空間従属性に対する交通ネットワークからの影響に着目したKrigingの拡張

筑波大学大学院
システム情報工学研究科１年次
村上大輔

空間内挿法Kriging

≒データ間の類似性

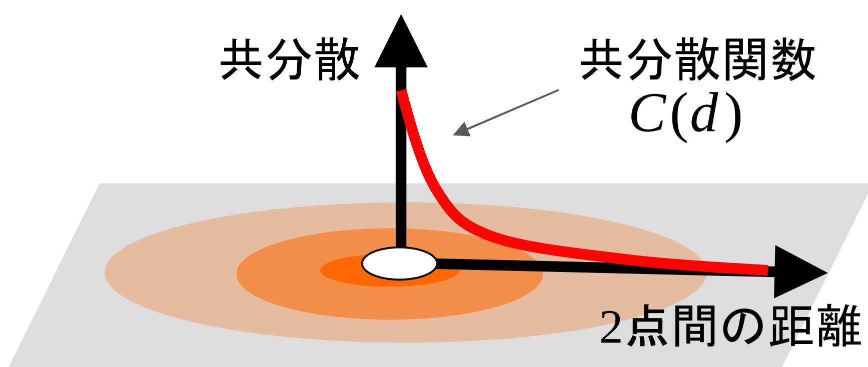
- 空間データ相互の従属関係（空間従属性）を考慮した空間内挿手法

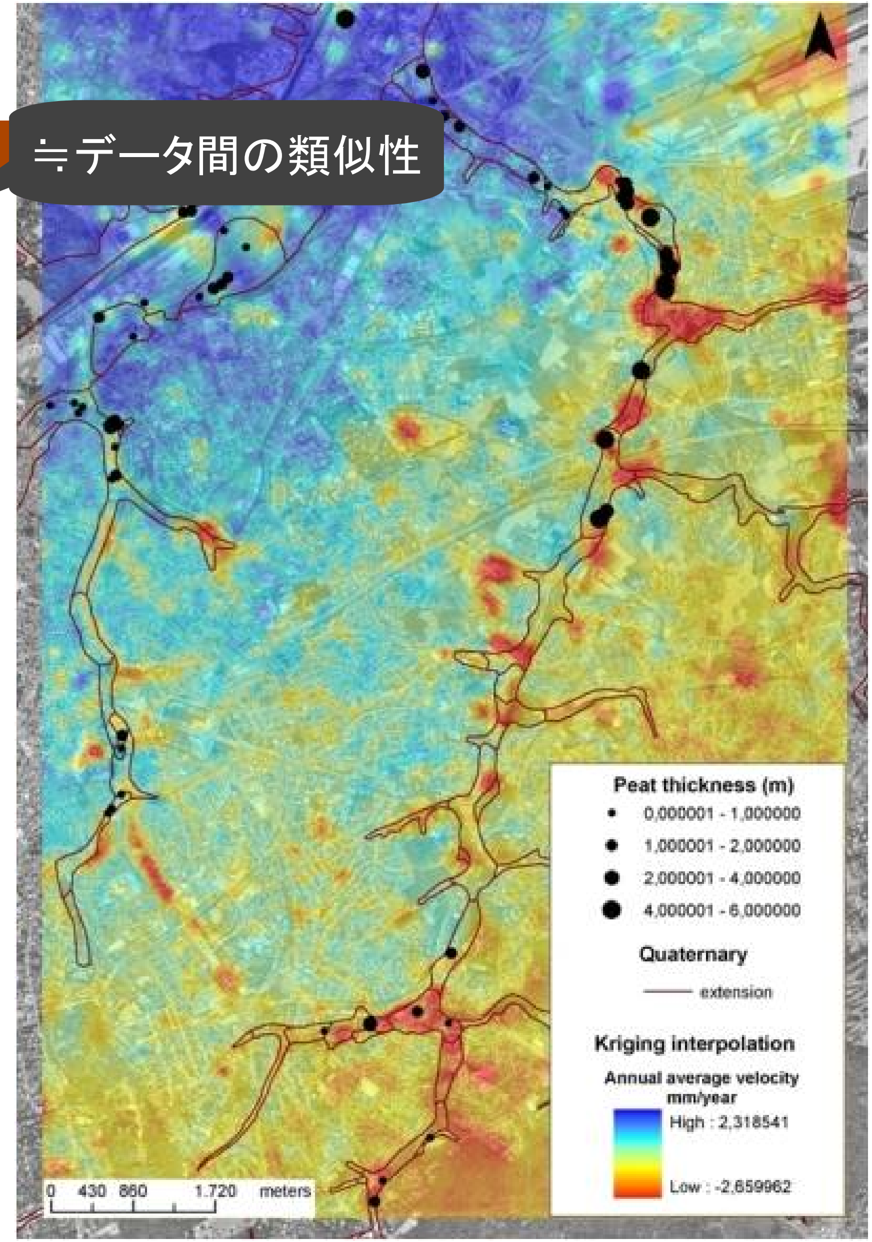
$$z(s_i) = \mu(s_i) + \varepsilon(s_i) \quad \varepsilon(s_i) \sim N(0, C(d_{ij}))$$

$z(s_i)$	:被説明変数	$\mathbf{s}_i = [x_i, y_i]'$	:地点iの位置ベクトル
$\mu(s_i)$	:トレンド項	$\varepsilon(s_i)$	:局所的変動成分
d_{ij}	:距離	$C(d_{ij})$	:共分散関数

■空間従属性の定式化

- 共分散にユークリッド距離の関数（共分散関数）を仮定





Krigingによる内挿の例：泥炭の深さ

出典：The Geological Society, URL： http://www.geolsoc.org.uk/index.html

空間従属性と都市活動

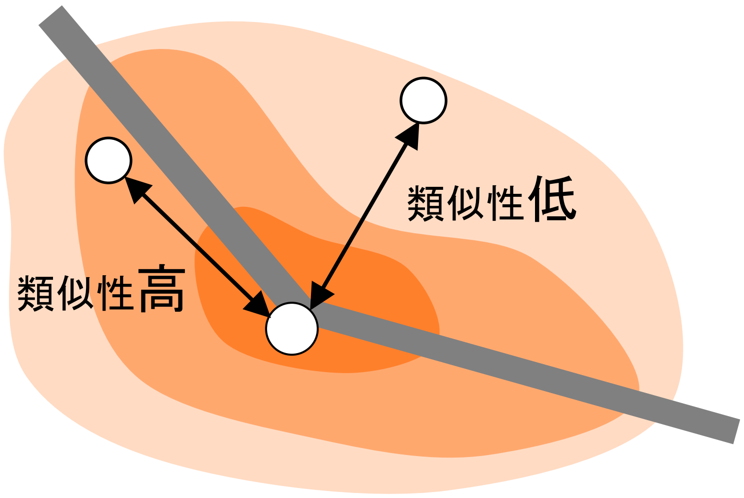
研究背景

■都市活動における移動

- ネットワーク上を人々が移動する
- ネットワーク上の最短経路距離が小さい程、移動コストは低い

■仮説

- ２地点がネットワークでつながれていれば空間従属性は強まる



■ネットワークを考慮したKriging内挿を行うための方法

- 共分散関数をネットワーク上の最短経路距離の関数で与える

最短経路距離の関数として
共分散関数を定義したKrigingの先行研究

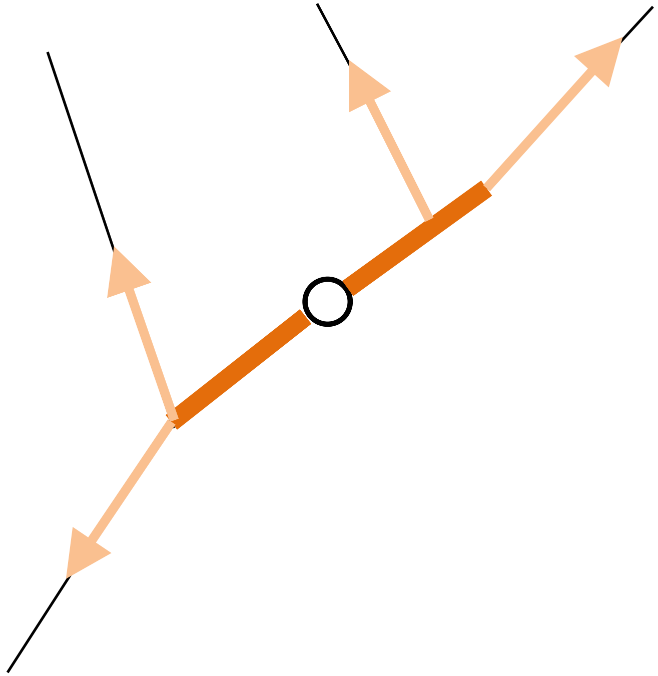
研究背景

■ ネットワーク空間内への内挿

- 河川(Hoef et al(2006),Cressie et al(2006)他多数)
- 道路(Curriero(1996))

