

Multiscale Geographically Weighted Regression (MGWR)

[Fotheringham et al., 2017]

곽지호

서울대학교 지리학과

2025년 5월 15일

1. Introduction
2. Model Formulation & Calibration, Bandwidth Selection
3. 시뮬레이션 데이터를 통한 GWR과 MGWR 비교
4. Irish Famine Dataset을 통한 GWR과 MGWR 비교
5. Discussion and Conclusion

1. Introduction
2. Model Formulation & Calibration, Bandwidth Selection
3. 시뮬레이션 데이터를 통한 GWR과 MGWR 비교
4. Irish Famine Dataset을 통한 GWR과 MGWR 비교
5. Discussion and Conclusion

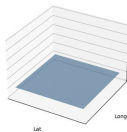
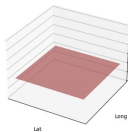
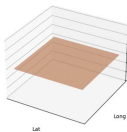
- 스케일은 지리학의 근본 개념 중 하나
 - Remark. 공간적 스케일의 종류: 지도 스케일, 지리 스케일, 측정 스케일, **작동 스케일**
- 서로 다른 프로세스가 서로 다른 스케일에서 작동하는 경우 빈번함
 - e.g. 어류 종 밀도의 감소 $\leftarrow$ 기후변화(global) + 남획(local)
- 그러나, 상이한 작동 스케일에 관한 정량적 연구는 거의 진행되지 않음
 - 예외: Bayesian Nonseparable Spatially Varing Coefficient (SVC) Models

- GWR은 공간적 이질성을 탐색할 수 있는 가장 대표적인 도구 [Brunsdon et al., 1996]
 - “평균적인” 전역적 관계 대신 설명변수와 반응변수의 관계가 공간적으로 변동하도록
OLS
 - 한 지역에서의 GWR 모델은 다른 지점의 데이터를 “빌려” 거리에 따라 가중치를 부여
 - Tobler의 지리학 제1법칙에 부합
 - 공간데이터의 공간적 의존성 또한 활용할 수 있음
- SGWR (Semiparametric GWR)의 개발
 - 전역적 관계와 국지적 관계를 동시에 고려할 수 있도록 함
 - 한계: 모든 설명-종속변수의 관계가 서로 다른 작동 스케일을 가진다면?

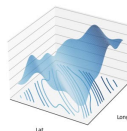
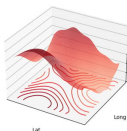
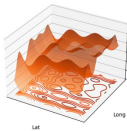
→ 가장 강력한 모델의 탄생: MGWR (Multiscale GWR)

Introduction

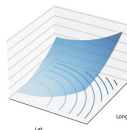
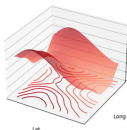
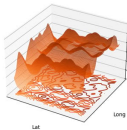
OLS



GWR



MGWR



1. Introduction
2. Model Formulation & Calibration, Bandwidth Selection
3. 시뮬레이션 데이터를 통한 GWR과 MGWR 비교
4. Irish Famine Dataset을 통한 GWR과 MGWR 비교
5. Discussion and Conclusion

Model Formulation & Calibration (1): GWR

$$y_i = \sum_{j=0}^m \beta_j(u_i, v_i) x_{ij} + \varepsilon_i, \quad (1)$$

- n 개 지점의 (1)을 하나로 표현하면: $\mathbf{y} = (\boldsymbol{\beta} \otimes \mathbf{X})\mathbf{1} + \boldsymbol{\epsilon}$ ($\otimes$ 는 요소별 곱)

- $\mathbf{y}$: $n \times 1$ 종속변수
- $\mathbf{X}$: $n \times (m+1)$ 행렬
- $\mathbf{1}$: $(m+1) \times 1$ 벡터 (각 원소: 1)

$$\bullet \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0(u_1, v_1) & \cdots & \beta_m(u_1, v_1) \\ \vdots & \ddots & \vdots \\ \beta_0(u_i, v_i) & \cdots & \beta_m(u_i, v_i) \\ \vdots & \ddots & \vdots \\ \beta_0(u_n, v_n) & \cdots & \beta_m(u_n, v_n) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}_{s_1}^\top \\ \vdots \\ \boldsymbol{\beta}_{s_i}^\top \\ \vdots \\ \boldsymbol{\beta}_{s_n}^\top \end{bmatrix}$$

Model Formulation & Calibration (1): GWR

$$y_i = \sum_{j=0}^m \beta_j(u_i, v_i) x_{ij} + \varepsilon_i, \quad (1)$$

- k 번째 지역의 위치를 s_k 라고 하면, 가중치 행렬은 $(\mathbf{W}_{s_k})_{ii} = K(d(s_i, s_k), h)$
 - $d(\cdot)$: 거리
 - K : 커널
 - h : 대역폭(Bandwidth)
- Model Calibration: $\hat{\beta}(s_k) = (\mathbf{X}^\top \mathbf{W}_{s_k} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}_{s_k} \mathbf{y}$
 - Weighted Least Squares Method로 추정

Model Formulation & Calibration (1): GWR

$$y_i = \sum_{j=0}^m \beta_j(u_i, v_i) x_{ij} + \varepsilon_i, \quad (1)$$

- Hat Matrix (Projection Matrix) [Joubert De Bellefon et al., 2018]

$$\mathbf{S} = \begin{bmatrix} \mathbf{x}_1^\top (\mathbf{X}^\top \mathbf{W}_{s_1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}_{s_1} \\ \vdots \\ \mathbf{x}_n^\top (\mathbf{X}^\top \mathbf{W}_{s_n} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}_{s_n} \end{bmatrix}_{n \times n}$$

- $\hat{\mathbf{y}} = (\hat{\boldsymbol{\beta}} \otimes \mathbf{X})\mathbf{1} = \mathbf{S}\mathbf{y}$
- $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}} = (\mathbf{I} - \mathbf{S})\mathbf{y}$
- cf. WLS $\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W} \mathbf{y}$

- GWR(과 SGWR)에서는 하나의 대역폭만 필요
- 대역폭 선정 기준: Goodness-of-Fit Measures

$$AIC_c = 2n \ln(\hat{\sigma}) + n \ln(2\pi) + n \frac{n + \text{tr}(\mathbf{S})}{n - 2 - \text{tr}(\mathbf{S})}$$

