similarity
monetary	policy	0.9518	gain	job	0.8720
go	forward	0.9263	fomc	vote	0.8710
korean	won	0.8685	price	stability	0.7943
agricultural	petroleum	0.8646	monetary	policy	0.7927
japanese	yen	0.8536	business	household	0.7693
home	abroad	0.8439	submit	request	0.7660

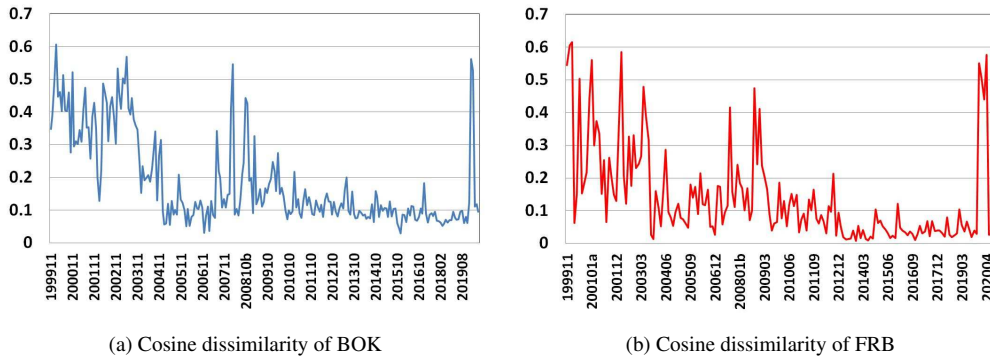


Figure 5: Cosine dissimilarity.

로 양국의 문서는 매우 유사한 의미를 지닌 단어들로 구성되어 있다. 한편, 전체 문서에 담긴 단어의 의미적 유사성을 확인해 보기 위해 Pennington 등 (2014)을 참조하여 워드 임베딩 기법을 적용한 뒤 단어 간의 코사인 유사성을 산출해 보았다. 텍스트 데이터의 규모가 크지 않기 때문에 알고리즘을 통한 학습(learning)이 이루어지기에 제한적이긴 하지만, 차원을 축소하고 유사성을 산출한 결과를 살펴보았다. Table 1에는 의미적으로 유사한 단어 쌍들 중 몇 가지가 정리되어 있다. “go”와 “forward”나 “monetary”와 “policy” 등에서처럼 문장에서 함께 사용되는 사례가 많은 단어들은 높은 의미적 관련성을 지닌 것으로 나타난다. 코사인 유사성이 0.7 이상인 단어들 중 이처럼 짝지어진 단어들이 매우 많이 나타났다. 주로 “won”과 “korean”, “japanese”와 “yen” 등과 같이 의미적으로 유관한 단어들이 분포하고 있다. “business”와 “household”가 짝지어 나타난 것도 경제 주체로서 의미상 연관성을 나타낸다고 할 수 있다. 다만, 간혹 “home”과 “abroad”처럼 상반된 의미의 단어가 추출된 경우도 발견할 수 있었는데, 이는 동일한 문맥에서 동시 출현한 빈도수에 따라 연관성을 판단하기 때문에 발생하는 워드 임베딩의 특징이자 유의점이기도 하다.

다만, 이제까지 살펴본 워드 클라우드와 워드 임베딩을 통한 연관어 분석은 전 기간에 걸친 통합 분석으로서 정제된 공식 문서의 의미상 유사성은 충분히 살펴볼 수 있으나 시간의 흐름에 따른 정책적 변화를 분석할 수는 없다. 한국은행 통화정책 결정문과 미 연준의 FOMC 의결문 간의 동조화현상을 분석해보기 위해 코사인 비유사성의 추이를 살펴보기로 한다.

