

OSS開発における共進化プロセス の理解のための遅延相関分析

山谷 陽亮 大平 雅雄 伊原 彰紀

60206052

岡坂亮汰

目次

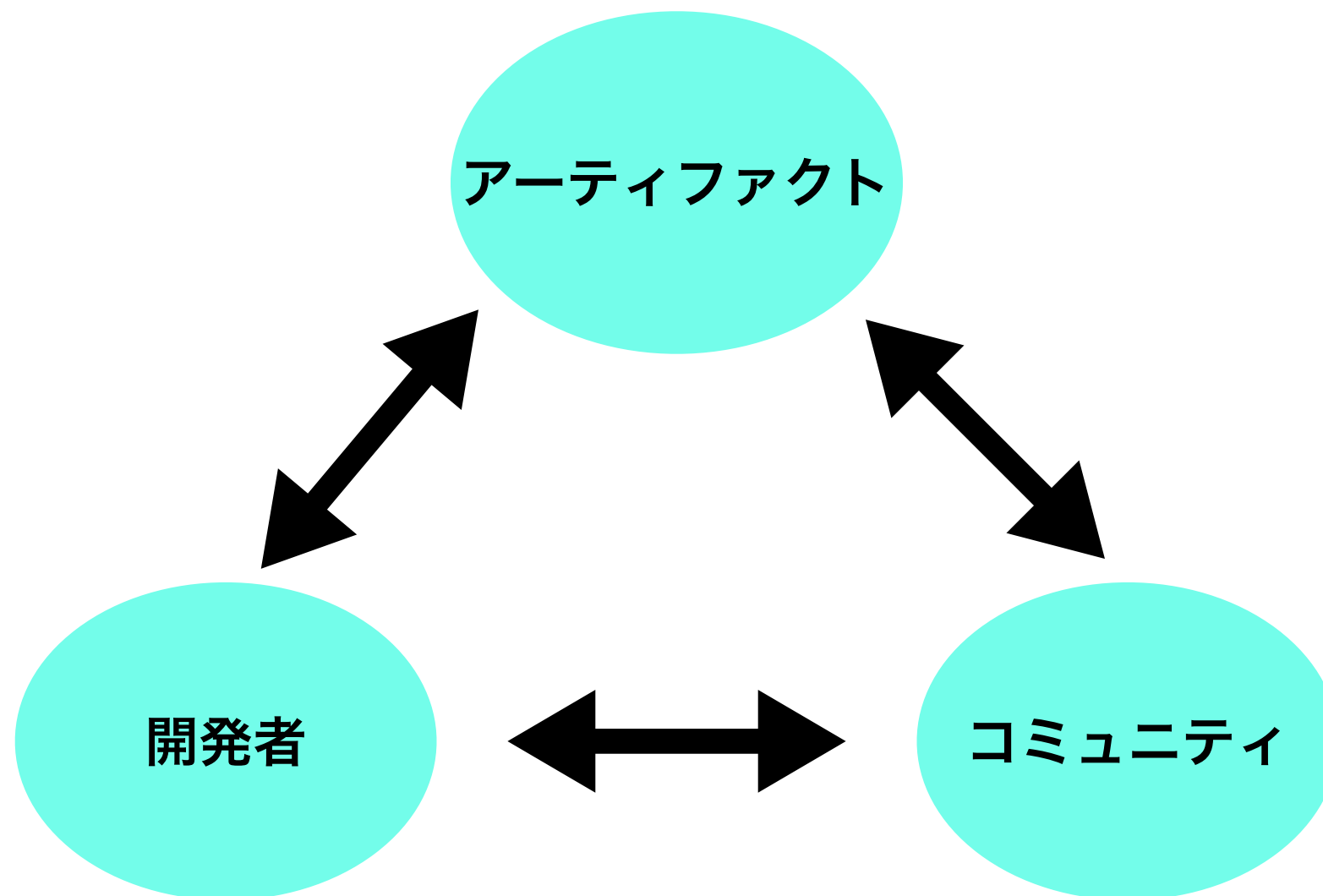
- OSSの進化について
- 遅延相関分析について
- OSSの進化に関するケーススタディ
- まとめ

OSSの進化

- OSSはボランティアによる開発者によって開発・保守されるので、プロジェクトが途中で崩壊することもある.
 - … この懸念を払拭するため、OSSの進化に関する研究が盛んに行われている.
- OSSにおける共進化のモデルが提案される.