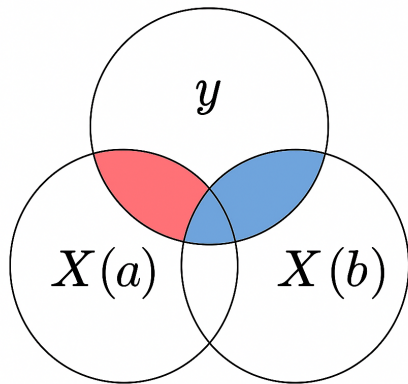
Model Formulation & Calibration (2): SGWR

$$y_i = \sum_{j=1}^{k_a} a_j x_{ij}(a) + \sum_{l=1}^{k_b} b_l(u_i, v_i) x_{il}(b) + \epsilon_i, \quad (2)$$

$$\Leftrightarrow \mathbf{y} = \mathbf{X}(a)\mathbf{a} + (\boldsymbol{\beta} \otimes \mathbf{X}(b))\mathbf{1} + \boldsymbol{\epsilon}$$

- $x_{ij}(a)$: 전역적 설명변수 ($j = 1, \dots, k_a$) $\rightarrow \mathbf{X}(a)$
- $x_{il}(b)$: 국지적 설명변수 ($l = 1, \dots, k_b$) $\rightarrow \mathbf{X}(b)$

Model Formulation & Calibration (2): SGWR



Model Formulation & Calibration (2): SGWR

- Model Calibration: $(k_a + 2)$ 번의 GWR과 1번의 OLS
 - 각각의 전역적 설명변수 $\mathbf{x}_j(a)$ 를 $\mathbf{X}(b)$ 에 대해 k_a 번 GWR 하고 종속변수 y 를 $\mathbf{X}(b)$ 에 대해 GWR 하여 잔차 $\mathbf{e}_{\mathbf{X}(a)}$ 와 $\mathbf{e}_y$ 를 구한다.

$$\mathbf{S}_{GWR}^* = \begin{bmatrix} \mathbf{x}_1(b)^\top (\mathbf{X}(b)^\top \mathbf{V}_{s_1} \mathbf{X}(b))^{-1} \mathbf{X}(b)^\top \mathbf{V}_{s_1} \\ \vdots \\ \mathbf{x}_n(b)^\top (\mathbf{X}(b)^\top \mathbf{V}_{s_n} \mathbf{X}(b))^{-1} \mathbf{X}(b)^\top \mathbf{V}_{s_n} \end{bmatrix}_{n \times n}$$

where $\mathbf{V}_{s_i}$ is the weight matrix

$$\begin{aligned} \mathbf{e}_{\mathbf{X}(a)} &= (\mathbf{I} - \mathbf{S}_{GWR}^*) \mathbf{X}(a) \\ \mathbf{e}_y &= (\mathbf{I} - \mathbf{S}_{GWR}^*) \mathbf{y} \end{aligned}$$

Model Formulation & Calibration (2): SGWR

- Model Calibration: ($k_a + 2$)번의 GWR과 1번의 OLS

2. e_y 를 $e_{\mathbf{X}(a)}$ 에 대해 OLS를 통해 전역적 (편)회귀계수 $\hat{a}$ 를 구한다.

$$\begin{aligned}\hat{a} &= (e_{\mathbf{X}(a)}^\top e_{\mathbf{X}(a)})^{-1} e_{\mathbf{X}(a)}^\top e_y \\ &= (\mathbf{X}(a)^\top (\mathbf{I} - \mathbf{S}_{GWR}^*)^\top (\mathbf{I} - \mathbf{S}_{GWR}^*) \mathbf{X}(a))^{-1} \mathbf{X}(a)^\top (\mathbf{I} - \mathbf{S}_{GWR}^*)^\top (\mathbf{I} - \mathbf{S}_{GWR}^*) \mathbf{y} \\ &= (\mathbf{X}(a)^\top \mathbf{W} \mathbf{X}(a))^{-1} \mathbf{X}(a)^\top \mathbf{W} \mathbf{y}\end{aligned}$$

3. $\mathbf{y} - \mathbf{X}(a)\hat{a}$ 를 $\mathbf{X}(b)$ 에 대해 GWR 하여 국지적 (편)회귀계수 $\hat{b}$ 를 구한다.

$$(\hat{b} \otimes \mathbf{X}(b))\mathbf{1} = \mathbf{S}_{GWR}^*(\mathbf{y} - \mathbf{X}(a)\hat{a})$$

Model Formulation & Calibration (2): SGWR

- Derivation of $\mathbf{S}_{SGWR}$ ($\rightarrow$ AICc에 대입, 대역폭 선택)

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{X}(a)\hat{\mathbf{a}} + (\hat{\mathbf{b}} \otimes \mathbf{X}(b))\mathbf{1} \\&= \mathbf{X}(a)\hat{\mathbf{a}} + \mathbf{S}_{GWR}^*(\mathbf{y} - \mathbf{X}(a)\hat{\mathbf{a}}) \\&= \mathbf{X}(a)\hat{\mathbf{a}} + \mathbf{S}_{GWR}^*\mathbf{y} - \mathbf{S}_{GWR}^*\mathbf{X}(a)\hat{\mathbf{a}} \\&= \mathbf{X}(a)\hat{\mathbf{a}} + \mathbf{S}_{GWR}^*\mathbf{y} - \underbrace{\mathbf{S}_{GWR}^*\mathbf{X}(a)(\mathbf{X}(a)^\top \mathbf{W}\mathbf{X}(a))^{-1}\mathbf{X}(a)^\top \mathbf{W}}_0 \mathbf{y} \\&\quad 0 \quad (\because \mathbf{W} = (\mathbf{I} - \mathbf{S}_{GWR}^*)^\top (\mathbf{I} - \mathbf{S}_{GWR}^*)) \\&= \mathbf{X}(a)\hat{\mathbf{a}} + \mathbf{S}_{GWR}^*\mathbf{y} \\&= \mathbf{X}(a)(\mathbf{X}(a)^\top \mathbf{W}\mathbf{X}(a))^{-1}\mathbf{X}(a)\mathbf{W}\mathbf{y} + \mathbf{S}_{GWR}^*\mathbf{y} \\&= \mathbf{S}_{SGWR}\mathbf{y}\end{aligned}$$

$$\therefore \mathbf{S}_{SGWR} = \mathbf{X}(a)(\mathbf{X}(a)^\top \mathbf{W}\mathbf{X}(a))^{-1}\mathbf{X}(a)\mathbf{W} + \mathbf{S}_{GWR}^*$$

