

김정은 정권 시기 북한 핵·경제 정책의 동학

2020. 02. 28.

[연구진]

연구책임자 : 김병연 (서울대학교 경제학부 교수)

공동연구원 : 최경희 (샌드연구소 대표)

김규일 (세종컨설팅 대표)

김명성 (샌드연구소 연구위원)

연구보조원 : 조용신 (서울대학교 경제학부 박사과정)

목 차

1. 연구배경과 목적	1
2. 자료개요	2
3. 분석방법론 및 적용방법	3
3.1 DTM과 TDM	3
3.2 연관분석	4
3.3 상관분석	4
4. 주요결과 및 시사점	5
4.1 “핵”과 “경제” 기간별 출현 빈도 분석	5
4.2 “핵”과 “경제” 출현 성격 분석	10
4.3 “핵”과 “경제” 동시출현 단어군 분석	14
4.4 “핵”, “경제”와 한반도 사건들과의 관계 분석	17
5. 결론	24
부록 : 분석방법론	26

표 목차

<표 1> 핵, 경제 단어 출현 분기별 빈도 (단어 수 기준)	7
<표 2> 핵과 경제 단어 분기별 출현 빈도 및 비율 (문단 수 기준)	8
<표 3> 기간별 핵과 경제 단어 출현 빈도 간 상관관계	9
<표 4> 핵과 경제 동시출현 비율 분석 결과	11
<표 5> 핵과 경제 관련단어군중 서로의 순위	12
<표 6> 핵과 경제 관련 단어군	16
<표 7> 경제·핵 관련 단어 별 공통 관련 단어 순위	16
<표 8> 종속변수 구성 내용	20
<표 9> 시계열 로짓 회귀분석 결과표	23
<표 10> OLS 회귀분석 결과표	24
<표 11> TDM 및 동시출현 빈도분석 표	26

그림 목차

<그림 1> 핵과 경제 단어 분기별 출현 빈도 변화 추이 (단어 수 기준)	7
<그림 2> 핵과 경제 단어 분기별 출현 빈도 변화 추이 (문단 수 기준)	8
<그림 3> 핵과 경제가 같이 출현한 문건 비중 분기별 추이 (문건 수 기준)	11
<그림 4> 핵, 경제 관련단어군 출현 비중과 순수 핵 관련단어군 출현 비중 추이	18
<그림 5> TDM, DTM의 형태	27
<그림 6> 토픽 모델링 개념	29
<그림 7> LDA 분석 과정	30
<그림 8> LDA 그래프 모델	30

1. 연구의 배경과 목적

- 북한 비핵화 협상에 있어 가장 중요한 질문 중 하나가 북한 정권의 비핵화 의지에 관한 것임. 즉 북한은 경제 건설을 위해 핵을 포기할 의사가 있는지에 관해 상반되는 가설이 존재함.
 - 일각에서는 김정은이 경제 발전을 위해 비핵화를 할 의지가 충분히 있으며 비핵화 협상에서 제시되는 조건에 따라 그 가능성이 열려 있다고 주장함.
 - 반면 김정은은 핵을 포기할 의사가 없으며 비핵화 협상에서 그가 의도하는 것은 실질적인 핵 보유국이 되는 것이라는 주장도 존재함.
- 이상의 중요한 질문에 대해 답하기는 어려움. 그러나 한 가지 시도할 수 있는 방법은 북한의 시각에서 이 관계를 살펴보는 것임.
- 이 연구는 북한 자료를 대상으로 텍스트 마이닝 기법을 이용, 핵과 경제의 관계를 살펴 봄. 즉 북한 지식인, 전문가, 또는 공식적인 프로퍼갠더가 이해하는 핵과 경제의 관계를 파악하려고 함.
 - 보다 구체적으로 핵과 경제의 관계에 대해서는 다음과 같은 세 가지 가능성이 존재함. 양의 상관관계(보완재), 상호 독립적인 관계(독립재), 그리고 음의 상관관계(대체재)임.
 - 텍스트 마이닝을 통해 핵과 경제를 구성하는 관련 단어군의 상당수가 서로 겹치고 그 의미가 서로를 강화하며 핵과 경제의 분기별 출현 빈도의 상관관계가 높다면 이는 북한은 이 두 단어의 관계를 보완재로 이해하고 있다고 볼 수 있을 것임. 반면 핵과 경제를 구성하는 관련 단어군이 거의 겹치지 않고 출현 빈도의 상관관계가 낮다면 독립재, 관련 단어군이 겹치지만 그 의미가 서로에 부정적으로 영향을 줄 수 있다면 대체제로 이해할 수 있을 것임.

- 이상의 분석을 통해 이 연구는 북한이 2013년에 채택했던 ‘핵·경제 병진노선’을 2018년에 ‘경제집중 노선’으로 전환한 이유가 무엇인지 이해하고자 함. 그리고 북한이 목표하고 있는 국가발전 전략과 향후 핵개발과 경제개발에 대한 함의를 도출하고자 함.

2. 자료 개요

- 본 연구에서는 북한의 대표적인 학술지인 『경제연구』, 『정치법률연구』, 『김일성종합대학 학보: 철학, 경제』를 중심으로 분석을 진행함. 이 세 잡지는 각각 경제, 정치·법, 그리고 정치·경제 이론에 있어 대표적인 학술지로 간주됨.

- 『경제연구』는 1956년 4월 10일 평양의 과학백과사전종합출판사에서 분기간행물로 창간되었으며 1964년부터는 사회과학원출판사에서, 1973년 4월부터는 『사회과학』 이론잡지에 통합되었다가 1985년 12월부터 다시 『사회과학』에서 분리되어 『경제연구』로 발간되고 있음.

- 이 잡지는 북한의 경제 현상과 경제 질서를 대내외의 경제학자와 지식인들에게 정당화하는 설득기제로 사용되고 있음으로 북한 경제를 논리적으로 설명하는 과정에서 그 구조와 변화를 추출할 수 있다는 측면에서 유의미한 자료로 간주됨.

- 『정치법률연구』는 2003년 제1호를 시작으로 2017년 제4호까지 발간된 분기간행물임.

- 『정치법률연구』는 15년간 북한 정치의 정당성을 법률적으로 뒷받침하는 역할을 하였고 이는 선군정치의 중요한 수단으로 기능함. 후속으로는 『법률연구』가 발간되고 있음.

- 주요 독자는 정치인, 법률가, 법학자들임.

- 이 자료는 선군정치의 산물인 북한 핵과 경제의 병진노선을 이해하는데 중요한 시사점을 제공해줌.
- 『김일성종합대학 학보: 철학, 경제』는 경제적 현실보다는 이념에 치중되어 있으며 경제적 성과와 이론의 정당성을 강조함.
 - 특히 전통과 세습의 정당화 논리에 초점을 맞추고 있음. 이 과정에서 북한은 변화된 경제 현상을 어떻게 인식하고 있으며 역사적 맥락에서 어떻게 현실을 정당화하고자 하는지를 추적할 수 있음.

3. 분석방법론 및 적용방법¹⁾

3.1 Document-Term Matrix(DTM)와 Term-Document Matrix(TDM)

□ 개요

- 정보검색 이론에서 문서를 단어로 검색 및 접근하기 위하여 문서를 단어의 배열 집합으로 표현한 형태로 설명됨.
- DTM을 통하여 문서별 단어의 출현 빈도를 도출할 수 있고, TDM을 통하여 단어별 문서의 출현 빈도를 알 수 있는 행렬을 구성함.