■ ネットワーク空間外への内挿

- なし

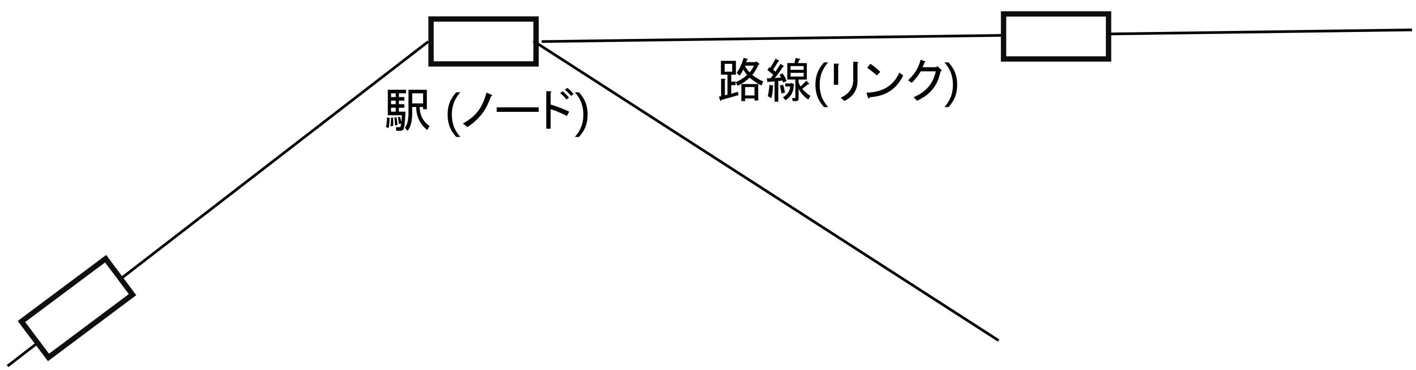


研究目標

最短経路距離を共分散関数に考慮した
ネットワーク外の空間への内挿のためのKrigingの拡張

対象とするネットワーク

- 鉄道ネットワーク(航路、高速道路等への応用も可能)