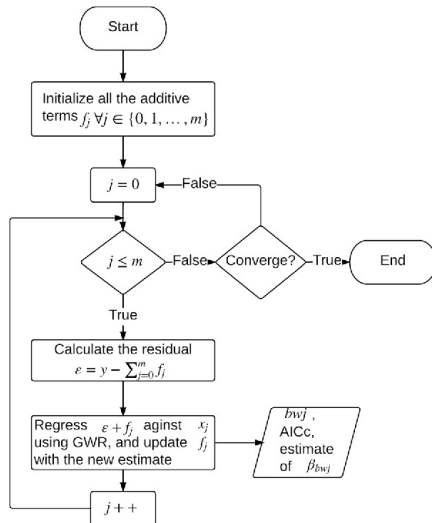
Model Formulation (3): MGWR

$$y_i = \sum_{j=0}^m \beta_{bw_j}(u_i, v_i) x_{ij} + \epsilon_i, \quad (3)$$

- 기존의 Model Calibration과 Bandwidth Selection 방법 적용하기 어려움
 - 각 위치에서 변수마다 서로 다른 대역폭을 지님
 - 가능한 변수별 대역폭 조합의 가짓수가 너무 많음
- GAM(generalized additive model)에서 사용하는 Back-Fitting Algorithm 이용

$$\mathbf{y} = \sum_{j=0}^m \mathbf{f}_j + \boldsymbol{\epsilon}$$

Back-Fitting Algorithm



Back-Fitting Algorithm

Back-Fitting Algorithm for a MGWR Model Calibration

1. 최초의 $\hat{f}_j$ 설정 ($j = 0, 1, 2, \dots, m$)
2. 잔차 계산: $\mathbf{e} = \mathbf{y} - \sum_{j=0}^m \hat{f}_j$
3. $\mathbf{e} + \hat{f}_k = \mathbf{y} - \sum_{j \neq k}^m \hat{f}_j$ 에 대해 $\mathbf{x}_k$ 로 GWR을 적합하여 bw_k , $AICc_k$, $\hat{\beta}_k$ 업데이트
4. 2, 3을 $k = 1, 2, \dots, m$ 에 대해 실행
5. 추정값의 차이가 수렴할 때까지 2, 3, 4 반복

Back-Fitting Algorithm: 두 가지 결정

1. 초기값 (국지 회귀계수 추정량) 설정

- 0
- 전역적 모델 (일반 OLS)의 추정량
- GWR의 추정량

→ Rule of Thumb: GWR 추정량

Back-Fitting Algorithm: 두 가지 결정

2. 알고리즘 종료 조건 (수렴성 판단)

- SOC-RSS: RSS의 변화율

$$SOC_{RSS} = \frac{|RSS_{new} - RSS_{old}|}{RSS_{new}}$$

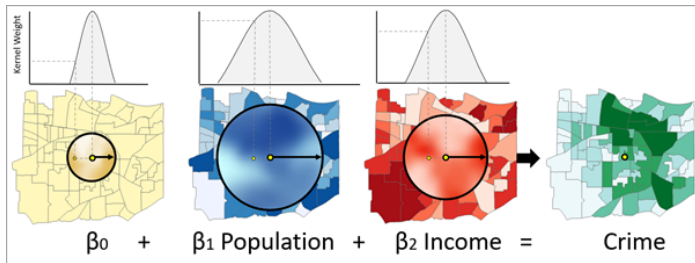
- SOC-f: GWR smoother의 변화율

$$SOC_f = \sqrt{\frac{\frac{1}{n} \sum_{j=0}^p \sum_{i=1}^n (\hat{f}_{ij}^{new} - \hat{f}_{ij}^{old})^2}{\sum (\sum \hat{f}_{ij}^{new})^2}}$$

→ 모두 scale-free; 아래 분석에서는 가법 항들의 차이에 주목하는 SOC-f 선택

Caveat

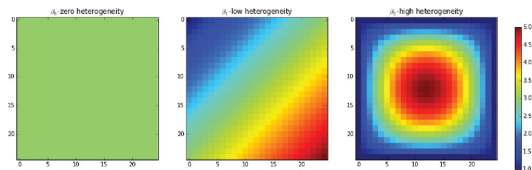
- 해석 및 비교를 용이하도록 하기 위해 각각의 변수를 평균 0, 표준편차 1이 되도록 정규화(standardize) 할 수 있다.
- 정규화 할 경우 채택된 대역폭은 설명변수와 종속변수 간 관계의 작동 스케일로 해석, 비교할 수 있다.



1. Introduction
2. Model Formulation & Calibration, Bandwidth Selection
3. 시뮬레이션 데이터를 통한 GWR과 MGWR 비교
4. Irish Famine Dataset을 통한 GWR과 MGWR 비교
5. Discussion and Conclusion

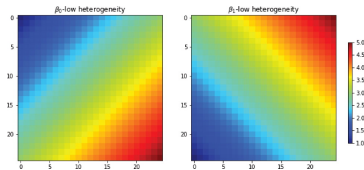
Simulation Design 1

- DGP: $y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \epsilon_i$
 - $x_1, x_2 \sim N(0, 1)$; $\epsilon \sim N(0, 0.5^2)$ 로 100개의 데이터 생성
- 다른 공간적 이질성 수준의 Parameter Surface: MGWROI 더 잘 모델링하는가?
 - $\beta_0 \rightarrow \beta_{zero} = 3$
 - $\beta_1 \rightarrow \beta_{low} = 1 + \frac{1}{12}(u + v)$
 - $\beta_2 \rightarrow \beta_{high} = 1 + \frac{1}{324} \left[36 - \left(6 - \frac{u}{2} \right)^2 \right] \left[36 - \left(6 - \frac{v}{2} \right)^2 \right]$