□ 적용방법

- DTM과 TDM 방법론은 ‘경제연구’, ‘정치법률연구’, ‘학보(철학,경제)’ 자료 내의 핵과 경제 관련 단어와 그 빈도수를 도출하는데 적용함.
- 기존의 방법은 문서별 단어의 매트릭스를 구하는 방식인데, 더 나아가 문장 및 문단을 기준으로 단어 매트릭스를 형성함으로써 문맥 중심의 의미를 파악함.

1) 본 연구에 활용된 텍스트마이닝 방법론들의 자세한 설명은 부록 참고.

3.2 연관 분석

□ 개요

- 앞서 설명한 TDM과 DTM 두 매트릭스를 곱함으로써 정방형 매트릭스를 만들어 단어별 상대적 연관 척도를 시각화하는 분석기법임.

□ 적용방법

- 연관 분석을 통하여 ‘경제연구’, ‘정치법률연구’, ‘학보(철학,경제)’, 자료 내의 핵, 경제와 연관성 높은 단어를 추출함.

3.3 상관 분석

□ 개요

- 상관 계수 및 조정된 결정 계수를 중심으로 모델을 평가하고, 변수에 대한 P-value를 통하여 각각의 변수 유의성에 대해 평가함과 동시에 데이터 범위에 따라 각각의 계수 중 가장 높은 값을 나타내는 모델을 채택함.
- 『경제연구』, 『정치법률연구』, 『학보(철학,경제)』에 등장하는 분기별 핵 단어 출현 빈도(비율)를 변수 1(X_1)로, 분기별 경제 단어 출현 빈도(비율)를 변수 2(X_2) 등으로 설정하여 각 시기별 변수간의 상관관계를 통계적으로 분석함.

4. 주요결과 및 시사점

4.1 “핵”과 “경제” 기간별 출현 빈도 분석

□ 앞서 언급한 ‘경제연구’, ‘정치법률연구’, 그리고 ‘학보(철학·정치)’ 자료에서 “핵” 단어와 “경제” 단어와 이 단어가 포함된 문단이 출현하는 빈도를 분기별로 도출함.

○ “핵”과 “경제”의 출현 빈도를 도출함으로써 두 요소 간의 상대적 중요도를 알아보고 기간 별 추이를 통하여 두 요소 간의 동태를 파악하고자 함.

□ 분석결과

○ 단어 수 기준과 문단 수 기준 분석 결과 경제 단어와 핵 단어의 추이는 각각의 특성을 갖고 있음을 알 수 있음.

- <그림 1>과 <그림 2>에 따르면 경제 단어 수와 문단 수는 2013년 1분기부터 2018년 3분까지의 기간 동안 비교적 변동 폭이 크지 않았던 것으로 나타남.

- 반면 핵 단어 수와 문단 수는 상대적으로 변동이 큰 것으로 나타났으며 특히 2016년 1분기부터 2017년 4분기까지 단어 수와 문단 수가 크게 증가한 것으로 나타남.

○ 두 단어의 출현빈도에 대하여 좀 더 엄밀히 알아보기 위해 두 자료의 상관계수를 도출하며 이때 전 기간 상관관계와 함께 상관성이 커 보이는 2016년 3분기 이후와 이전으로 기간을 나눈 상관계수도 도출함.

- <표 3>에 따르면 두 단어의 빈도 변화 추이 간 상관계수는 전 기간 단어 수 기준 0.4198을 기록하였으며 이는 5% 신뢰수준에서

통계적으로 유의미한 것으로 나타남.

- 또한 2016년 3분기 이후의 기간에는 상관관계가 더욱 커져 상관계수 0.8212를 기록하였으며 이 또한 5% 신뢰수준에서 통계적으로 유의미한 반면 2016년 3분기 이전의 기간에는 상관계수 0.3345로 양의 상관관계를 갖지만 그 값은 상대적으로 낮은 편이며 통계적 유의미성 또한 사라지는 것으로 나타남.

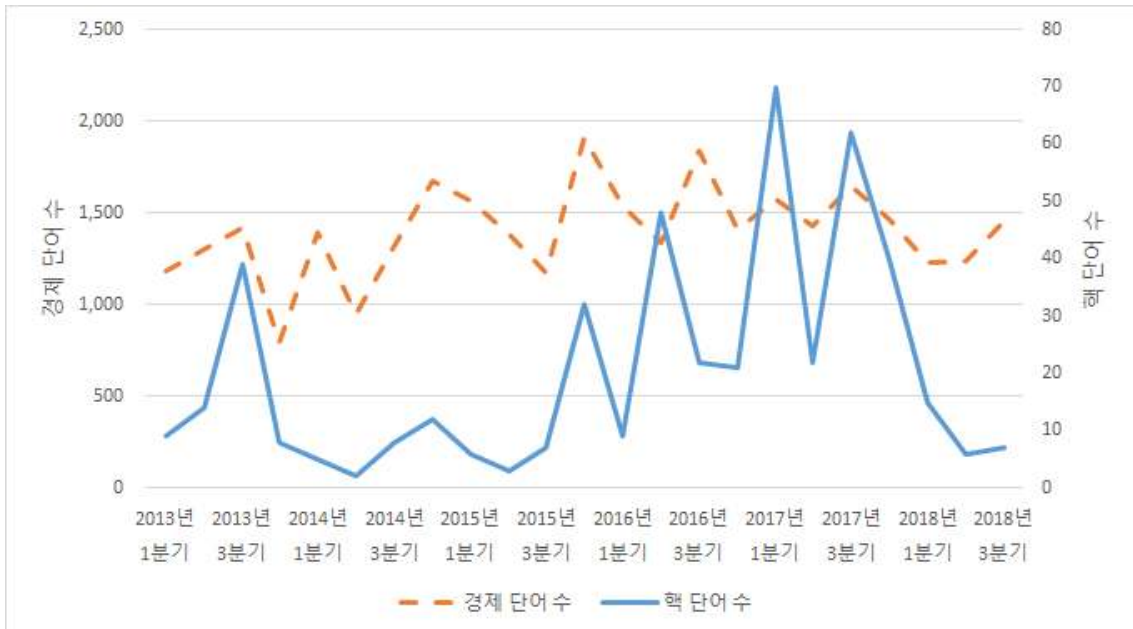
- 문단 수 기준에서도 단어 수의 경우와 유사한 결론을 도출할 수 있음.

○ 요약하면 단어 수와 문단 수 빈도 분석에 따르면 핵과 경제는 2013년 1분기부터 2018년 3분기까지 전체 기간 동안 대체적으로 양의 관계를 보임. 기간을 나누어서 검토하면 2013년 1분기부터 2016년 2분기까지는 독립재, 2016년 3분기부터 2018년 3분기까지는 보완재로 이해됨.

<표 1> 핵, 경제 단어 출현 분기별 빈도 (단어 수 기준)

분기	총 단어 수	핵 단어 수	경제 단어 수	핵 단어 비율	경제 단어 비율
2013.1	80,766	9	1,181	0.00011	0.015
2013.2	84,016	14	1,301	0.00017	0.015
2013.3	79,196	39	1,423	0.00049	0.018
2013.4	61,574	8	792	0.00013	0.013
2014.1	73,416	5	1,388	0.00007	0.019
2014.2	61,764	2	951	0.00003	0.015
2014.3	74,515	8	1,319	0.00011	0.018
2014.4	79,744	12	1,678	0.00015	0.021
2015.1	80,431	6	1,569	0.00007	0.020
2015.2	82,818	3	1,381	0.00004	0.017
2015.3	82,715	7	1,172	0.00008	0.014
2015.4	89,421	32	1,915	0.00036	0.021
2016.1	89,199	9	1,541	0.00010	0.017
2016.2	81,019	48	1,334	0.00059	0.016
2016.3	113,017	22	1,836	0.00019	0.016
2016.4	87,326	21	1,409	0.00024	0.016
2017.1	83,759	70	1,576	0.00084	0.019
2017.2	82,246	22	1,430	0.00027	0.017
2017.3	85,003	62	1,648	0.00073	0.019
2017.4	91,246	40	1,468	0.00044	0.016
2018.1	75,671	15	1,232	0.00020	0.016
2018.2	82,918	6	1,235	0.00007	0.015
2018.3	82,896	7	1,453	0.00008	0.018