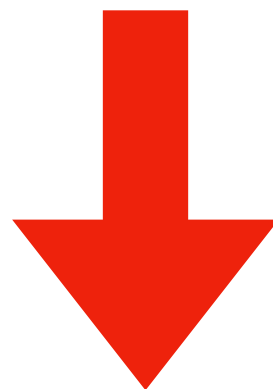
共進化

- ・ OSSのシステムとコミュニティが互いに影響を与え進化していく過程.



共進化

- ・ OSSにおける共進化のプロセスを定量的に表現することで、進化プロセスを予測することに繋がると考えられる.



遅延相関分析

遅延相関分析

- ・ 説明変数の値が累積した結果, 一定期間後に目的変数の値の変化に影響を与えると想定した相関分析手法.

※ 説明変数 … 何かの原因となっている変数.

目的変数 … その原因を受けて発生した結果となっている変数.

- ・ 3つのパラメータを定義.
 - 加算係数, 差分係数, 遅延係数

遅延相関分析

- 加算係数

- … 説明変数の値を累積させた期間

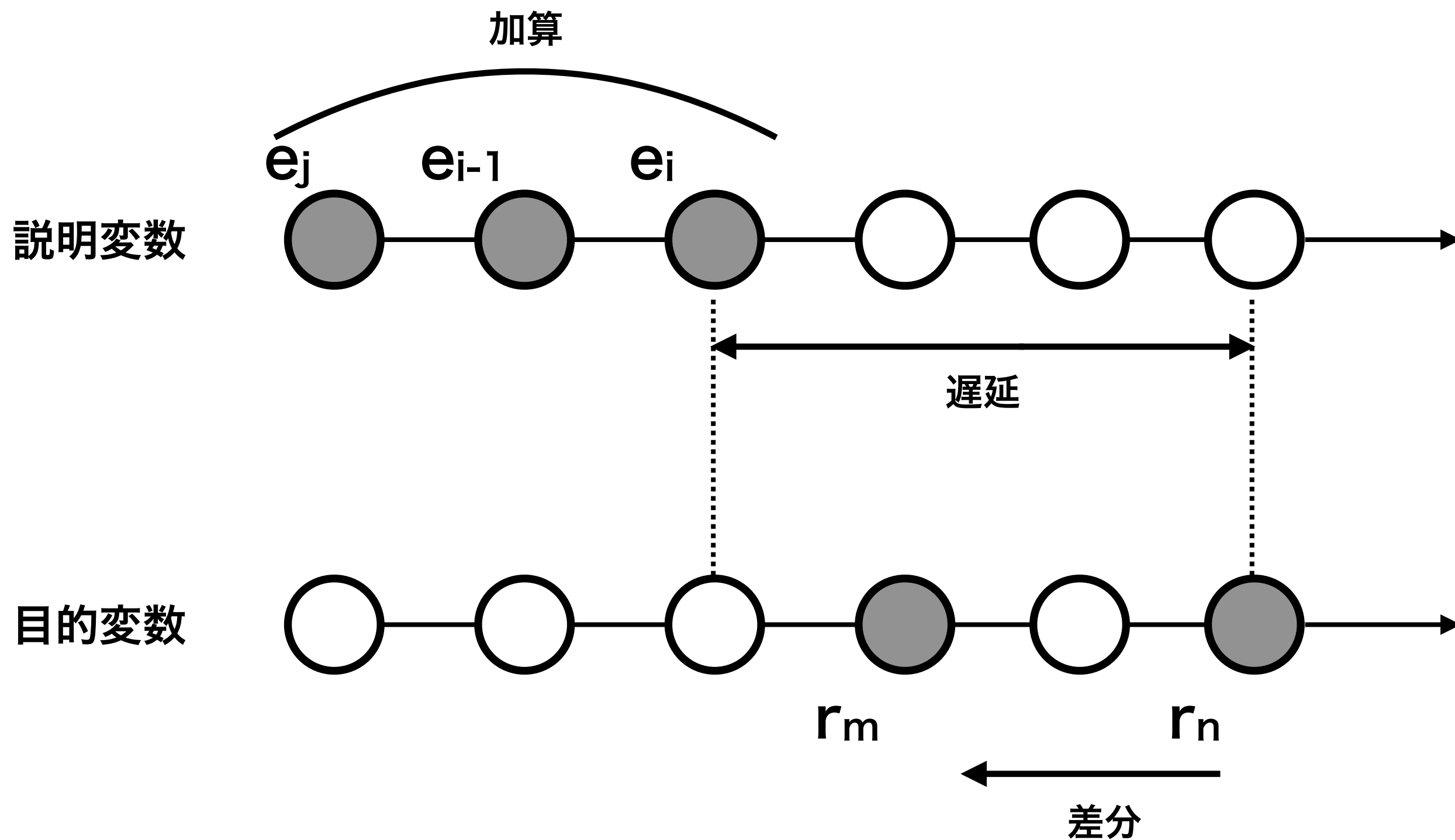
- 差分係数

- … 目的変数の値の変化を考慮する期間

- 遅延係数

- … 説明変数の累積が目的変数の変化に影響を与える期間

遅延相関分析



ケーススタディ

- アーティファクト, 開発者, コミュニティそれぞれにさまざまなメトリクスを定義し, ケーススタディを行う.
- Eclipse Platformプロジェクトを対象とし, 分析を行う.

ケーススタディ

進化に関するレイヤー	メトリクス名	メトリクスの説明
アーティファクト	AddLine	追加した行数
	AvgCyclomatic	各メソッドの複雑度の平均
	CountDeclClass	クラス宣言の数
	CountDeclFunction	関数宣言の数
	CountLineBlank	空白の行数
	CountLineCode	総コード行数
	CountLineCodeDecl	宣言部のコード行数
	CountLineCodeExe	処理を行うコード行数