Simulation Design 2

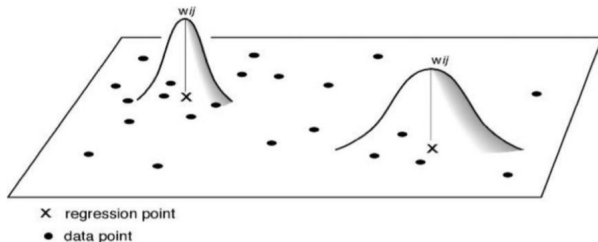
- DGP: $y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \epsilon_i$
 - $x_1, x_2 \sim N(0, 1)$; $\epsilon \sim N(0, 0.5^2)$ 로 100개의 데이터 생성
- **같은** 공간적 이질성 수준의 Parameter Surface: GWR과 MGWR 비슷한 결과?
 - $\beta_0 \rightarrow \beta_{low} = 1 + \frac{1}{12}(u + v)$
 - $\beta_1 \rightarrow \beta_{low}$ 를 90도 회전



Evaluation Methods

1. 대역폭 비교

- 가설
 - MGWR은 공간적으로 동질적인 관계와 이질적인 관계를 잘 구분한다.
 - GWR의 단일 대역폭은 MGWR의 변수별 대역폭의 평균이다.
- 비교에는 적응형 대역폭($:=$ 국지 회귀에 이용되는 최근린 이웃 수)을 이용한다.



[Oshan et al., 2019]

1. 대역폭 비교 (cont.)

- Adaptive Bisquare Kernel은 다음과 같이 정의된다:

$$w_{ij} = \begin{cases} \left[1 - \left[\frac{d_{ij}}{G_i}\right]^2\right]^2 & \text{if } d_{ij} < G_i \\ 0 & \text{Otherwise} \end{cases}$$

- d_{ij} : i 와 j 지점 간 거리
- G_i : M^{th} Nearest Neighbor와의 거리
- M : AICc 비교를 통해 구한 최적의 Nearest Neighbor 개수

2. 회귀계수 추정의 정확도

- GWR과 MGWR 추정의 정확도는 회귀계수별 $RMSE_j$ 를 통해 평가된다.

$$RMSE_j = \sqrt{\frac{1}{n} \sum_{i=1}^n (\beta_j(u_i, v_i) - \hat{\beta}_j(u_i, v_i))^2}$$

3. 적합도 (Goodness of Fit)

- RSS (MGWR에서는 변수별 weight matrix가 상이하므로 AICc 계산 불가*)
 - Note. RSS는 단위-의존적이고 모델 복잡도에 대한 페널티를 주지 않음

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

4. 계산 효율성

- 한 번의 Model Calibration에 걸리는 시간으로 정의한다.
- MGWRO가 GWR에 비해 더 많은 시간이 걸릴 것으로 예상된다.

Initialization and Convergence Criteria

Iterations	AICc ₀	AICc ₁	AICc ₂	b ₀	b ₁	b ₂	SOC _f
1	3,544.475178	3,107.084069	1,223.544247	622	310	46	0.039566412
2	1,118.033996	966.721046	984.2842977	444	62	44	0.002339227
3	904.1492707	954.050443	984.5833555	524	62	44	0.000155162
4	902.270467	953.5527079	984.7089459	536	62	44	1.70×10^{-5}
5	902.1610436	953.4854188	984.7075472	536	62	44	2.03×10^{-6}
6	902.1504341	953.4776078	984.7065979	536	62	44	3.29×10^{-7}
7	902.148657	953.4763536	984.7063954	536	62	44	5.60×10^{-8}

Simulation Design 1 - Data Set 4의 Back-Fitting Algorithm 작동 결과 (최초 추정량: 0)

- $SOC_f \leq 10^{-5}$, 즉 5번째 반복 이후 부터 대역폭이 수렴하는 것을 확인할 수 있다.

Initialization and Convergence Criteria

Iterations	Zeros			GWR			OLS		
	b_0	b_1	b_2	b_0	b_1	b_2	b_0	b_1	b_2
1	622	310	46	260	62	44	413	146	46
2	444	62	44	524	62	44	512	62	44
3	524	62	44	536	62	44	536	62	44
4	536	62	44	536	62	44	536	62	44
5	536	62	44				536	62	44

최초 추정량 설정에 따른 Simulation Design 1 - Data Set 4의 대역폭 수렴 과정 ($SOC_f \leq 10^{-5}$ 면 종료)

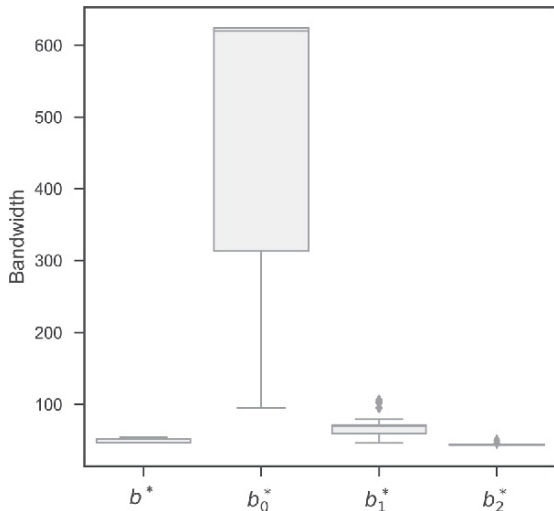
- 최초 추정량의 설정과 무관하게 동일한 값으로 수렴이 이루어진다.
- 0, GWR-추정량, OLS-추정량 중 **GWR-추정량**을 초기값으로 설정했을 때 가장 수렴 속도가 빠르다.