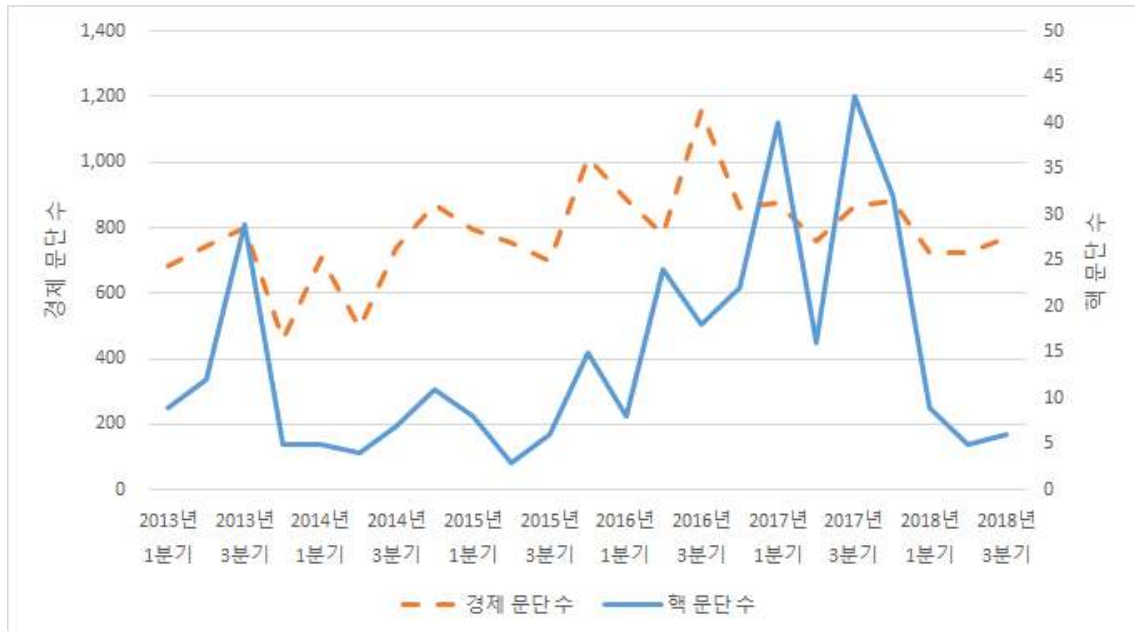
<그림 1> 핵과 경제 단어 분기별 출현 빈도 변화 추이 (단어 수 기준)



<표 2> 핵과 경제 단어 분기별 출현 빈도 및 비율 (문단 수 기준)

구분	총 문단 수	핵 문단 수	경제 문단 수	핵 문단 비율	경제 문단 비율
2013.1	3,656	9	683	0.00246	0.187
2013.2	3,858	12	745	0.00311	0.193
2013.3	3,468	29	799	0.00836	0.230
2013.4	3,076	5	459	0.00163	0.149
2014.1	3,222	5	711	0.00155	0.221
2014.2	2,848	4	494	0.00140	0.173
2014.3	3,383	7	738	0.00207	0.218
2014.4	3,534	11	871	0.00311	0.246
2015.1	3,579	8	794	0.00224	0.222
2015.2	3,529	3	754	0.00085	0.214
2015.3	3,918	6	700	0.00153	0.179
2015.4	4,082	15	1,013	0.00367	0.248
2016.1	3,782	8	889	0.00212	0.235
2016.2	3,675	24	778	0.00653	0.212
2016.3	5,245	18	1,158	0.00343	0.221
2016.4	4,062	22	860	0.00542	0.212
2017.1	3,909	40	879	0.01023	0.225
2017.2	3,574	16	757	0.00448	0.212
2017.3	3,940	43	865	0.01091	0.220
2017.4	4,181	32	884	0.00765	0.211
2018.1	3,453	9	725	0.00261	0.210
2018.2	3,620	5	726	0.00138	0.201
2018.3	3,721	6	771	0.00161	0.207

<그림 2> 핵과 경제 단어 분기별 출현 빈도 변화 추이 (문단 수 기준)



<표 3> 기간별 핵과 경제 단어 출현 빈도 간 상관관계

	기준 분류	전 기간	2016년 3분기 이전	2016년 3분기 이후
상관계수 (p-값)	단어 수	0.4198* (0.0461)	0.3345 (0.2424)	0.8212* (0.0124)
	문단 수	0.4637* (0.0259)	0.4144 (0.1407)	0.8874* (0.0033)

※(*)표시는 신뢰수준 5%에서 통계적으로 유의미하다는 의미임

□ 시사점

- 핵과 경제라는 두 단어와 이 두 단어를 담은 문단의 출현 빈도가 양의 상관관계를 가지며 그 추이가 전체적으로 유사한 모습을 보이고 있다는 의미는 모든 출판물에 대한 검열을 실시하는 북한의 상황을 감안할 때 북한이 2013년 핵·경제 병진노선을 천명한 후 북한의 대표적 학술지가 이러한 당의 기본 정책방향을 반영하고 있는 것으로 추측됨
- 그러나 핵과 경제의 상관계수가 2016년 3분기 이후에 크게 높아지고 통계적으로도 유의해 진 사실은 2016년을 시작으로 2017년 말까지 이어진 핵과 대륙간탄도미사일(ICBM) 실험 등 북한의 일련의 군사적 도발과 이에 대한 국제 사회의 대응과 연관된 것으로 해석됨.
 - 군사·정치적 긴장상황이 높아질수록 북한은 국제사회의 비난여론에 노출됨으로 대내적으로 핵 개발의 명분을 더욱 공고히 할 필요가 있었을 것임. 이 일환으로 학술지 등 출판물에서의 핵과 경제, 특히 핵의 출현빈도가 높아진 것으로 추측됨.

4.2 “핵”과 “경제” 출현 성격 분석

- 북한 당국의 핵과 경제에 대한 이해를 파악하기 위하여 분석 문건 기준 핵과 경제의 동시출현 비율 추이와 핵과 경제 각각의 동시출현 단어 군에서 각각의 출현 빈도를 분석함.