4.2. 인접 문서 간 비유사성 추이 분석

한국은행과 미 연준의 통화정책 결정문과 의결문의 의미는 각 문서에 포함된 단어들의 구성으로 판단해 볼 수 있다. 문서 내에 포함된 단어의 구성이 어떻게 변화하는지에 따라 정책기조의 변화 여부를 가늠해 볼 수 있는 것이다. 이를 위해서 코사인 비유사성을 산출하고 그 추이를 살펴보았다. 비유사성 측도는 코사인 비유사성을 사용했으며 비교의 대상이 되는 문서는 공개 시점 t 를 기준으로 $t - 1$ 시점이다. 따라서 비유사성은 1999년 11

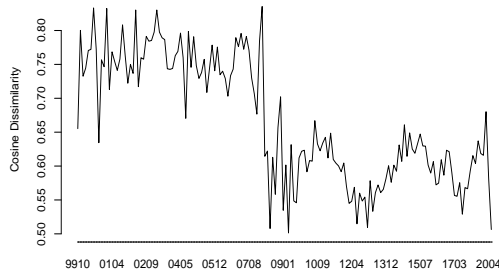


Figure 6: Dissimilarity between BOK and FRB.

월부터 계산할 수 있다. 미 연준의 경우 전 기간, 한국은행의 2017년 이후에는 문서가 매일 공표되지 않았으므로 인접 문서라는 의미가 인접한 두 달 간 문서의 비교가 아니라라는 점에 유의하도록 한다. 특정 월에 2회 이상의 결문이 공표되었을 때는 200101a 또는 200810b처럼 문자로 구별하여 표시하였다.

Figure 5에서는 한국은행이나 미 연준의 경우 모두 비유사성이 매우 큰 폭의 등락을 거듭하고 있는 부분을 찾아볼 수 있다. 이는 정책방향의 전환이 빈번했던 시기를 의미하며 정책금리의 변화 시점과 대체로 일치한다. 특히 한국은행의 경우를 보면, 1999년부터 2000년대 초반까지 상대적으로 높은 수준을 보이던 비유사성이 점차 낮아져 상당히 낮은 수준에서 안정적인 모습을 보이고 있다. 다만, 2007년 및 2008년, 그리고 2020년에 접어들어 급격한 등락의 모습을 보이는데, 글로벌 금융위기와 최근의 Covid-19 팬데믹 위기 시점과 일치한다. 미 연준의 경우에도 2008년 글로벌 금융위기 전후, 2020년 Covid-19 팬데믹 위기에 따른 중앙은행의 정책기조 변화가 반영된 점에서는 한국은행과 유사하다. 그러나, 전체적인 추이를 살펴보면 한국은행과 다소 상이한 모습을 찾아볼 수 있다. 이는 1999년부터 2000년대 초반의 한국의 외환위기 및 가계 신용카드 대출 부실 사태 등이 국내에 국한된 경제적 충격이었던 데 반해, 2000년 후반 및 최근의 팬데믹 상황 등은 글로벌 경제 전반에 영향을 미치는 충격이라는 차이점에 기인한 것으로 보인다. 각국의 통화당국은 이러한 경제 환경 변화에 따라 신속하게 정책적인 대응을 해왔음을 시사한다 하겠다.

한국은행의 통화정책 결정문과 미 연준의 의결문 간의 유사성을 직접 비교하면 글로벌 환경 변화에 따른 정책변화의 기조는 어떻게 변화했는지 보다 명확하게 판단할 수 있다. 이를 위해서 동일 시점의 양국 문서 간 코사인 비유사성을 계산하고 이의 추이를 살펴보았다. Figure 6을 보면, 2007년 중반을 기준으로 이전과 이후의 양국 문서 간 비유사성의 수준이 차이를 나타내고 있음을 알 수 있다. 앞서 살펴본 한국은행과 미 연준 개별 문서의 추이만으로는 충분히 파악할 수 없었던 정책결정 문서의 의미 상 동조화 현상을 명확히 파악할 수 있다. 이는 조각별 회귀나무 모형을 적용하면 더욱 면밀하게 살펴볼 수 있다.