研究の流れ

1. 最短経路距離を考慮した共分散関数の提案
2. 提案された共分散関数を用いた実証

共分散関数提案の際の前提条件

- 用いる距離と導かれる共分散は以下の性質を満たさなくてはならない

条件1：三角不等式：距離の公理

$$d(i,j) \le d(i,k) + d(k,j)$$

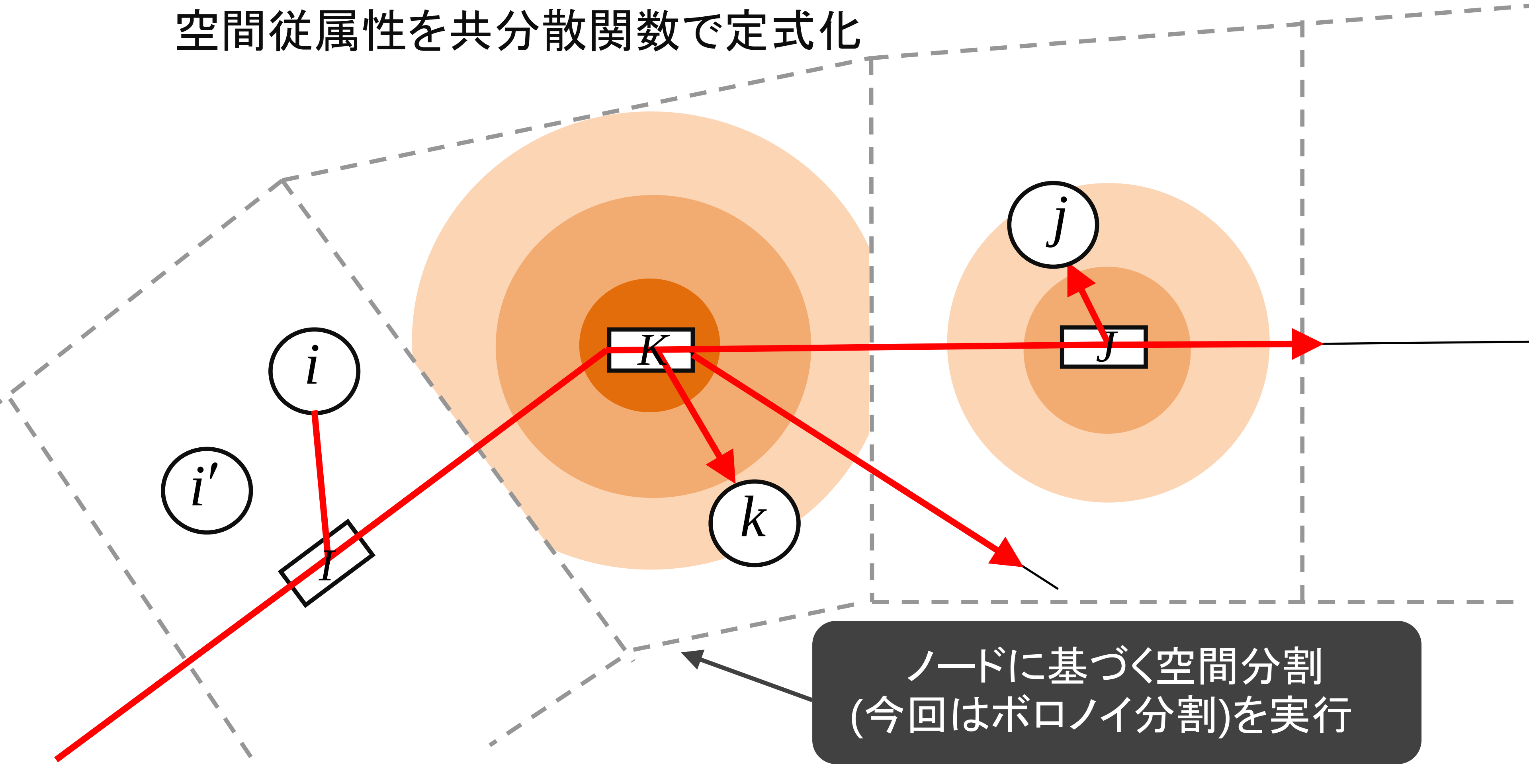
条件2：正定値性：共分散が満たさなくてはならない条件

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(s_i, s_j) \ge 0$$

共分散関数で定式化する空間従属性

■地点 i からの空間従属性のイメージ

- ネットワークを通して各領域内に広がる下図のような空間従属性を共分散関数で定式化



共分散関数提案の流れ

□距離とそれを用いた共分散関数の定義

- 前提条件に整合する距離の定義と共分散関数の導出を行う

空間領域間について

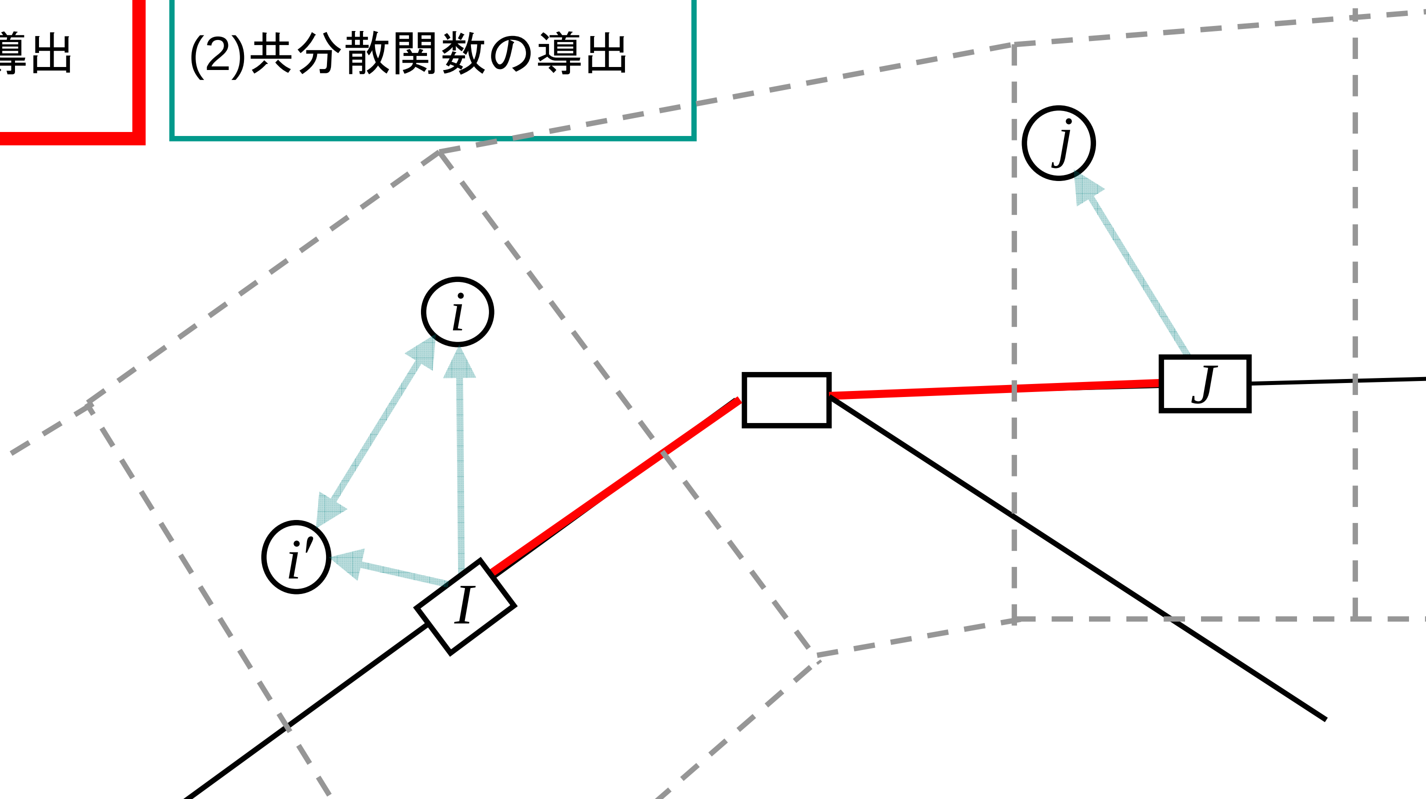
(1)距離の定義

(2)共分散関数の導出

空間領域内について

(1)距離の定義

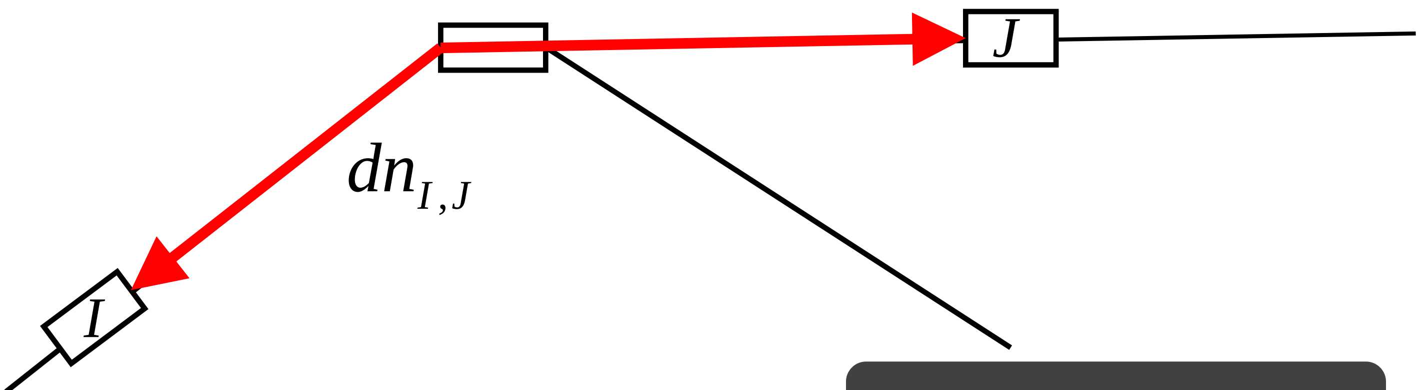
(2)共分散関数の導出



空間領域間の距離と共分散関数の定義（1）

□距離の定義

条件1：三角不等式・・・自明



次スライドで確認

□共分散関数の導出

条件2：正定値性・・・指数型共分散関数により満足される

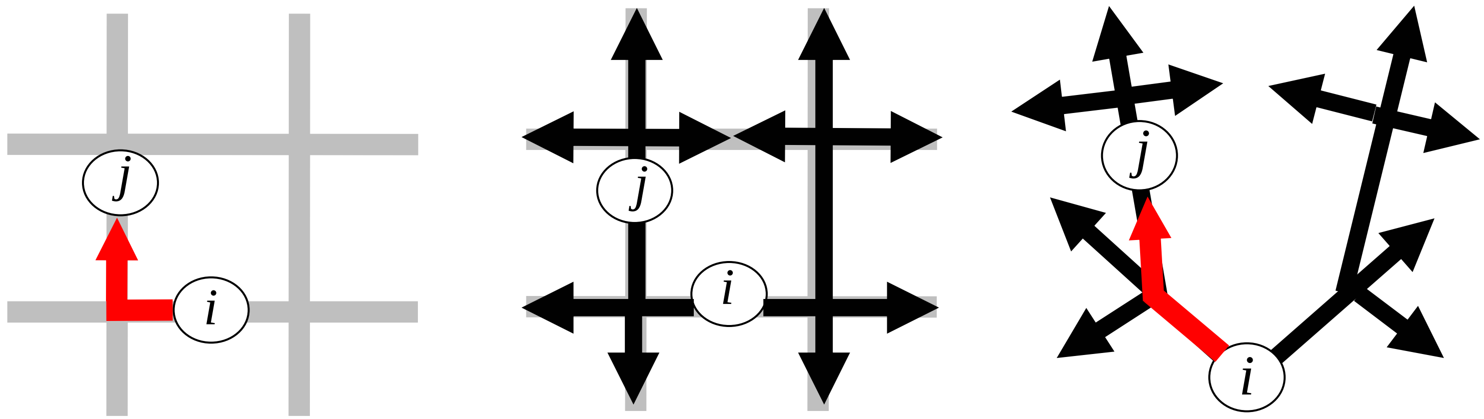
$$C(dn_{I,J}) = \begin{cases} \tau^2 + \sigma^2 \\ \sigma^2 \exp\left(-\dfrac{dn_{I,J}}{w_{dn}}\right) \end{cases}$$

※球型モデルなど一般に用いられる
他の共分散関数は不可

空間領域間の距離と共分散関数の定義(2)

■指数型共分散関数による正定値満足の確認

(1) ツリー(最短経路が一通りのみのグラフ)への変換により、
ネットワーク上の任意地点間の距離はツリー上の距離として表現可能



一般のネットワーク　　　　　　　　　　　　ツリー

(2) ツリー上の距離に指数型共分散関数を適用することで正定値性は保障されるHoef et al(2006)