4.2.1 “핵”과 “경제” 동시출현 비율

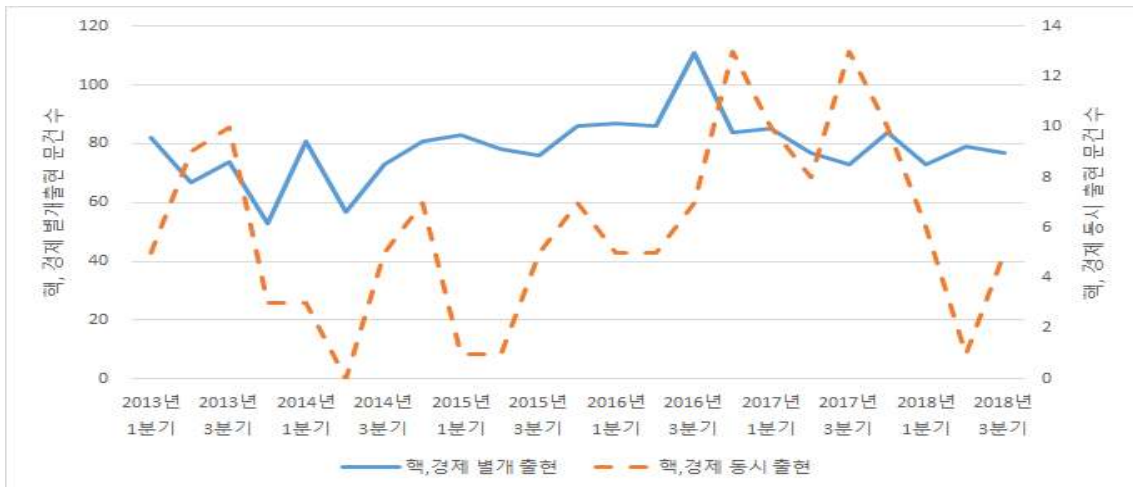
□ 분석결과

- 아래 <표 4>에 따르면 총 분석 문건 수 중 경제 단어가 출현한 문건 수(2번 열)는 핵 단어가 출현한 문건 수(3번 열)에 비하여 월등히 많은 것으로 나타남.
- 핵과 경제 단어가 동시에 출현한 문건 수(5번 열) 보다 핵과 경제가 별개로 출현한 문건의 수(6번 열)가 월등히 많은 것으로 나타남.
- 핵이 출현한 문건 수(3번 열) 중 경제가 함께 출현한 문건 수(5번 열)의 비율은 상당히 높은 반면 경제가 출현한 문건 수(2번 열) 중에 핵이 함께 출현한 문건 수(5번 열)는 적은 것으로 나타남.
- 핵과 경제가 출현한 모든 문건 수(4번 열)중 핵과 경제가 함께 출현한 문건 수(5번 열)의 비율(7번 열)은 최저 0%에서 최고 약 15%로 높지 않음.
- 핵과 경제가 별개로 출현한 문건 수와 동시에 출현한 문건 수의 추이 (<그림 3>)를 비교해보면 핵과 경제가 별개로 출현한 문건 수의 경우 비교적 변동 폭이 적음. 반면 핵과 경제가 동시에 출현한 문건 수의 경우 2016년 3분기부터 비중이 크게 상승하며 2017년 3분기를 기점으로 하락하지만 2018년 3분기에는 다시 상승하는 것으로 나타남.

<표 4> 핵과 경제 동시출현 비율 분석 결과

구분	(1) 총 문건 수	(2) 경제 문건 수	(3) 핵 문건 수	(4) 핵, 경제 출현 문건 총 수	(5) 동시 출현 문건 수	(6) 별개 출현 문건 수	(7) (5)/(4)
2013.1	112	86	6	87	5	82	0.05747
2013.2	112	75	10	76	9	67	0.11842
2013.3	107	83	11	84	10	74	0.11905
2013.4	86	55	4	56	3	53	0.05357
2014.1	105	82	5	84	3	81	0.03571
2014.2	87	55	2	57	0	57	0.00000
2014.3	103	77	6	78	5	73	0.06410
2014.4	109	86	9	88	7	81	0.07955
2015.1	112	80	5	84	1	83	0.01190
2015.2	110	77	3	79	1	78	0.01266
2015.3	122	80	6	81	5	76	0.06173
2015.4	125	92	8	93	7	86	0.07527
2016.1	132	91	6	92	5	87	0.05435
2016.2	119	89	7	91	5	86	0.05495
2016.3	159	114	11	118	7	111	0.05932
2016.4	138	95	15	97	13	84	0.13402
2017.1	127	90	15	95	10	85	0.10526
2017.2	117	84	9	85	8	77	0.09412
2017.3	114	85	14	86	13	73	0.15116
2017.4	128	90	14	94	10	84	0.10638
2018.1	108	78	7	79	6	73	0.07595
2018.2	109	80	1	80	1	79	0.01250
2018.3	109	82	5	82	5	77	0.06098

<그림 3> 핵과 경제가 같이 출현한 문건 비중 분기별 추이 (문건 수 기준)



4.2.2 “핵”과 “경제” 출현 관련성 분석

□ 분석결과

○ <표 5>에 따르면 핵 관련단어군 중 경제 단어의 순위는 전 분기에서 경제 관련단어군 중 핵 단어의 순위보다 월등히 높은 수준임이 확인됨.

- 경제 관련단어군 중 핵 단어는 많은 경우 1,000위를 넘는 반면 핵 관련단어군 중 경제 단어는 최저 순위가 178위일 만큼 두 경우 사이에는 많은 차이가 존재함.

<표 5> 핵과 경제 관련단어군중 서로의 순위

분기	핵 관련단어군 중 경제 순위	경제 관련단어군 중 핵 순위
2013.1	6	608
2013.2	9	854
2013.3	29	710
2013.4	173	1628
2014.1	20	1716
2014.2	107	1911
2014.3	85	1804
2014.4	45	1938
2015.1	76	2491
2015.2	16	1566
2015.3	148	2522
2015.4	28	900
2016.1	81	2306
2016.2	178	1529
2016.3	65	1495
2016.4	1	637
2017.1	52	808
2017.2	31	1291
2017.3	15	894
2017.4	20	875
2018.1	61	1279
2018.2	31	1723
2018.3	2	1204

※ 순위는 각 단어의 절대적 빈도(frequency)가 아닌 동시 출현 단어 매트릭스가 도출되었을 때, 행렬 곱 결과 수치(sum)를 기준으로 순위화를 진행한 결과임.

□ 시사점

- <표 4>의 결과 경제가 출현한 문건 수가 핵이 출현한 문건 수 보다 많은 결과는 분석대상에 경제 분야 학술지가 포함되어 있고 대중 잡지와 대비하여 학술지가 갖는 성격에 일정 부분 기인한 것으로 보임.
 - 경제의 경우 학술적 설명이 용이한 반면, 핵의 경우 학술적 설명보다는 정치적 설명이 더욱 용이할 것으로 판단됨. 따라서 이 연구에서 사용하는 분석대상이 학술지이기 때문에 핵에 비해 경제가 출현한 빈도수가 높을 수 있음을 고려해야 할 것임.
- <표 4>의 결과 핵과 경제가 동시에 출현하는 문건 수와 별개로 출현하는 문건 수를 비교해보면 두 단어가 별개로 출현하는 문건 수가 월등히 많은 것으로 분석됨.
 - 이는 핵·경제 병진노선이 핵과 경제를 잇는 논리적 고리에 의해 연결되어 있기보다 그 각각의 정당성과 효용을 주장하는 경향이 있음을 시사함.
- <표 5>에 따르면 핵을 설명하는데 있어서는 경제를 적극 활용하고 있는 반면, 경제를 설명하는데 있어서 핵을 사용하는 경우는 드물. 이 결과는 핵과 경제는 상호보완적이기 보다는 핵에 있어 경제는 보완적인 반면 경제에 있어서는 핵의 영향이 없는 관계, 즉 일방적 보완관계에 있다고 판단됨.
 - 이는 북한 주민의 시각에서 핵보다는 경제의 효용가치가 더 높을 수 있음을 학술지에서도 인지하고 있는 것으로 해석할 수 있음. 즉 핵만 으로서는 북한 주민의 생존이 어렵기 때문에 경제를 함께 강조하는 경향이 있는 반면 경제를 설명할 때는 핵이 불필요함을 알거나 이 둘을 연결하는 내적인 연관성을 찾지 못함에 기인하는 것으로 파악됨.

4.3 “핵”과 “경제” 동시출현 단어군 분석

□ 핵과 경제 단어가 출현할 시 동시에 출현할 확률이 높은 관련 단어군 분석을 통하여 북한 당국이 이해하고 있는 핵과 경제의 개념과 의미에 대하여 알아보고자 함.

□ 분석결과

○ <표 6>에 따르면 핵과 경제의 동시출현 단어군 순위화 분석결과 두 단어의 관련단어군 상위 25개의 구성은 8개 단어를 제외하고 모두 다른 것으로 나타남.