4.3. 나무모형을 통한 비유사성 추이 분석

Figure 7은 조각별 회귀나무 적합 결과이다. 통화정책방향 결정문과 의결문의 공개 시점 t 라고 할 때, 해당 시점의 두 문서 간 비유사성 측도(CD)를 출력변수로, $t-1, t-2, t-3$ 시점의 비유사성 측도(CD1, CD2, CD3) 및 t 등을 입력변수로 설정하여 각 노드에서 분기변수 후보 또는 최종 노드에서 선형모형의 독립변수로 적합될 수 있도록 허용하여 나무모형을 적합하였다. 적합 결과, 뿌리노드에서 가지가 나뉘는 기준은 2008년 4월 이전인지 여부로 나타났고, 한번 더 분기가 일어난 노드 3에서의 분기변수는 $t-1$ 시점의 비유사성 측도였다. 각 최종노드 아래에는 관측치 수와 비유사성의 평균이 함께 표시되어있는데, 2008년 4월 이전의 평균적인 비유사성의 수준이 0.76으로 그 이후 시점의 0.59나 0.61에 비해 높은 수준을 나타내어 앞서 살펴본 대로 글로벌 경제환경의 변화가 통화정책에 영향을 미치면서 문서 간 의미상 동조화가 이루어졌음을 알 수 있다. 각 최종노드에서 적합한 선형 회귀모형은 Table 2와 같다.

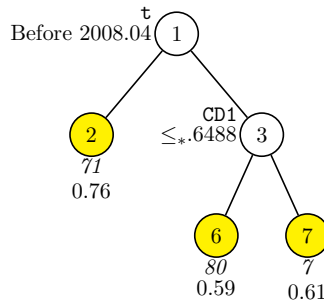


Figure 7: GUIDE v.27.4 0.00-SE piecewise linear least-squares regression tree with stepwise variable selection for predicting CD. At each split, an observation goes to the left branch if and only if the condition is satisfied. The symbol ' \leq_* ' stands for ' \leq or missing'. Sample size (in italics) and mean of CD printed below nodes. Second best split variable at root node is CD1.

Table 2: Coefficients of least squares regression at terminal nodes

Terminal node	Regressor	Coefficient	T	P -value
2	Constant	0.7585	163.07	0.0000
	Constant	0.2557	3.75	0.0003
6	CD1	0.4261	3.56	0.0006
	CD2	0.4040	4.60	0.0000
	CD3	-0.2601	-2.95	0.0043
7	Constant	0.6149	358.1	0.0000

5. 결론

본 연구는 한국은행의 통화정책결정문과 미 연준의 FOMC 의결문 텍스트 데이터를 바탕으로 의미상 동질성을 산출하여 통화정책이 동조화를 이루고 있다는 점을 분석한 것이다. 분석결과로부터 한국은행과 미 연준의 문서 모두 정제된 단어로 효과적으로 시장과 소통하고 있으며 명확한 의사 전달을 위해 단어 선택에도 신중을 기하고 있음을 알 수 있었다. 정책의 실제 집행에 앞서 정책 시그널의 도구로 사용된 각 문서의 의미가 잘 나타나고 있어, Woo와 Chang (2016)에서 분석한 바와 같이 향후 정책의 기초적인 변화 포착을 위한 텍스트 마이닝 기법 활용의 적절성을 확인해준 분석결과라고 할 수 있다. 특히, 글로벌 금융위기나 Covid-19 팬데믹으로 인한 급격한 경제 환경 변화에 신속하게 대응한 각국의 통화정책의 방향이 문서에 뚜렷하게 나타난 것으로 분석되었다. 이는 외환위기 및 가계 신용카드 대출 부실 사태 등의 국내 경제상황의 악화에 대응했던 2000년 전후의 시기와는 확연하게 다른 모습으로서, 전 세계적인 경제 위기에 적극적으로 대응한 각국의 통화정책 방향성이 유사하게 반영된 데 기인한다. 이런 동조화 현상을 다년에 걸친 텍스트 데이터의 분석으로부터 확인하였다는 점에서 본 연구의 의의를 찾을 수 있겠다. 다만, 1999년 이래로 공개된 한국은행의 통화정책방향 결정문과 미 연준의 의결문 텍스트 데이터가 양적인 측면에서 충분하지 않다는 점과, 한국은행 문서의 경우 영문 번역본이라는 점에서 번역의 일관성이 결여될 가능성이 있다는 점 등은 본 연구의 한계점으로 지적할 수 있다.