(1),(2)より

空間領域間の距離に指数型共分散関数を適用した場合も正定値性保障

共分散関数提案の流れ

□距離とその関数である共分散関数の定義

- 前提条件に整合する距離の定義と共分散関数の導出を行う

空間領域間について

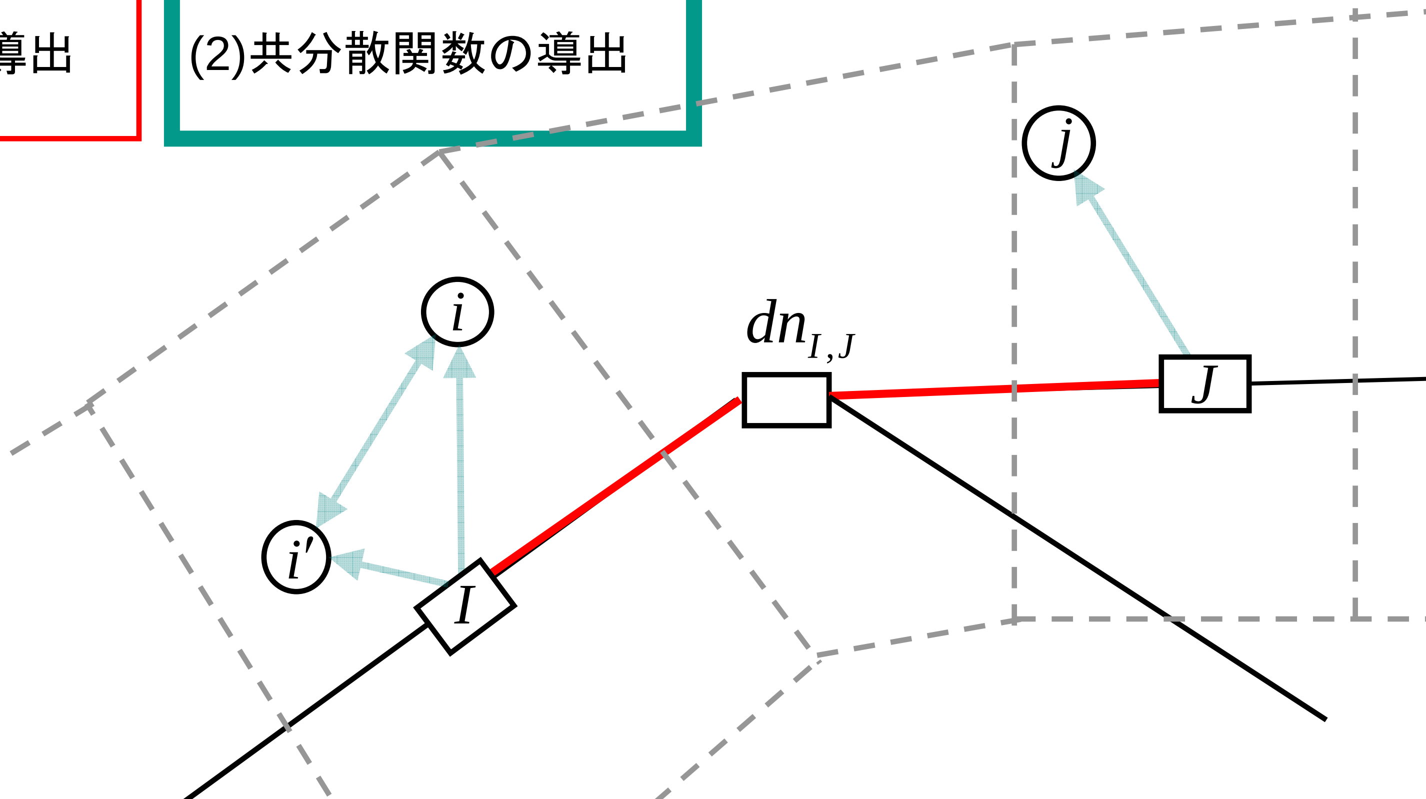
(1)距離の定義

(2)共分散関数の導出

空間領域内について

(1)距離の定義

(2)共分散関数の導出



空間領域内の距離と共分散関数の定義（1）

□距離の定義

- ネットワークからの空間従属性の波及を説明しうる距離を与える

■ 方法

- 疑似ネットワークを想定し、リンクの属性に重みを与えることで距離を定義

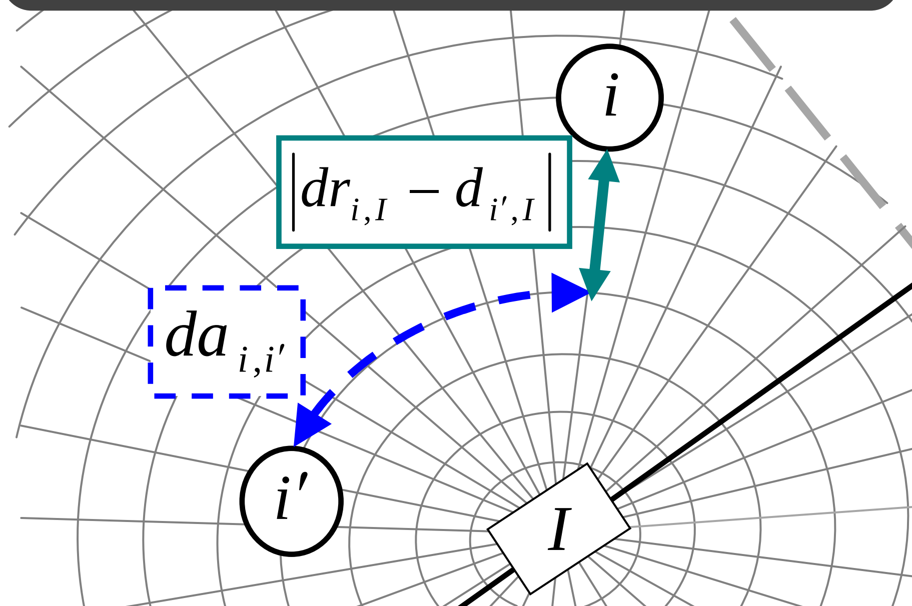
$\left| dr_{i,I} - d_{i',I} \right|$ ：ノードからの距離の差

$da_{i,i'}$ ：環状方向の距離

空間領域内での重み付け距離

$$d^{*}_{i,i'} = \frac{\left| dr_{i,I} - dr_{i',I} \right|}{w_{dr}} + \frac{da_{i,i'}}{w_{da}}$$

疑似ネットワーク
下のようなネットワークを高密に配置



空間領域内の距離と共分散関数の定義（2）

■ 条件1：三角不等式

- 距離関数は $d_{i,i'} \le d_{i,i''} + d_{i'',i'}$ の満足が不可欠

● 2種類の重み付け距離の比較

(1)2点間の経路距離：$d^*_{i,i'} = \dfrac{\left|dr_{i,I} - dr_{i',I}\right|}{w_{dr}} + \dfrac{da_{i,i'}}{w_{da}}$

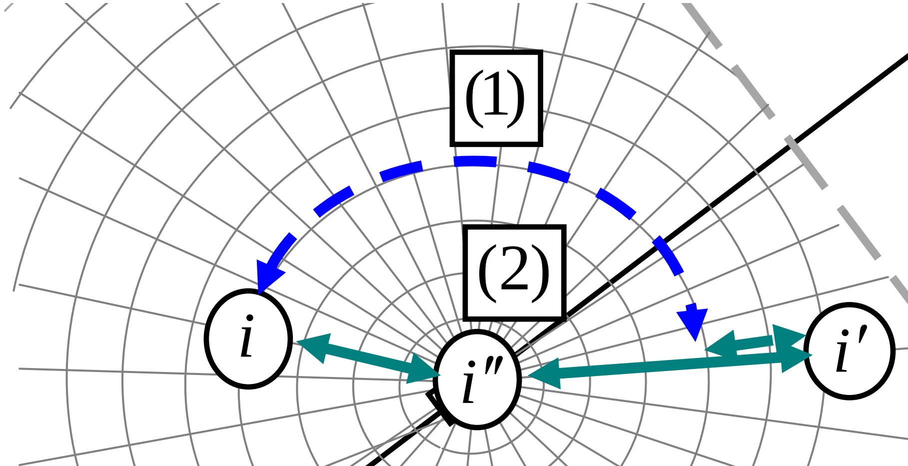
(2)ノードに近い中継点 i'' を経由した場合の経路距離

$$d^*_{i,i''} + d^*_{i'',i'} = \frac{\left|dr_{i,I} - dr_{i'',I}\right| + \left|dr_{i'',I} - dr_{i',I}\right|}{w_{dr}} + \frac{da_{i,i''} + da_{i'',i'}}{w_{da}}$$

三角不等式満足の条件

$$(1) \le (2)$$

$$w_{dr} \le (2/\pi) w_{da}$$



空間領域内の距離と共分散関数の定義(3)

□共分散関数の導出

■条件2:正定値性

空間領域内で定義したのは擬似的なネットワーク上の距離

↓

一般のネットワークと同様に指数型共分散関数により
正定性を満足することができるかは定かでない

↓

S-PLUSを用いたシミュレーションによる検証を実行

空間領域内の距離と共分散関数の定義(4)