- 핵의 경우 주로 기술, 전쟁, 에네르기, 무기, 세계, 나라, 강국, 미 제 등의 단어들로 설명하고 있음.

- 경제의 경우는 발전, 사회주의, 인민, 나라, 관리, 건설, 국가, 생산 등의 단어들로 대부분 설명하고 있음.

○ 총 8개의 단어는 경제관련 단어군과 핵관련 단어군 모두에 포함되어 있는 것으로 분석됨.

○ 두 관련단어군 모두에 포함된 단어는 발전, 사회주의, 인민, 나라, 건설, 강국, 과학기술, 기술 등임.

- 이 중 발전, 사회주의, 인민, 나라, 건설, 강국은 핵 개발과 경제 발전의 결과를 보여주는 관점에서 연결되어 있는 것으로 판단됨.

- 과학기술과 기술은 핵과 경제발전에 다 적용 가능한 공통된 수단으로 이해됨.

- 그러나 핵과 경제발전을 연결하는 논리적 고리는 여전히 발견되지 않음.

- 공통 관련단어가 8개 존재함에도 그 순위는 경제 관련 단어군과 핵 관련 단어군에서 비교적 명확한 차이를 보임.
- <표 7>에 따르면 경제 관련 단어군에서 각각 1위, 2위, 3위를 차지한 발전, 사회주의, 인민 등의 단어는 핵 관련 단어군에서는 비교적 낮은 순위인 21위와 16위, 11위를 기록함.
- 반면 핵 관련 단어군에서 1위를 기록한 “기술”은 경제 관련 단어에서 17위로 낮은 순위를 기록하는 등, 핵 관련 단어군에서 높은 순위를 기록한 단어들은 경제 관련 단어군에서는 비교적 낮은 순위를 기록함.
- 또한 핵 관련 단어군에서 상위권을 차지한 기타 단어들은 경제 관련 단어군에서 찾아볼 수 없음.

<표 6> 핵과 경제 관련 단어군

경제 관련 단어				핵 관련 단어			
	term	sum	freq		term	sum	freq
1	발전	14,229	19,341	1	기술	259	299
2	사회주의	9,345	17,713	2	전쟁	160	116
3	인민	9,206	24,163	3	에네르기	138	116
4	나라	9,088	19,420	4	무기	133	85
5	관리	8,863	8,938	5	세계	126	98
6	건설	8,707	14,967	6	나라	122	107
7	국가	6,600	13,471	7	강국	108	94
8	생산	6,392	17,799	8	시험	97	60
9	요구	5,456	11,834	9	공화국	95	70
10	사업	5,428	12,496	10	미제	94	76
11	강국	5,422	5,993	11	인민	86	108
12	개발	4,615	5,517	12	무력	86	66
13	계획	4,409	5,981	13	우주	82	71
14	활동	4,374	8,259	14	나노기술	78	88
15	사회	4,294	16,540	15	과학기술	75	114
16	과학기술	4,142	8,917	16	사회주의	74	88
17	기술	4,139	8,748	17	정보기술	73	87
18	보장	3,690	7,560	18	책동	72	58
19	방법	3,571	7,402	19	건설	72	83
20	지도	3,557	3,323	20	미국	69	45
21	과학	3,131	6,469	21	발전	68	95
22	기초	2,979	6,574	22	위력	64	60
23	전략	2,912	3,577	23	세력	64	53
24	세계	2,483	6,663	24	첨단	63	93
25	자본주의	2,479	4,722	25	경제	62	74

※ 순위는 각 단어의 절대적 빈도(frequency)가 아닌 동시 출현 단어 매트릭스가 도출되었을 때, 행렬곱 결과 수치(sum)를 기준으로 순위화를 진행한 결과임. 안의 단어는 경제관련, 핵 관련 단어군 모두에 포함된 단어임.

<표 7> 경제·핵 관련 단어 별 공통 관련 단어 순위

공통 관련 단어	경제 관련 단어군 순위	핵 관련 단어군 순위
발전	1위	21위
사회주의	2위	16위
인민	3위	11위
나라	4위	6위
건설	6위	19위
강국	11위	7위
과학기술	16위	15위
기술	17위	1위

□ 시사점

- 핵과 경제 관련 단어군 순위화 결과는 25개의 관련단어군 중 8개만이 두 단어군 모두에 공통으로 포함되어 있어 두 단어가 기본적으로 서로 다른 개념으로 쓰이고 있다는 것을 시사함.
 - 또한 공통단어군 내에서도 공통단어들이 핵과 경제 각각의 관련단어군에서 차지하고 있는 순위에 비교적 큰 차이가 존재함은 핵과 경제를 독립된 개념으로 인식하고 있음을 시사함.
- 핵과 경제 각각의 관련단어군의 단어들의 성격을 살펴보면, 핵 관련 단어의 경우 기술, 전쟁, 무기, 강국, 미제, 과학기술 등 주로 기술관련 단어와 전쟁, 미제 등 북한이 주장하는 핵개발 이유로 구성되어 있음. 반면 경제 관련단어의 경우 발전, 사회주의, 인민, 건설, 생산, 개발 등 주로 경제발전과 관련된 단어들로 구성되어 이루고 있음.
- 요약하면 핵은 미국으로부터의 전쟁위협 등 국제정치적 위협으로부터 인민을 지키기 위한 수단으로 설명하는 경향이 있음. 경제의 경우는 경제발전을 위해 필요한 요소들을 설명을 하고 있어 두 단어 사이에는 기본적으로 이질성이 존재함을 시사함. 그리고 이 두 단어를 연결시켜 주는 내적인 연관 고리도 제시되어 있지 않음.

4.4 “핵”, “경제”와 한반도 사건들과의 관계 분석

- 전체 문서의 모든 단어에서 핵 관련단어군과 경제 관련단어군의 비중이 핵실험 등 한반도 평화에 부정적인 영향을 미친 사건들과 남북 간 회담, 북미 간 회담 등 한반도 평화에 긍정적인 영향을 미친 사건들 간의 어떠한 관계가 있는지 로짓(Logit) 회귀분석을 통하여 알아봄.
- 우선 전체 단어 수에서 <표 5>의 핵 관련단어군과 경제 관련단어군이 차지하는 비중 추이를 구하여 독립변수로 설정하고 앞서 다른 “경제” 단어와 “핵”단어의 단순 출현 빈도 또한 독립변수로 활용함.

- 추가로, 핵 관련단어군 중에는 경제 관련단어군이 다수 포함되어 있으므로 순수 핵 관련단어군을 도출하기 위하여 핵 관련단어군 중에서 경제 관련단어군을 제거하여 “순수 핵 관련단어군”을 도출하여 독립변수로 활용함.²⁾

○ 핵과 경제 두 관련단어군의 분기별 출현 비중 추이와 순수 핵 관련단어군 출현 비중 추이는 다음과 같음.

<그림 4> 핵,경제 관련단어군 출현 비중과 순수 핵 관련단어군 출현 비중 추이



2) 핵 관련단어군에서 경제 관련단어군을 제거하는 방법은 두 관련단어군의 평균값을 구하여 경제 관련단어군 평균값을 핵 관련단어군 평균값을 나누고 여기에 해당 분기의 핵 관련단어군을 곱하여 도출함. ($\frac{\text{평균 경제 관련단어군 출현 비중}}{\text{평균 핵 관련단어군 출현 비중}} \times \text{분기별 핵 관련단어군 출현 비중}$).