References

Board of Governors of the Federal Reserve System, <http://www.federalreserve.gov/monetarypolicy/default.htm>

- Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*, Wadsworth, Belmont, CA.
- Chang, Y. and Kim, H. (2011). Tree-structured nonlinear regression, *The Korean Journal of Applied Statistics*, **24**, 759–768.
- Chang, Y. (2020). The asymmetry of economic perceptions analyzed with business survey index, *Journal of the Korean Data Analysis Society*, **22**, 1863–1874.
- Choi, J., Han, H. S., Lee, M. Y., and Ahn, J. M. (2015). The prediction of corporate bankruptcy using text-mining methodology, *Productivity Review*, **29**, 201–228.
- FRED, Federal Reserve Bank of St. Louis Economic Data, <https://fred.stlouisfed.org/>
- Loh, W. Y. (2002). Regression trees with unbiased variable selection and interaction detection, *Statistica Sinica*, **12**, 361–386.
- Loh, W. Y. (2008). Regression by parts: Fitting visually interpretable models with GUIDE, *Handbook of Computational Statistics*, **III**, 447–469, Springer.
- Loh, W. Y. (2014). Fifty years of classification and regression trees (with discussion), *International Statistical Review*, **34**, 329–370.
- Shin, D. H. and Kim, C. B. (2016). Query extension of retrieve system using hangul word embedding and apriori, *The Journal of Korea Navigation Institute*, **20**, 671–624.
- Son, B. E. and Ko, H. K. (2018). The frequency analysis of teacher’s emotional response in mathematics class, *Communications of Mathematical Education*, **32**, 555–573.
- Pennington, J., Socher, R., and Manning, C. D. (2013). *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543.
- The Bank of Korea, <https://www.bok.or.kr/eng/bbs/E0000627/list.do?menuNo=400022>
- Woo, S. and Chang, Y. (2016). An analysis of FOMC statements by text mining methods, *Journal of the Korean Data Analysis Society*, **18**, 179–188.

Received November 16, 2020; Revised November 29, 2020; Accepted November 29, 2020

통화정책 결정문에 나타난 한미 통화정책 동조화 현상 분석

장영재^{1,a}

^a한국방송통신대학교 통계·데이터과학과

요 약

중앙은행은 통화정책을 운용하면서 통화정책 방향에 관한 보고서를 통해 시장과 소통하고 있다. 최근의 Covid-19 팬데믹은 세계적인 경제의 급격한 위축을 초래하였다. 2008년 글로벌 금융위기 시와 비교해 보더라도 불확실성이 적지 않은 상황이다. 그 파급효과가 전 세계적으로 확산되면서 경기침체의 장기화에 관한 우려도 증폭되고 있다. 본 논문에서는 미 연준과 한국은행의 통화정책을 담고 있는 통화정책방향 결정문과 의결문의 특징을 분석하고 세계적인 위기에 어떠한 영향을 받았는지 살펴보았다. 분석을 위해 1999년 10월부터 2020년 9월까지 공표된 양 국가의 통화정책방향 보고서 텍스트 자료를 수집하였으며 워드 클라우드 및 워드 임베딩 등을 이용하여 의미상 특징을 살펴보았다. 조각별 회귀나무 모형을 통해 양국 문서의 비유사성 추이도 분석해 보았다. 분석 결과 한국은행과 미 연준 모두 시장과의 투명하고 효과적인 소통을 위해 명확한 의미를 지닌 단어로 정제된 문서 자료를 작성하고 있는 것으로 나타났다. 또한, 급격한 글로벌 경제환경의 변화가 통화정책에 영향을 미치면서 문서 간 의미상 동조화가 이루진 것으로 나타났다.

주요용어: 조각별 회귀나무 모형, 텍스트 마이닝, 통화정책방향 결정문, 통화정책 의결문