■S-PLUSを用いた正定値性シミュレーションの概要

(1)(X,Y)=(0,0)にノードを仮定、X[-1,1]、Y[-1,1]の範囲に100地点配置

(2) 地点間の重み付け距離算出

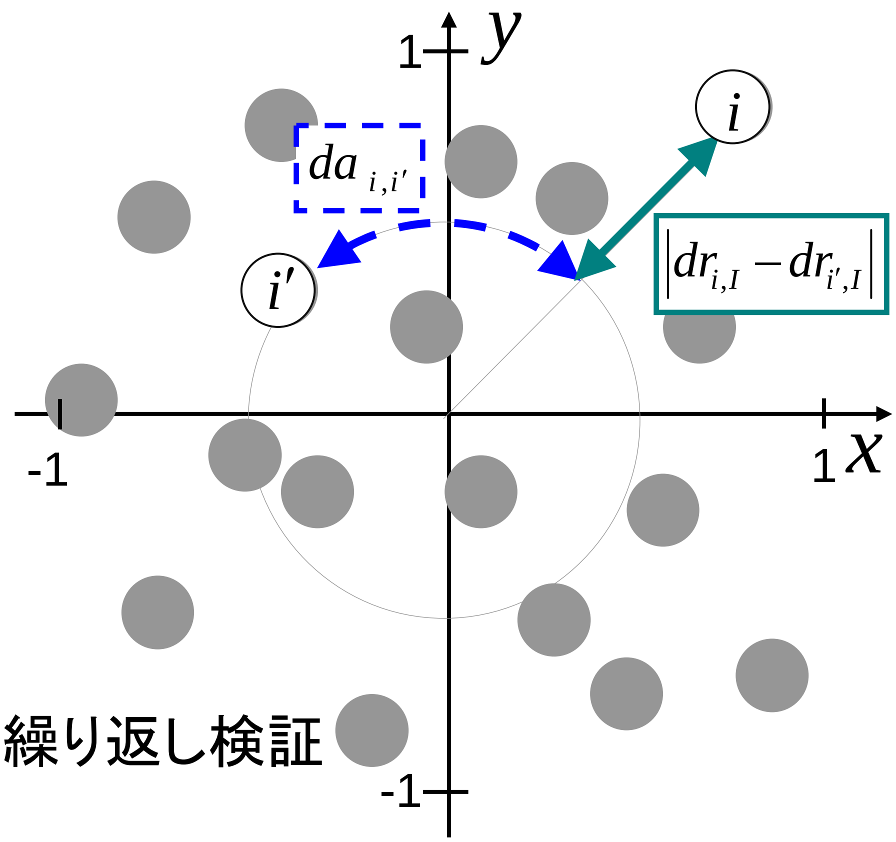
$$d^*_{i,i'} = \frac{\left| dr_{i,I} - dr_{i',I} \right|}{w_{dr}} + \frac{da_{i,i'}}{w_{da}}$$

(3)指数型共分散関数によって共分散算出

$$C(d^*_{i,i'}) = \sigma^2 \exp(d^*_{i,i'})$$

(4)得られた共分散の正定値性の検証

100地点の配置を100パターン変えながら繰り返し検証



正定値性崩れず＝指数型共分散関数は適用可能

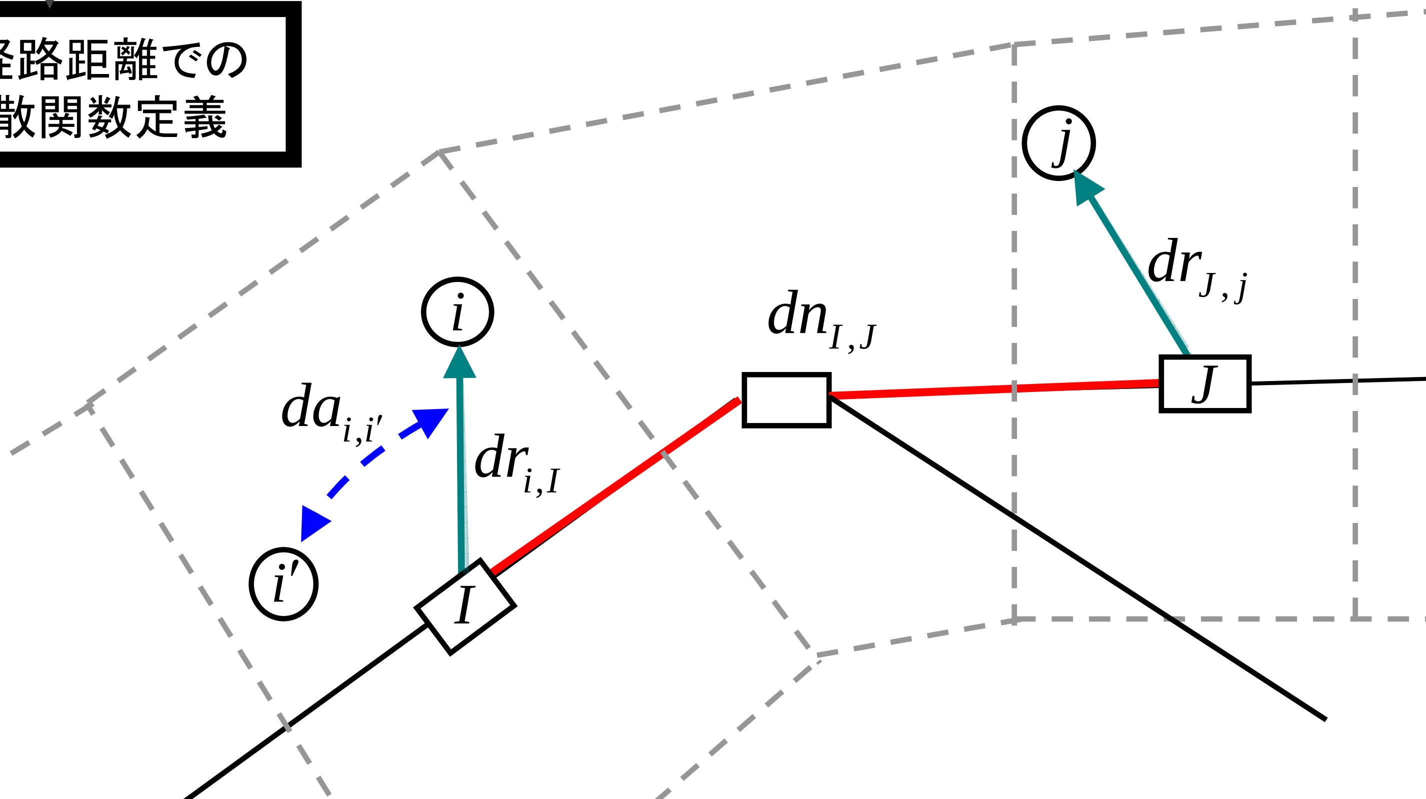
共分散関数提案の流れ

□距離とその関数である共分散関数の定義

空間領域間での定義
共分散関数: $C(dn)$

空間領域内での定義
共分散関数: $C(dr, da)$

最短経路距離での
共分散関数定義



最短経路距離の共分散関数

- 空間領域内、間どちらも指数型共分散関数が有効であったことから、全体に対して指数型共分散関数を適用

■重み付け最短経路距離

（1）異なる空間領域内の2点

$$d^*_{net} = \frac{dn_{I,J}}{w_{dn}} + \frac{dr_{i,I} + dr_{J,j}}{w_{dr}}$$

（2）同一空間領域内の2点

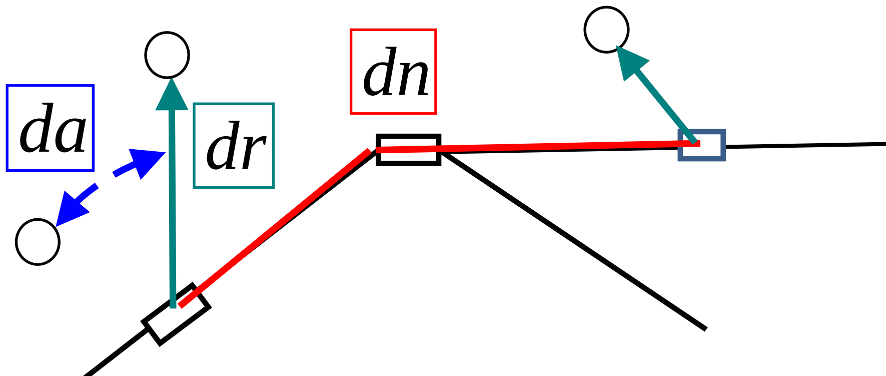
$$d^*_{net} = \frac{\left|dr_{i,I} - dr_{i',I}\right|}{w_{dr}} + \frac{da_{i,i'}}{w_{da}}$$