- 전체적으로 경제 관련단어군의 비중이 핵 관련 단어군의 비중보다 전 기간 높은 수준을 유지함.
- 이와 함께 북한의 핵실험 등 한반도 평화에 부정적 영향을 미친 사건과 남북고위급접촉 및 남북정상회담, 북미고위급접촉 및 북미정상회담 등 평화적 사건들의 더미변수를 생성하여 종속변수로 설정함.
 - 부정적 사건 변수에서 북한이 핵실험, 또는 탄도미사일을 시행한 분기는 1, 기타 분기는 0의 값을 갖는 더미변수(Dummy variable)를 생성함.
 - 긍정적 사건 변수에서 남북, 북미 간 고위급회담과 정상회담이 진행된 분기는 1, 기타 분기는 0의 값을 갖는 더미변수를 생성함.
 - 위의 기준에 따라 도출한 긍정적 사건과 부정적 사건 변수 내용은 아래 <표 7>과 같음.
- 추가로, 한반도 평화에 부정적인 영향을 미친 사건들이 이후 학술지의 핵과 경제 단어의 단순 출현빈도, 관련단어군 출현 비중, 순수 핵 관련단어군 출현 비중에 어떠한 영향을 미쳤는지 분석함.
 - 부정적 사건, 즉 핵실험, 탄도미사일 발사 등의 사건이 이후 핵 단어 및 관련단어 출현에 미치는 영향을 분석하여 학술지 자료가 부정적 사건 발생 이후 해당 사건에 대한 정당화를 시도하고 있는지 살펴봄.

<표 8> 종속변수 구성 내용

	부정적 사건 변수		긍정적 사건 변수	
	사건	변수값	사건	변수값
2013.1	제3차 핵실험	1	-	0
2013.2	단거리 미사일 발사	1	-	0
2013.3	-	0	-	0
2013.4	-	0	-	0
2014.1	-	0	정치 고위급회담	1
2014.2	중거리 미사일 발사	1	-	0
2014.3	-	0	-	0
2014.4	-	0	군사 고위급회담	1
2015.1	스커드 계열 미사일 발사	1	-	0
2015.2	KN 계열 미사일 발사	1	-	0
2015.3	-	0	정치 고위급회담	1
2015.4	-	0	정치 고위급회담	1
2016.1	제4차 핵실험	1	-	0
2016.2	중거리 탄도미사일 발사	1	-	0
2016.3	잠수함탄도미사일 발사	1	-	0
2016.4	제5차 핵실험	1	-	0
2017.1	탄도 미사일 발사	1	-	0
2017.2	탄도 미사일 발사	1	-	0
2017.3	제6차 핵실험	1	-	0
2017.4	탄도 미사일 발사	1	-	0
2018.1	-	0	북 고위급대표단 방남, 대북특사 방북	1
2018.2	-	0	남북정상회담 북미정상회담	1
2018.3	-	0	남북정상회담	1

※출처: 통일부 홈페이지 <남북관계 주요일지> 각 월호, 통일부 남북회담본부 홈페이지 <남북회담정보>

- 또한 핵과 경제의 관련단어군과 부정/긍정적 사건, 즉 종속변수와 독립변수 사이에는 동 분기 발생하는 영향도 있을 수 있지만, 시차를 두고 영향이 발생할 수 있음을 고려하여 독립변수의 시차변수(Lagged

variable)를 2기, 즉 이전 2분기까지 생성하여 분석을 진행함.

□ 분석결과

- <표 9>의 결과에 따르면 단순 단어 출현 비중을 독립변수로 설정한 모델(1번 열, 4번열)은 부정적 사건과 긍정적 사건 모두에 통계적 유의미성을 보이지 않음.
- 순수 핵 관련단어군 출현 비중의 1기 시차변수는 부정적 사건에 5% 신뢰수준에서 통계적 유의미성을 갖는 양(+)의 영향을 미치는 것으로 나타났음(3번 열). 반면 긍정적 사건에 미치는 영향은 통계적 유의미성을 갖지 못함(6번 열).
 - 즉, 이전 분기 순수 핵 관련단어군 출현 비중이 높으면 높을수록 현재 분기에 핵실험, 미사일 발사 등 부정적 사건이 일어날 확률이 높아지는 것으로 추정됨.
- 경제 관련단어군 출현 비중의 1기 시차변수는 긍정적 사건에 5% 신뢰수준에서 통계적 유의미성을 갖는 양(+)의 영향을 미치는 것으로 나타남(5번 열).
 - 이전 분기의 경제 관련단어군 출현 비중이 높으면 높을수록 현재 분기에 긍정적 사건이 일어날 확률이 높아지는 것으로 추정됨.
- 핵과 경제 각 단어의 단순 출현빈도는 긍정/부정 사건에 영향을 주지 못하는 것으로 나타나는 반면, 1기 시차변수(Lagged Variable)는 각 사건에 유의미한 영향을 미치는 것으로 나타남.
 - 1기 시차 변수와 사건 발생 사이의 관계가 논리적으로 설명될 수 있을 지는 확실하지 않음. 핵과 미사일 실험의 분위기나 회담 분위기가 사전에 탐지되어 이에 미리 대응하려는 경향 때문에 이런 관계가 발생했을 가능성도 완전히 배제할 수 없음. 그러나 우연한 통계적 현상일 가능성도 존재함.

- <표 10>의 결과에 따르면 한반도에 부정적 사건이 발생할 시 해당분기의 핵 단어 단순 출현빈도가 증가하며, 2분기 이전의 부정적 사건 발생도 해당분기의 핵 단어 단순 출현빈도를 증가함. 또한 부정적 사건 발생 분기의 경제 관련단어군의 출현 비중을 감소시킴.

□ 시사점

- 학술지에 출현하는 핵과 경제 각 단어의 관련단어군 출현 비중은 한반도 평화의 긍정/부정적인 사건과 밀접한 연관이 있는 것으로 보임.
- 한반도 평화에 부정적인 사건은 핵 단어의 단순출현 빈도를 증가시키는 것으로 나타나 핵 실험, 탄도미사일 발사 등의 사건 이후 그러한 사건들의 정당화가 어느 정도 이루어지고 있는 것으로 보임.
- 반면 부정적 사건이 발생할 시 경제 관련단어군의 출현비중은 감소시키는 것으로 나타나 핵실험, 탄도미사일 발사 등의 사건이 일어날 시, 경제부문에 대한 관심은 상대적으로 낮아지는 것으로 보임.

<표 9> 시계열 로짓 회귀분석 결과표

	종속변수					
	부정적 사건			긍정적 사건		
	(1)	(2)	(3)	(4)	(5)	(6)
	“핵” 단순 출현빈도	핵 관련단어 군 비중	순수 “핵” 관련단어 군 비중	“경제” 단순 출현빈도	경제 관련단어 군 비중	순수 “핵” 관련단어 군 비중
핵 단어 출현 빈도 _t	0.1231 (0.1440)					
핵 단어 출현 빈도 _{t-1}	0.1483 (0.1561)					
핵 단어 출현 빈도 _{t-2}	-0.0150 (0.1415)					
핵 관련단어군 비중 _t		-152.5 (102.23)				
핵 관련단어군 비중 _{t-1}		-0.404 (84.11)				
핵 관련단어군 비중 _{t-2}		15.38 (63.82)				
순수 핵 관련단어군 비중 _t			7.742 (72.85)			-27.75 (102.54)
순수 핵 관련단어군 비중 _{t-1}			138.17* (61.24)			-147.20 (75.38)
순수 핵 관련단어군 비중 _{t-2}			38.99 (68.55)			-12.00 (76.30)
경제 단어 출현 빈도 _t				0.0447 (0.0415)		
경제 단어 출현 빈도 _{t-1}				-0.0697 (0.0543)		
경제 단어 출현 빈도 _{t-2}				-0.0641 (0.0472)		
경제 관련단어군 비중 _t					98.88 (86.22)	
경제 관련단어군 비중 _{t-1}					176.10* (70.07)	
경제 관련단어군 비중 _{t-2}					59.29 (69.22)	
R-squared	0.1349	0.1120	0.1665	0.1663	0.2298	0.1665
관측치	21	21	21	21	21	21

* 괄호 안의 수치는 강건표준오차(Robust Standard Error)임. (*) 표시는 5% 신뢰수준에서 유의미함을 의미하며 굵은체로 표시함. 각 종속변수마다 모든 독립변수의 각 t 시점별로 각각 추정된 결과도 위의 결과와 유의미성과 계수에 큰 차이가 없음.