→ 아래 분석에서 최초 추정량은 GWR 추정량, 종료 기준은 $SOC_f \leq 10^{-5}$ 으로 설정

Simulation Design 1: Results

1. 최적 대역폭 선정 결과

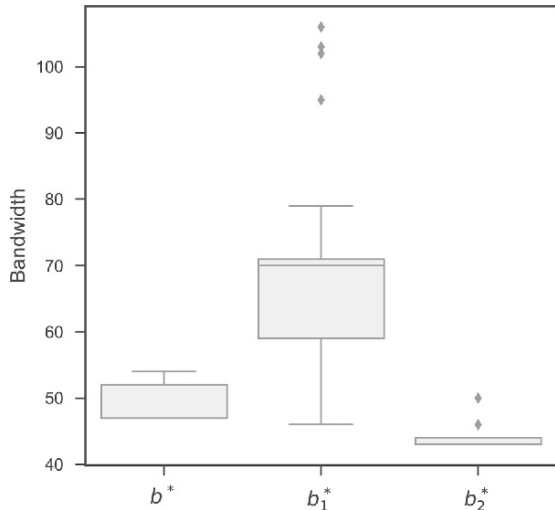
- $GWR(b^*)$ vs. $MGWR(b_0^*, b_1^*, b_2^*)$
 - MGWR이 세 가지 스케일 포착함 (Heterogeneity $\uparrow$, Bandwidth $\downarrow$)
 - b_0^* : 큰 대역폭 $\rightarrow$ 전역적 프로세스
 - b_1^*, b_2^* : 작은 대역폭 $\rightarrow$ 국지적 프로세스
 - b^* : 공간적 이질성 수준의 가중 평균으로 이해 가능 (가중치: 변수별 설명력)



Simulation Design 1: Results

1. 최적 대역폭 선정 결과

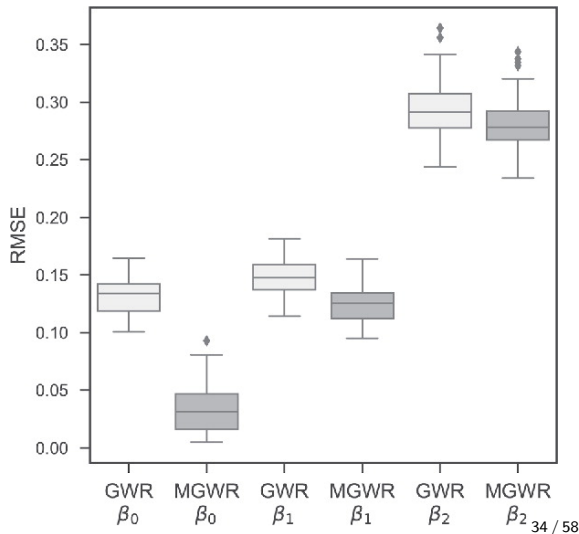
- $y \sim x_1(b_1^*)$ vs. $y \sim x_2(b_2^*)$
 - b_1^* 이 b_2^* 보다 항상 낮은 값
→ $y \sim x_1$ 이 $y \sim x_2$ 보다 더 국지적인 프로세스라는 사실이 확인된다.



Simulation Design 1: Results

2. 국지 회귀계수 추정 정확도

- RMSE
 - 모든 parameter에서 MGWR이 더 낮은 RMSE를 보임
 - β_0 에서 GWR-MGWR 추정량의 차이 큼

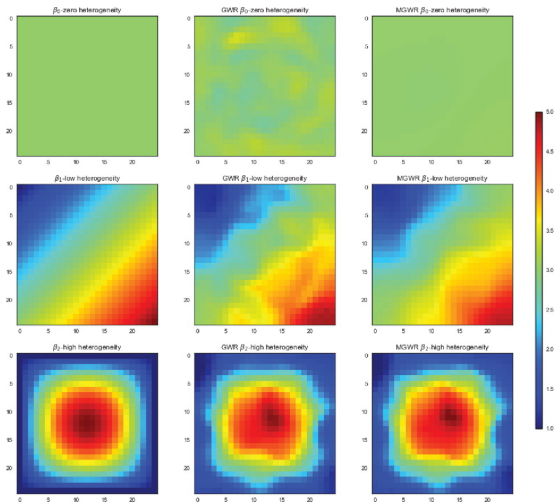


Simulation Design 1: Results

2. 국지 회귀계수 추정 정확도

- True vs. Estimated Parameter Surface

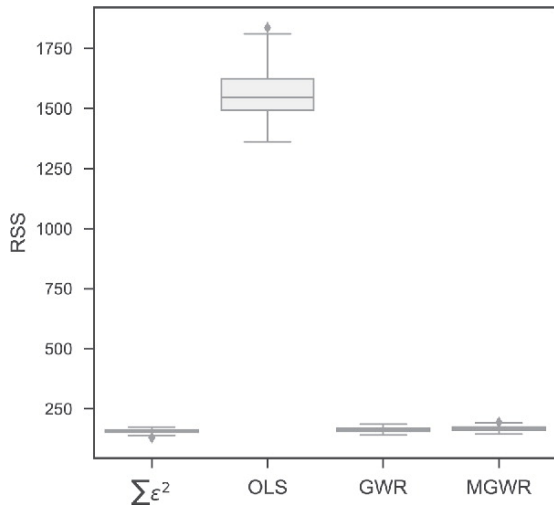
- β_1, β_2 에 비해 β_0 의 차이가 크다.
- GWR의 “평균” 대역폭의 크기가 β_1, β_2 의 MGWR 대역폭과 유사한 반면 β_0 의 대역폭은 훨씬 크기 때문



Simulation Design 1: Results

3. 적합도

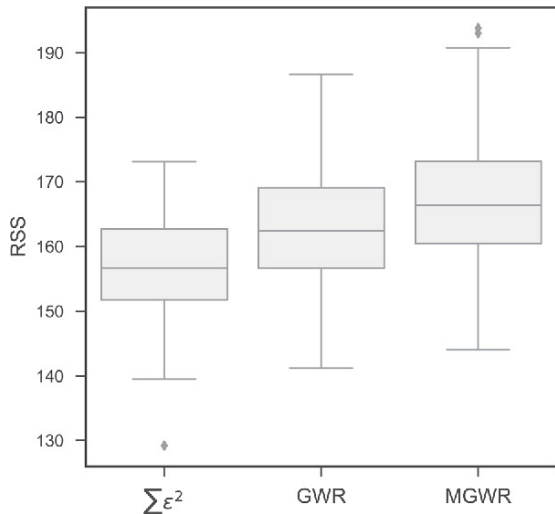
- RSS, $\sum \epsilon^2$
- OLS는 데이터의 공간적 이질성을 충분히 설명하지 못함