$$w_{dr} \le (\pi/2) w_{da}$$

■共分散関数

$$C(d^*_{net}) = \begin{cases} \tau^2 + \sigma^2 \\ \sigma^2 \exp(-d^*_{net}) \end{cases}$$

$$w_{dr} \le (\pi/2) w_{da}$$



共分散関数提案の流れ

□距離とその関数である共分散関数の定義

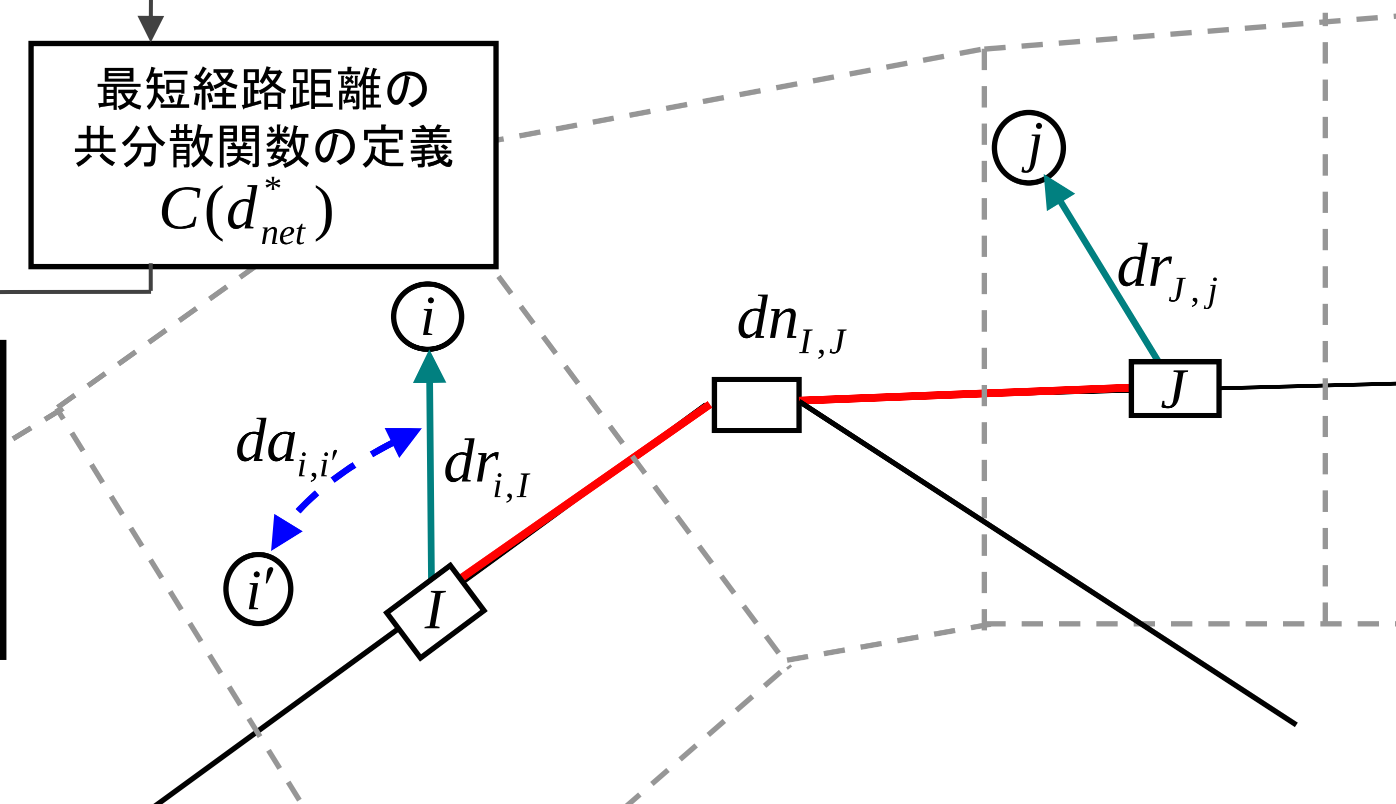
空間領域間での定義
共分散関数：$C(dn)$

空間領域内での定義
共分散関数：$C(dr, da)$

ユークリッド距離
の共分散関数
$C(d^{*}_{euclid})$

最短経路距離の
共分散関数の定義
$C(d^{*}_{net})$

ユークリッド距離と
最短経路距離の
共分散関数の定義
$C(d^{*}_{net}, d^{*}_{euclid})$



提案する共分散関数

□河川内挿での手法Cressie et al(2006)を応用

$$C(d^*_{net}, d^*_{euclid}) = \begin{cases} \tau^2 + \sigma^2 & if \quad s_i = s_j \\ \lambda C(d^*_{net}) + (1-\lambda) C(d^*_{euclid}) & if \quad s_i \neq s_j \end{cases}$$

■最短経路距離による影響

※ここれまでのスライドで導出した関数

・異なる空間領域内の2点

・同一空間領域内の2点

$$C(d^*_{net}) = \begin{cases} \sigma^2 \exp\left(-\frac{dn_{I,J}}{w_{dn}} - \frac{dr_{i,I} + dr_{J,j}}{w_{dr}} \right) \\ \sigma^2 \exp\left(-\frac{\left| dr_{i,I} - dr_{i',I} \right|}{w_{dr}} - \frac{da_{i,i'}}{w_{da}} \right) \end{cases}$$