<표 10> OLS 회귀분석 결과표

	종속변수				
	(1)	(2)	(3)	(4)	(5)
	“핵” 단순 출현빈도	핵 관련단어군 비중	순수 “핵” 관련단어군 비중	“경제” 단순 출현빈도	“경제” 관련단어군 비중
부정적 사건 _t	2.800* (1.206)	-0.0053 (0.0035)	0.0006 (0.0037)	6.180 (5.171)	-0.0086* (0.0033)
부정적 사건 _{t-1}	2.586 (1.437)	0.0026 (0.0036)	0.0006 (0.0052)	4.866 (6.078)	0.0034 (0.0036)
부정적 사건 _{t-2}	4.300** (1.482)	-0.0007 (0.0034)	0.0020 (0.0048)	5.880 (5.233)	-0.0030 (0.0035)
R-squared	0.5827	0.006	0.0091	0.2029	0.0070
관측치	21	21	21	21	21

※ 괄호 안의 수치는 강건표준오차(Robust Standard Error)임. (*) 표시는 5% 신뢰수준, (**) 표시는 1% 신뢰수준에서 유의미함을 의미하며 굵은체로 표시함. 각 종속변수마다 모든 독립변수의 각 t 시점별로 각각 추정된 결과도 위의 결과와 유의미성과 계수에 큰 차이가 없음.

5. 결론

- 북한 학술지를 통해 분석한 결과, 핵과 경제는 내적인 연결 고리라는 측면에서는 독립된 개념으로 인식되고 있음. 핵과 경제는 공통의 목표에 의해 연결되어 있을 뿐 이 둘을 연결하는 내적 연관성은 보이지 않음.
 - 다만 핵과 경제를 잇는 연관성을 미제국주의의 적대시정책에 두고 있는 것으로 핵 경제 병진노선의 수동적 성격이 파악됨.
- 핵과 경제를 이해하는 데 있어서 핵에는 경제가 보완적이지만 경제에는 핵이 필요하지 않은 일방적 보완관계라고 인식하는 것으로 판단됨. 이는 암묵적으로 핵만으로써는 주민으로부터의 정당성 확보가 어렵다는 사실을 인정하는 자기인식이 반영된 것으로 보임.

- 핵과 경제 단어의 출현 빈도 추이는 비슷한 양상을 보이며 이는 출판물에 대한 검열을 실시하는 북한 정권의 특성을 고려했을 때, 두 단어 간의 이질성에도 불구하고 학술지에 대한 검열을 통해 핵·경제 병진노선이라는 국가 정책의 방향을 주입한 것으로 해석할 수 있음.
- 1분기 이전 핵 관련단어의 출현 비중이 높아질수록 핵실험, 미사일 발사 등 한반도 평화에 부정적 사건이 일어날 확률이 높아지며, 1분기 이전 경제 관련단어의 출현 비중이 높아질수록 남북 간의 대화 등 한반도에 평화적 사건이 발생할 확률이 높아지는 것으로 나타남.
- 한반도 평화에 부정적인 사건 발생은 이후 핵 단어 출현 빈도를 증가시키는 효과가 있는 반면, 경제 관련단어의 출현 비중은 감소시키는 효과를 보임. 이는 핵과 미사일 실험의 사후적 정당성 확보를 위한 것으로 판단됨.

부록: 분석방법론

1. 동시출현(co-occurrence) 빈도 분석

□ 개요

- 다수의 문서에서 등장하는 각 단어들의 빈도를 행렬로 표현한 것으로, 문서별 단어의 출현 빈도를 알 수 있음.
- n 개의 문서(document)에 총 m 가지의 단어(term)가 존재한다고 가정할 때, 행렬의 각 행에 등장하는 m 가지의 단어를 제시하고 각 열에는 등장하는 n 개의 문서를 대응시키며, 행렬의 각 원소는 특정 문서에서 나타나는 특정 단어의 빈도를 표시하는 규칙임.
- 단어 문서 행렬(Term-Document Matrix, TDM)를 통하여 단어별 문서의 출현 빈도를, 문서 단어 행렬(Document-Term Matrix, DTM)을 통하여 문서별 단어의 출현 빈도를 알 수 있음.
- TDM은 다량의 문서를 단어와 단어, 단어와 문서, 문서와 문서 간의 동시출현(co-occurrence)과 빈도(frequency)로 이해하는 방법임.
- 단어들 사이에 나타나는 동시출현 관계를 찾은 후 그 빈도를 저장하는 행렬을 구성하는데 이를 동시출현 행렬(Co-occurrence Matrix)이라 하며 이 행렬을 기본 입력으로 이용하여 키워드 추출을 수행함.

<그림 5> TDM, DTM의 형태

	Tweet 1	Tweet 2	Tweet 3	...	Tweet N
Term 1	0	0	0	0	0
Term 2	1	1	0	0	0
Term 3	1	0	0	0	0
...	0	0	3	1	1
Term M	0	0	0	1	0

Term Document Matrix (TDM)

	Term 1	Term 2	Term 3	...	Term M
Tweet 1	0	1	1	0	0
Tweet 2	0	1	0	0	0
Tweet 3	0	0	0	3	0
...	0	0	0	1	1
Tweet N	0	0	0	1	0

Document Term Matrix (DTM)

□ 적용방법

- 한국어로 동시출현 행렬을 만들 때에는 조사를 제외한 체언/용언 등의 형태소만 남긴 후 만드는 것이 일반적임.

<예 시>

문서1) 나는 학교에 간다. 나의 학교는 멀다.
 (형태소 분석 결과) >> 나, 집, 가다, 나, 집, 멀다.

문서2) 학교에 친구가 있다.
 (형태소 분석 결과) >> 학교, 친구, 있다.

- 두 개의 문장이 주어지고 명사만을 추출하는 경우 문장 1과 문장 2에서 서로 동시출현 관계를 맺는 단어가 발생함.
- 형태소 분석 결과에 따른 TDM 및 동시출현 빈도분석 결과는 다음과 같음.