Simulation Design 1: Results

3. 적합도

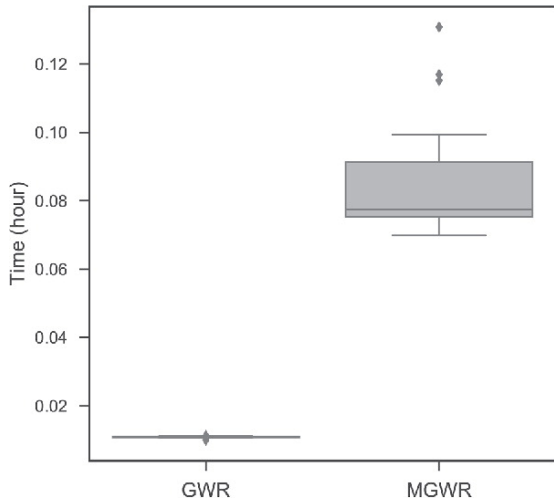
- $RSS, \sum \epsilon^2$
 - GWR의 RSS가 MGWR에 비해 더 낮음
 - GWR에서는 절편이 공간적으로 변이
→ 잔차 효과의 일부를 설명
 - 이는 Overfitting에 기인한 것으로 실제 DGP와는 맞지 않는 정보



Simulation Design 1: Results

4. 계산 효율성

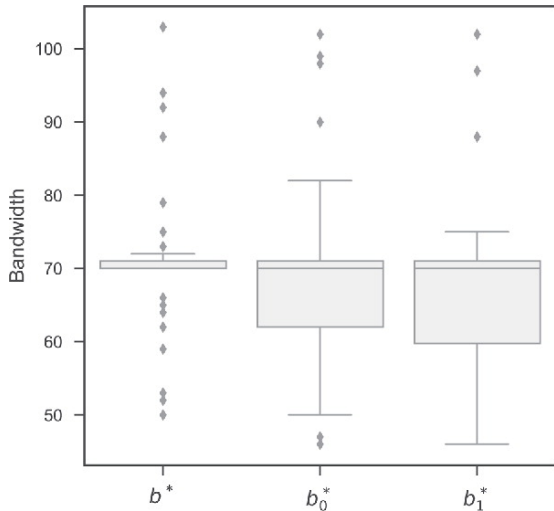
- 모델 적합에 걸린 시간
 - Initial Values: GWR 추정량
 - Convergence Criterion: $SOC_f \leq 10^{-5}$
 - MGWR 평균 소요 시간: 4.8분 (GWR의 약 8배)



Simulation Design 2: Results

1. 최적 대역폭 선정 결과

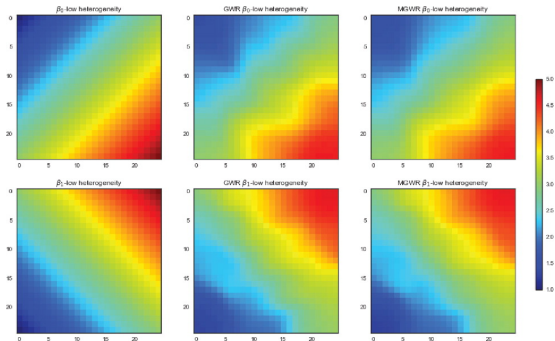
- $GWR(b^*)$ vs. $MGWR(b_0^*, b_1^*)$
 - 중간값: 거의 동일
($\because$ 동일한 공간적 이질성 수준)
 - 분산: $b^* < b_0^*, b_1^*$
($\because b^*$ 는 b_0^*, b_1^* 의 평균)



Simulation Design 2: Results

2. 국지 회귀계수 추정 정확도

- True vs. Estimated Parameter Surface
 - 변수별 공간적 이질성 수준이 동일하므로 예상과 같이 GWR과 MGWR의 추정량이 거의 같다.

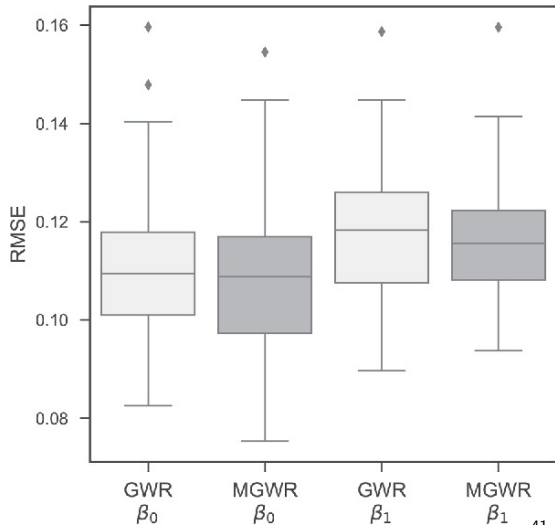


Simulation Design 2: Results

2. 국지 회귀계수 추정 정확도

- RMSE

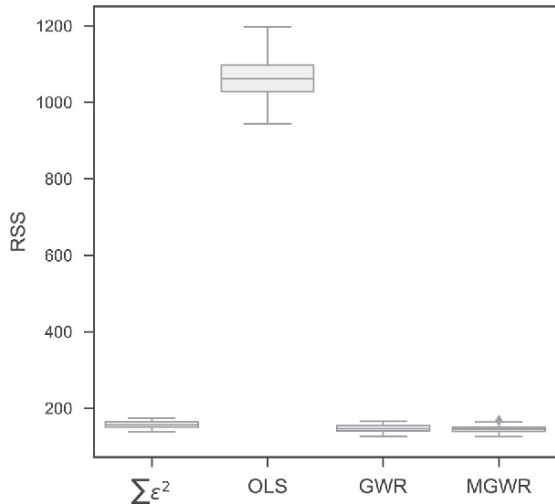
- 변수별 공간적 이질성 수준이 동일하므로 예상과 같이 GWR과 MGWR의 추정량이 거의 같다.