$$w_{dr} \leq (\pi/2) w_{da}$$

■ユークリッド距離による影響

$$C(d^*_{euclid}) = \sigma^2 \exp(-\frac{d_{euclid}}{w_{euclid}})$$

※三角不等式、正定値性は簡単に示すことができる

共分散関数提案のまとめ

□距離とその関数である共分散関数の定義

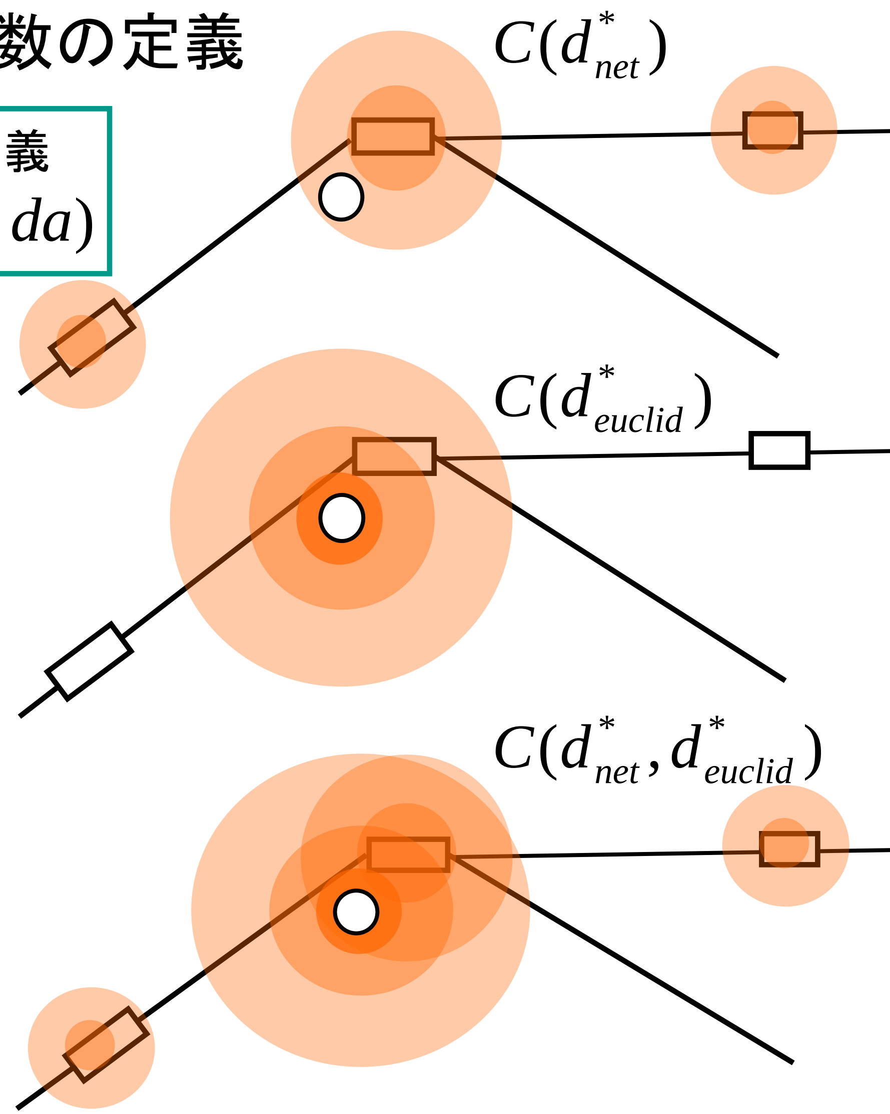
空間領域間での定義
共分散関数 : $C(dn)$

空間領域内での定義
共分散関数 : $C(dr, da)$

ユークリッド距離の共分散関数
$C(d^{*}_{euclid})$

最短経路距離の共分散関数の定義
$C(d^{*}_{net})$

ユークリッド距離と最短経路距離の共分散関数の定義
$C(d^{*}_{net}, d^{*}_{euclid})$



提案手法の有効性検証

実証

ロ方法：従来手法(ユークリッド距離のみによる共分散関数使用)との比較

- S-PLUSにより関数を作成
- データを8分割し、パラメータ推定を8回実行
- 各試行について残りの7/8の地点への
  地価推定を実行し、予測精度を比較

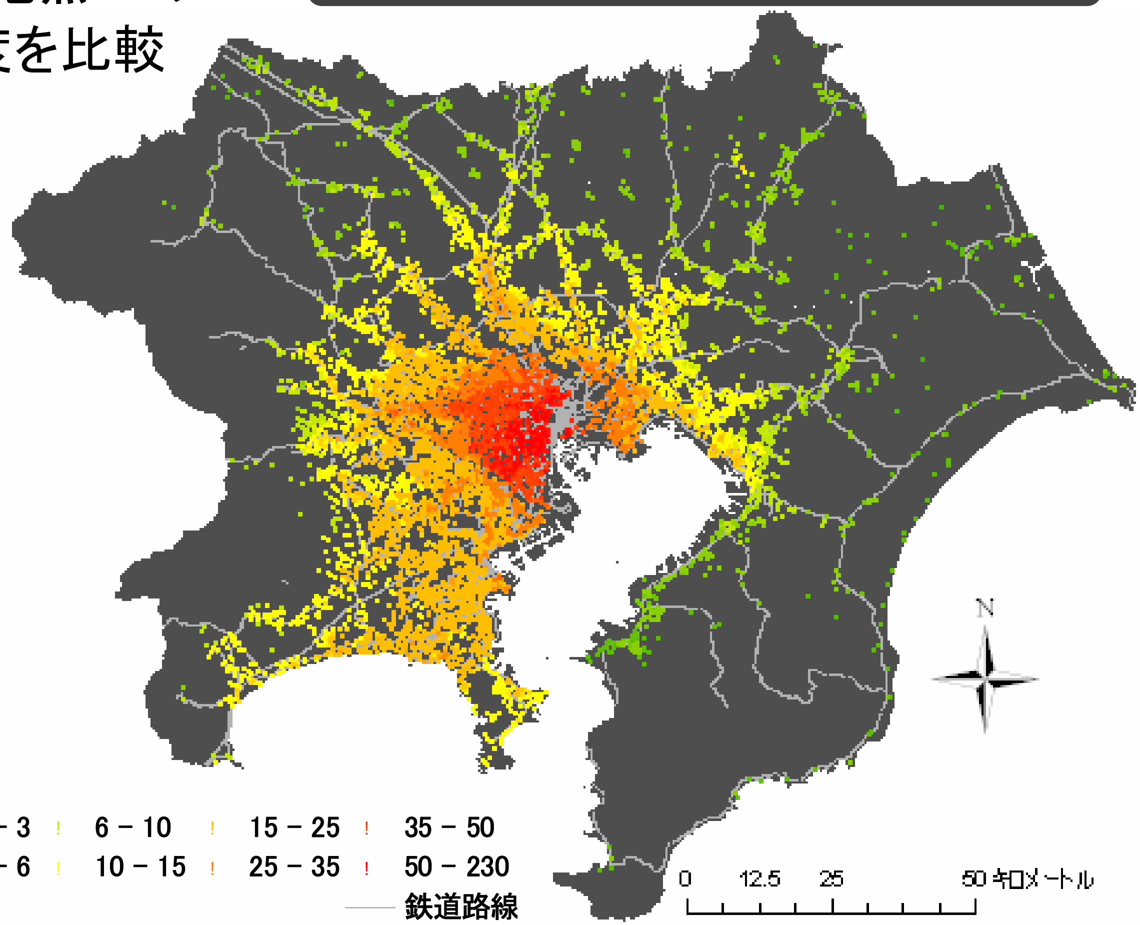
■ データ

- 首都圏の平成17年度住宅地
  公示地価データ6204点

■ 説明変数

- 最寄駅までの距離
- 最寄駅から東京駅までの
  鉄道ネットワーク距離
- 人口密度

地価分布の視覚化(ArcGIS利用)



パラメータ平均推定値

- 最尤法による8回のパラメータ推定結果の平均値

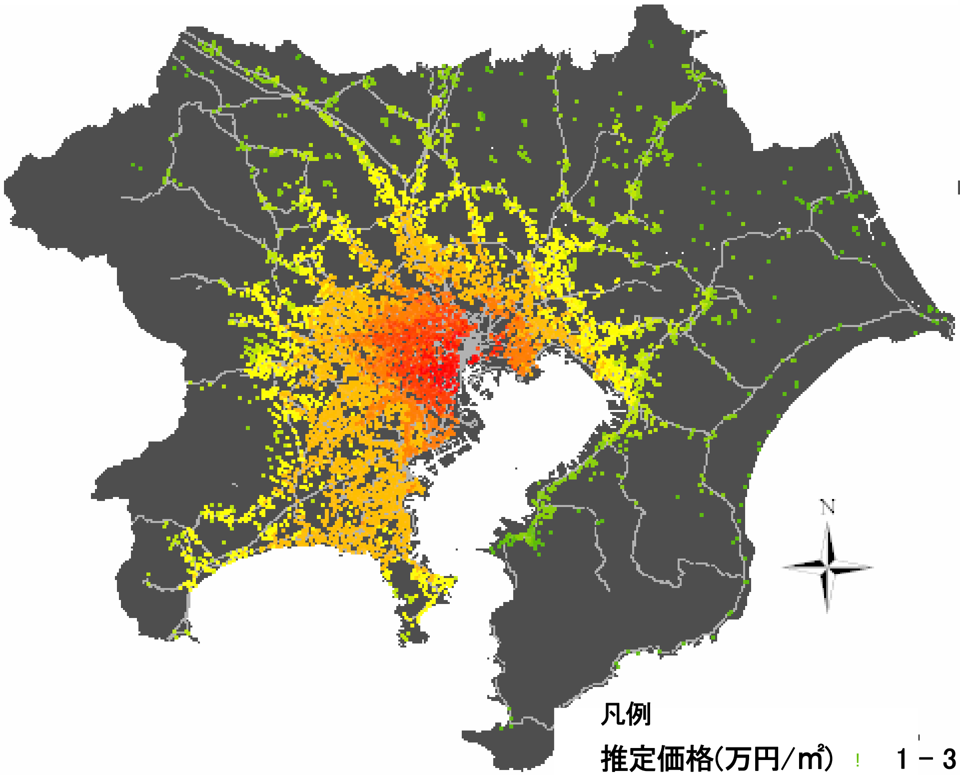
		提案	従来
各距離の空間従属性の強さ	w_{euclid}	39.39	39.36
	w_{dn}	69.74	
	w_{dr}	8.54	
	w_{da}	13.17	
最短経路距離による影響の割合	λ	0.50	
モデルのあてはまりの良さ	AIC	885	1138