<표 11> TDM 및 동시출현 빈도분석 표

단어	문서 1	문서 2	빈도	동시출현
가다	1	0	1	0
나	2	0	2	1
친구	0	1	1	0
멀다	1	0	1	0
있다	0	1	1	0
학교	2	1	3	2

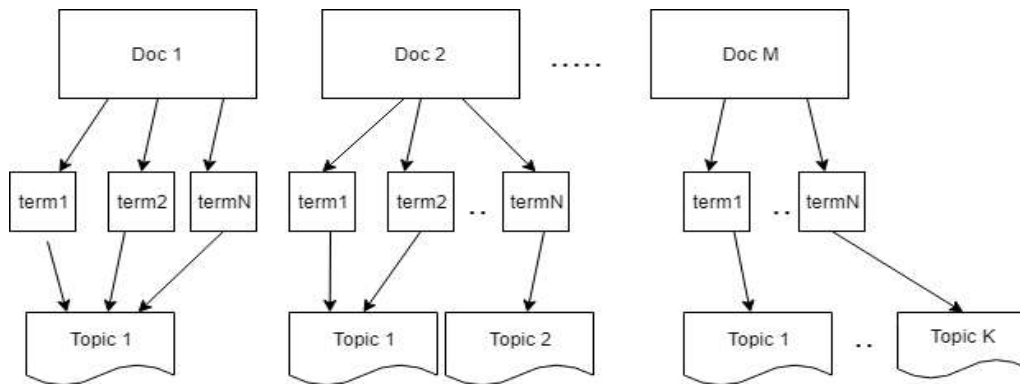
- TDM 또는 DTM의 원소는 특정 단어가 특정 문서에서 사용되었는지 여부와 사용 빈도를 제시함.
- “동시출현”은 두 개 이상의 단어가 하나의 문서에 동시에 나타나거나, 하나의 단어가 여러 개의 문서에 동시에 나타남을 의미함.
- “빈도”는 특정 문서 또는 문서 전체에 특정 단어가 얼마나 자주 나타남을 의미함.
- 일반적으로 하나의 문서에서 자주 사용된 단어는 그렇지 않은 단어보다 더 중요하다고 인식되며, 하나의 문서에서 두 개의 단어가 동시에 나타나는 빈도가 높으면 두 단어는 연관성이 높다고 보는 것이 일반적임.
- 특정 문서에서 나타나는 단어의 빈도는 해당 단어의 중요성을 나타내지만, 특정 문서의 길이가 다른 문서의 길이보다 매우 길면 빈도는 중요성보다는 단순히 특정 문서의 길이가 다른 문서보다 길다는 것을 의미할 수도 있음.
- TDM 또는 DTM은 동시출현과 빈도를 사용하여 특정 문서에서 특정 단어가 얼마나 많이 쓰이는지, 동일 문서에서 특정 단어가 어떠한 단어와 같이 쓰이는지, 특정 단어가 얼마나 많은 문서에서 나타나는지를 통해 다량의 텍스트 자료를 분석할 수 있음.

2. 토픽 모델링(Topic modeling)

□ 개요

- DTM을 이용하여 다량의 문서에서 사용된 주요 단어, 즉 주제어들의 빈도 및 동시출현 패턴을 바탕으로 문서 자료에 나타나는 다수의 주제를 자동으로 추출하여 분석하는 기법임.
- 토픽 모델링에는 일반적으로 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)이 사용됨.

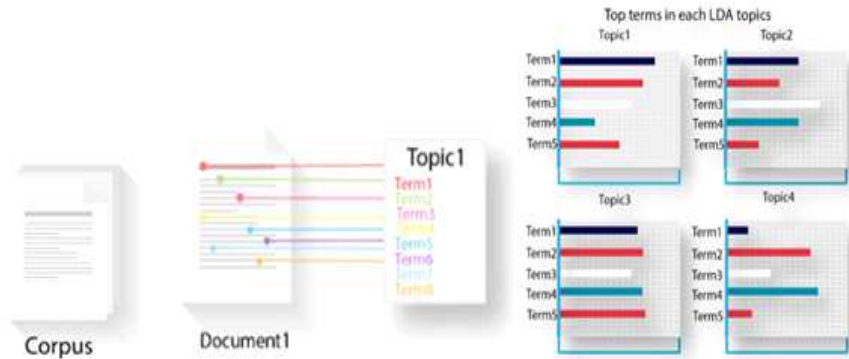
<그림 6> 토픽 모델링 개념



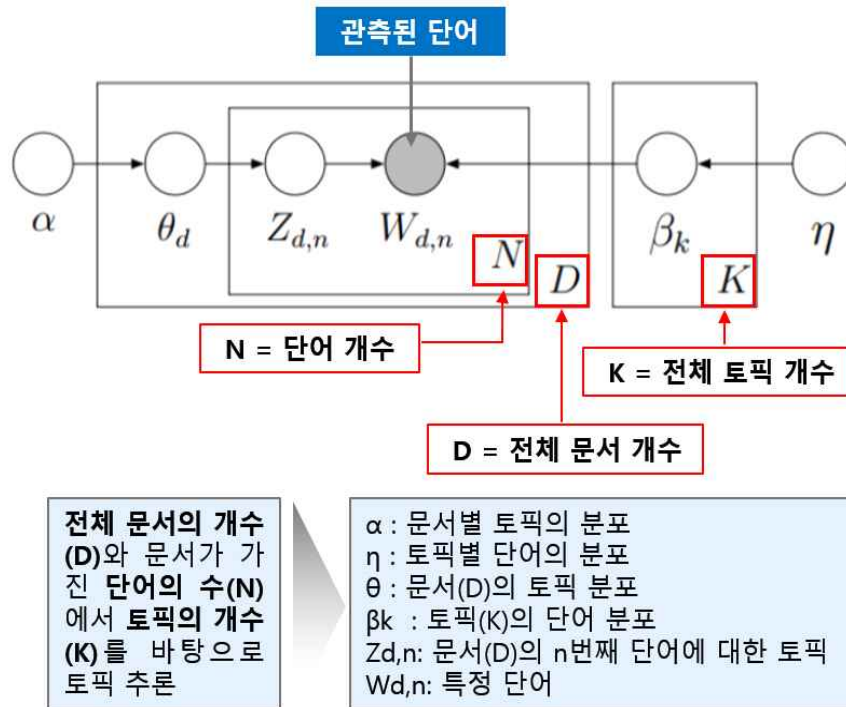
□ 분석기법

- LDA는 전체 문서에서 어떤 주제들이 존재하는지 알아보기 위해 토픽별로 단어의 빈도, 문서별 토픽의 비중을 확률모형으로 나타내어 추정하는 분석기법임.

<그림 7> LDA 분석 과정



<그림 8> LDA 그래프 모델



○ 문서집단(Corpus)에서 토픽을 구분한 후 각 토픽을 대표하는 단어를 중요도의 순서에 따라 추출함.

○ 각 문서를 토픽별로 매칭하여 확률값을 계산함.

○ LDA의 주요 가정은 다음과 같음.

- 문서들은 관측되지 않는 다수의 주제로 구성되어 있으며, 하나의 문서에서 다수의 주제가 나타날 수 있다.

예) 언론 기사에서 “대통령”이라는 단어는 정치면, 사회면, 경제면 등에서 등장

- 하나의 문서에 나타나는 다수의 주제는 구성비율로 나타낼 수 있으며 이 비율은 디리클레 분포를 따른다.

예) “대통령”이 문화공연장을 방문한 기사의 경우, “문화예술” 또는 “정치” 주제로 볼 수 있으며, 하나의 언론 기사에 다양한 주제가 나타날 수 있고, 이것이 상대적 비율로 표시될 수 있다.

- 각 주제는 다양한 단어로 표현되며, 특정 단어가 특정 주제에 따른 확률은 다항 분포에 따른다.

- 즉, 다량의 문서에 사전적으로 주어지는 “주제”가 결국에는 자주(높은 빈도), 동시에 나타나는(높은 동시출현) 단어들의 결합으로 정의된다.

예) 대통령, 국회의원, 정당 등의 단어가 빈번하게 나오는 문서의 주제는 “정치”가 될 확률이 높고, GDP, 기업, 임금, 경기 등의 단어가 빈번하게 나오는 문서의 주제는 “경제”가 될 확률이 높다.

참고문헌

문혜정·손원·원중호, 2017, “텍스트 마이닝 기법을 이용한 경제심리 관련 문서 분류”, BOK 국민계정리뷰.

김수현·이영준·신진영·박기영, 2019, “경제 분석을 위한 텍스트 마이닝 (Text Mining for Economic Analysis)”, BOK Working Paper, 2019-18.

한국재정정보원, 2019, “재정정보 분야 텍스트마이닝 활용 방안 연구”, 한국재정정보원 분석보고서.