Simulation Design 2: Results

3. 적합도

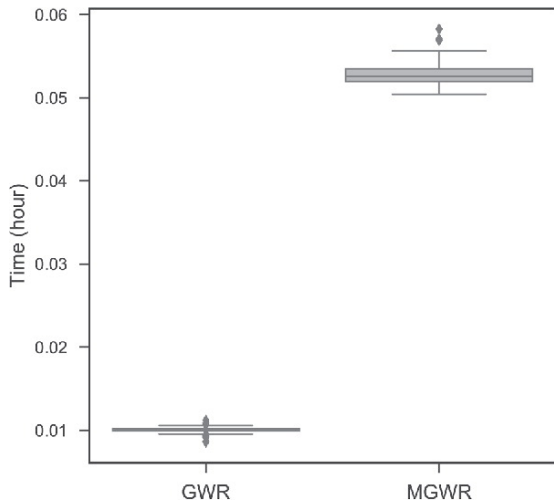
- RSS, $\sum \epsilon^2$
 - OLS는 데이터의 공간적 이질성을 충분히 설명하지 못함
 - GWR과 MGWR 모두 높은 정확도로 원래 데이터를 추정함



Simulation Design 2: Results

4. 계산 효율성

- 모델 적합에 걸린 시간
 - Initial Values: GWR 추정량
 - Convergence Criterion: $SOC_f \leq 10^{-5}$
 - 0.6분 (GWR) vs. 3.15분 (MGWR)



1. Introduction
2. Model Formulation & Calibration, Bandwidth Selection
3. 시뮬레이션 데이터를 통한 GWR과 MGWR 비교
4. Irish Famine Dataset을 통한 GWR과 MGWR 비교
5. Discussion and Conclusion

Irish Famine Dataset [Fotheringham et al., 2013]

- 관측지 개수: 2,317개 선거구 (Electoral Divisions)
- 종속변수: 1841년과 1851년 간 인구 증감 (Note. 아일랜드 대기근: 1841-1850)
- 설명변수: 8개
 - Grain Cultivation (곡물 수확량)
 - Mean Elevation (평균 고도)
 - Average Holding Size (평균 토지 면적)
 - Potato Cultivation (감자 수확량)
 - Urban Proximity (도심 근접성)
 - Workhouse Proximity (구빈원 근접성)
 - Population Density on Cropped Land (경작지 인구밀도)
 - Distance to Coast (해안까지의 거리)

Irish Famine Dataset

Predictor Variable	GWR	MGWR
Grain Cultivation	136	2,316
Mean Elevation	136	81
Average Holding Size	136	365
Potato Cultivation	136	86
Urban Proximity	136	66
Workhouse Proximity	136	210
Population Density On Cropped Land	136	91
Distance To Coast	136	2,316

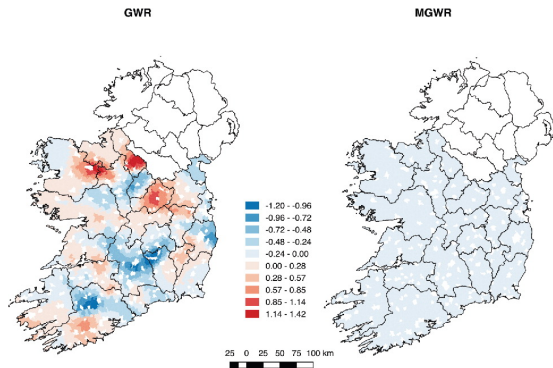
- GWR과 MGWR으로 구한 변수별 최적 대역폭
- 인구 증감에 영향을 미치는 변수별 작동 스케일이 다름 (global-regional-local)

Irish Famine Dataset

1. No Variation (Global)

- Distance to the Coast

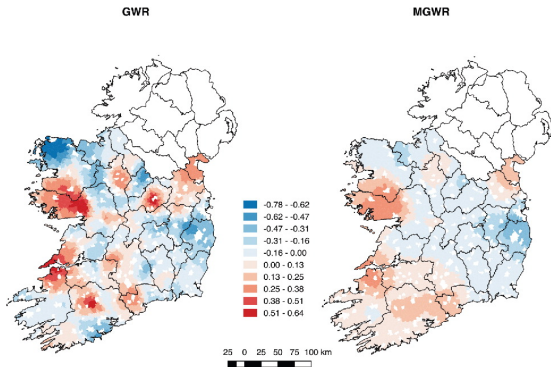
- MGWR (2,316)
 - 전역적으로 0에 가까운 음수
- GWR (136)
 - 지역에 따라 음수, 양수 혼재



2. Regional Variation

- Workhouse Proximity

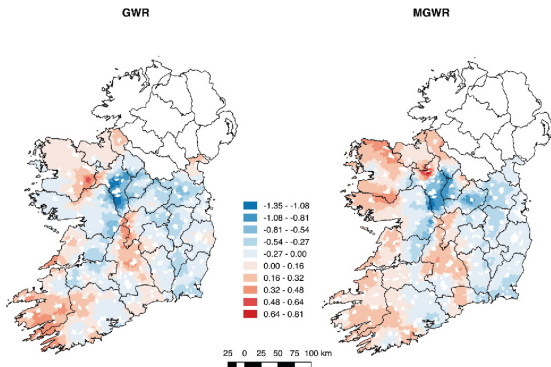
- MGWR (210) vs. GWR(136)
- 추정량 거의 비슷 ($\therefore$ 대역폭 유사)
- MGWR 추정량이 좀 더 지역적(덜 국지적)이고 0에 가까움
- 서부와 남서부에서는 구빈원이 인구 감소를 억제하는 역할



3. Local Variation

- Mean Elevation

- MGWR (81) vs. GWR (136)
- 추정량 거의 비슷 ($\therefore$ 대역폭 유사)
- MGWR 추정량이 좀 더 국지적
- 서부: 고저차가 커서 높은 지대에서 감자 재배보다 방목을 함
→ 대기근으로 인한 피해 적음
- 다른 지역: 고저차가 적어 높은 지대에서 감자 재배 (배수 좋음)
→ 대기근으로 인한 피해 큼



1. Introduction
2. Model Formulation & Calibration, Bandwidth Selection
3. 시뮬레이션 데이터를 통한 GWR과 MGWR 비교
4. Irish Famine Dataset을 통한 GWR과 MGWR 비교
5. Discussion and Conclusion