- 空間従属性の50%は最短経路距離により説明される
- AICは8回の全試行について向上

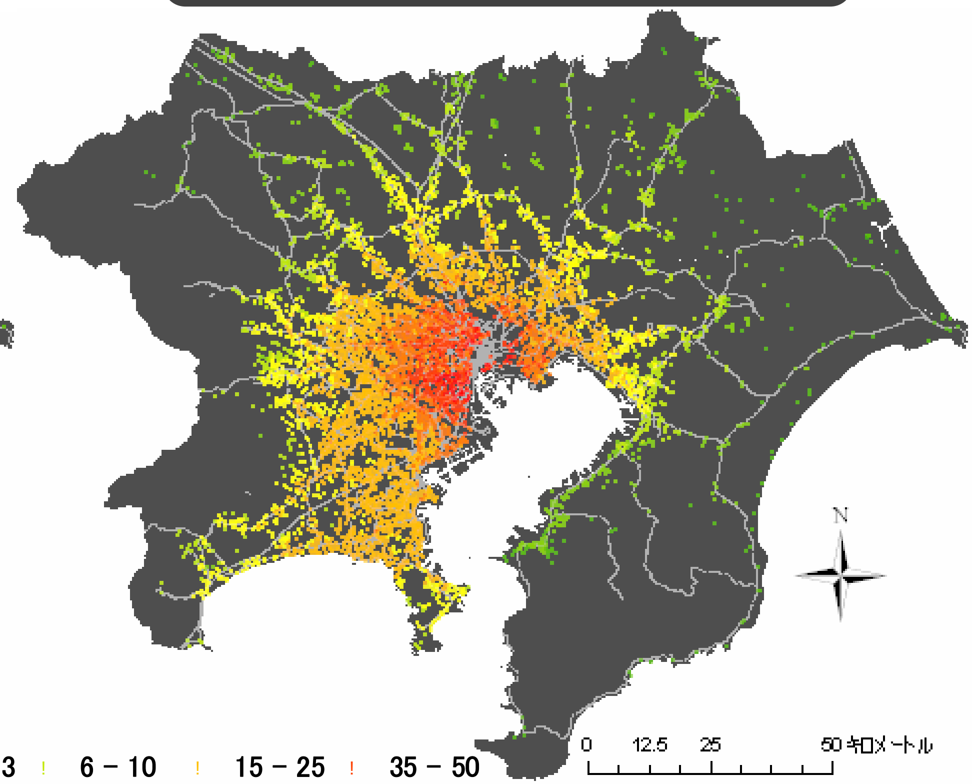
地価推定結果

実証

提案手法



従来手法



都心部及び都心から西へ30km程度の地域に比較的大きな差異

予測精度の比較

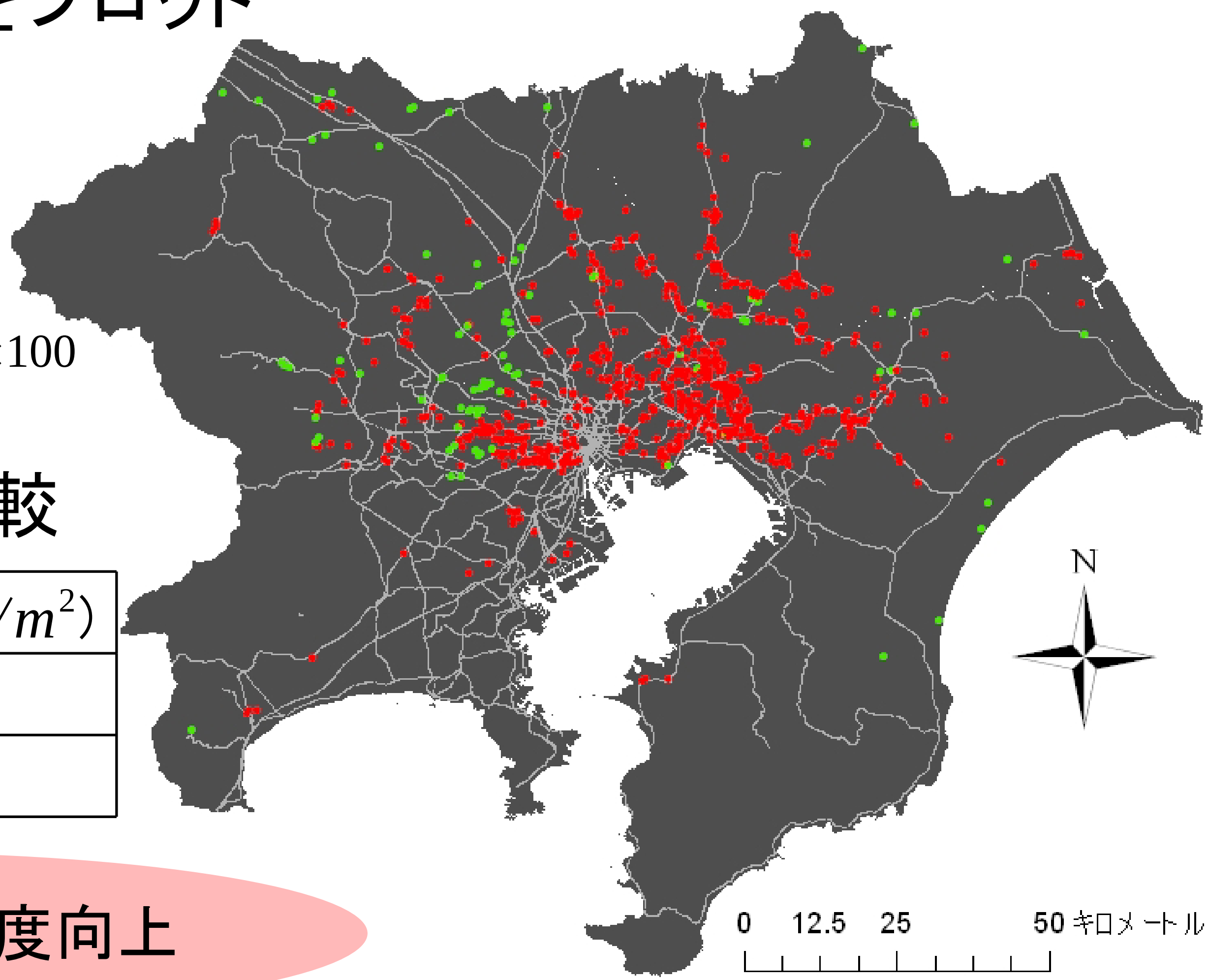
実証

☐平均誤差率の差をプロット

赤・・・15%以上精度向上

緑・・・15%以上精度悪化

$$\text{誤差率} = \sqrt{\sum_{i=1}^{n}\left(\frac{z(s_i)-\hat{z}(s_i)}{z(s_i)}\right)^2} \times 100$$



☐RMSE平均値の比較

	RMSE(単位:万円/m^2)
提案手法	7.21
従来手法	8.79

約20%の予測精度向上

まとめ

□研究成果のまとめ

- 最短経路距離を共分散関数に考慮したネットワーク外へ空間への内挿のためのKrigingの提案

■提案手法の特徴

- 空間従属性に及ぼす影響力をユークリッド距離と最短経路距離とで分離して計測することが可能

■実証から

- 最短経路距離が空間従属性に及ぼす影響の存在確認
- 首都圏における地価の精度向上

□S-PLUS使用に関するまとめ

- 扱いやすく、また、非常に多くの関数が実装されていたため、モデル構築やシミュレーションを容易に行うことができた。今後も、さらに高度な分析のためにプログラミング能力を身につけたいと思う。

参考文献

- 奥貫圭一(2008),GISを活用した空間分析,*地学雑誌*,117(2), pp324-330.
- Cressie, N. (1993). *Statistics for Spatial Data, Revised　Edition*. John Wiley & Sons.
- Cressie, N. Frey, J. Harch, B, and Smith, M. (2006). Spatial Prediction on a River Network. *Journal of Agricultural, Biological, and Environmental Statistics*. 11(2), 127-150.
- Curriero, F, C. (1996). On the Use of Non-Euclidean　Distance Measures in Geostatistics. *Mathematical Geology*, 38 (8), 907-926.
- Krivoruchko, K. Gribov, A.(2004). Geostatistical Interpolation and Simulation with Non-Euclidean Distances.Kluwer Academic Publishers, Barcelona,Spain,pp.331-342.
- Ver Hoef, J, M. Peterson, E. Theobald, D. (2006). Spatial Statistical Models that Use Flow and Stream Distance. *Environmental and Ecological Statistics*, 449-464.
- The Geological Society, http://www.geolsoc.org.uk/index.html /2009/10/29/15:00/最終アクセス