- MGWR: GWR의 확장
 - 변수별 영향력의 공간적 이질성이 동일하다는 가정을 완화
 - 다른 스케일의 프로세스가 복합적으로 영향을 미치는 경우 고려할 수 있게 됨
- Back-Fitting Algorithm을 이용한 Model Calibration
 - 결과: 설명변수마다 서로 다른 대역폭 지정
 - 변수를 표준화할 경우, 각각의 대역폭은 일종의 “작동 스케일”로 해석 가능
- 두 종류의 시뮬레이션 결과: 변수별 영향력의 작동 스케일 정확히 파악
 - (1) 변수별로 다른 공간적 이질성 수준 → (GWR과 달리) 서로 다른 대역폭 지정
 - (2) 변수별로 같은 공간적 이질성 수준 → (GWR과 유사한) 비슷한 대역폭 지정

1. MGWR에 활용할 수 있는 AICc와 같은 Goodness-of-Fit Measure의 개발
2. Local R^2 , Local 표준오차, Local t -값에 대응하는 통계량 개발
3. MGWR의 계산 효율성 제고
4. MGWR과 SVC 모델을 비교할 수 있는 지표 개발
5. MGWR에 관한 통계적 추론: 유의성 검정 절차의 개발

References



Brunsdon, C., Fotheringham, A. S., and Charlton, M. E. (1996).
Geographically weighted regression: A method for exploring spatial nonstationarity.
Geographical Analysis, 28(4):281–298.



Fotheringham, A. S., Kelly, M. H., and Charlton, M. (2013).
The demographic impacts of the irish famine: towards a greater geographical understanding.
Transactions of the Institute of British Geographers, 38(2):221–237.



Fotheringham, A. S., Yang, W., and and, W. K. (2017).
Multiscale geographically weighted regression (mgwr).
Annals of the American Association of Geographers, 107(6):1247–1265.



Joubert De Bellefon, M.-P., Loonis, V., Fontaine, M., and Costemalle, V. (2018).
Handbook of Spatial Analysis with R Insee-Eurostat.

 Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., and Fotheringham, A. S. (2019).

mgwr: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale.

ISPRS International Journal of Geo-Information, 8(6).

 Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., and Wolf, L. J. (2020).

Inference in multiscale geographically weighted regression.

Geographical Analysis, 52(1):87–106.

Additional Materials: Inference in MGWR [Yu et al., 2020]

- GWR as a GAM

$$\mathbf{y} = \sum_{j=1}^k \mathbf{f}_j + \boldsymbol{\epsilon}$$

$$\hat{\mathbf{y}} = \sum_{j=1}^k \hat{\mathbf{f}}_j \text{ where } \hat{\mathbf{f}}_j = \begin{bmatrix} x_{1j}\hat{\beta}_{1j} \\ \vdots \\ x_{nj}\hat{\beta}_{nj} \end{bmatrix}$$

Since $\hat{\beta}_{ij} = \mathbf{e}_j(\mathbf{X}^\top \mathbf{W}_i \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}_i \mathbf{y}$,

$$\hat{\mathbf{f}}_j = \underbrace{\begin{bmatrix} x_{1j}\mathbf{e}_j(\mathbf{X}^\top \mathbf{W}_1 \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}_1 \\ \vdots \\ x_{nj}\mathbf{e}_j(\mathbf{X}^\top \mathbf{W}_n \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}_n \end{bmatrix}}_{\mathbf{R}_j: \text{ Hat Matrix for } \hat{\mathbf{f}}_j} \mathbf{y}$$

Additional Materials: Inference in MGWR

- The Hat Matrix in MGWR
 - Initialization using GWR Estimates:

$$\hat{\mathbf{y}} = \sum_{j=1}^k \hat{\mathbf{f}}_j = \left(\sum_{j=1}^k \mathbf{R}_j \right) \mathbf{y} = \mathbf{S} \mathbf{y}$$

- Regress $\mathbf{x}_j$ against $\hat{\mathbf{f}}_j + \mathbf{e}$: (GWR)

$$\hat{\mathbf{f}}_j + \mathbf{e} = \hat{\mathbf{f}}_j + \mathbf{y} - \sum_{j=1}^k \hat{\mathbf{f}}_j = \mathbf{R}_j \mathbf{y} + \mathbf{y} - \mathbf{S} \mathbf{y}$$

$$\begin{aligned} \hat{\mathbf{f}}_j^* &= \mathbf{A}_j(\hat{\mathbf{f}}_j + \mathbf{e}) \\ &= \mathbf{A}_j(\mathbf{R}_j \mathbf{y} + \mathbf{y} - \mathbf{S} \mathbf{y}) \\ &= \underbrace{(\mathbf{A}_j \mathbf{R}_j + \mathbf{A}_j - \mathbf{A}_j \mathbf{S})}_{\mathbf{R}_j^*: \text{Updated Hat Matrix}} \mathbf{y} \end{aligned}$$

Additional Materials: Inference in MGWR

- The Hat Matrix in MGWR (after regressing $\mathbf{x}_j$)

$$\mathbf{S}^* = \mathbf{S} - \mathbf{R}_j + \mathbf{R}_j^*$$

- Local Standard Errors of the Parameter Estimates in MGWR

$$\hat{\mathbf{f}}_j = \begin{bmatrix} x_{1j}\hat{\beta}_{1j} \\ x_{2j}\hat{\beta}_{2j} \\ \vdots \\ x_{nj}\hat{\beta}_{nj} \end{bmatrix} = \begin{bmatrix} x_{1j} & 0 & \cdots & 0 \\ 0 & x_{2j} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & x_{nj} \end{bmatrix} \hat{\boldsymbol{\beta}}_j = \text{diag}(\mathbf{x}_j)\hat{\boldsymbol{\beta}}_j$$

$$\hat{\boldsymbol{\beta}}_j = [\text{diag}(\mathbf{x}_j)]^{-1} \hat{\mathbf{f}}_j = [\text{diag}(\mathbf{x}_j)]^{-1} \mathbf{R}_j \mathbf{y} = \mathbf{C} \mathbf{y}$$

$$\text{Var}(\hat{\boldsymbol{\beta}}_j) = \text{diag}(\mathbf{C} \mathbf{C}^\top \hat{\sigma}^2) \text{ where } \hat{\sigma}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n - \text{tr}(\mathbf{S})} \rightarrow \frac{\hat{\beta}_{ij}}{\sqrt{\text{Var}(\hat{\beta}_{ij})}} \sim t_{n - \text{tr}(\mathbf{S})}$$

감사합니다.