	CountLineComment	コメントの行数
	CountStmtDecl	宣言部のステートメント数
	CountStmtExe	処理を行うステートメント数
	NumberOfRevisions	ファイルを改定した回数
開発者	Committer	ソースコードを書き換えた人数
	Reporter	バグを報告した人数
	Fixer	バグを修正した人数
コミュニティ	NumberOfForums	Eclipse Community Forums に投稿された回数
	NumberOfPeople	Eclipse Community Forums に投稿した人数

各レイヤーのメトリクス

アーティファクトと開発者の関係

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
追加した行数	バグを報告した人数	1	3	0	0.687
追加した行数	バグを修正した人数	1	3	0	0.523
各メソッドの複数度の平均	バグを修正した人数	3	0	3	-0.479
ファイルを改定した回数	ソースコードを書き換えた人	1	3	3	-0.588
ファイルを改定した回数	バグを報告した人数	1	3	0	0.591
ファイルを改定した回数	バグを修正した人数	2	3	3	-0.670

アーティファクト→開発者での結果

アーティファクトと開発者の関係

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
ソースコードを書き換えた人数	各メソッドの複数度の平均	3	3	3	0.533
ソースコードを書き換えた人数	総コード行数	2	3	0	0.491
ソースコードを書き換えた人数	宣言部のコード行数	2	3	0	0.533
ソースコードを書き換えた人数	宣言部のステートメント数	2	3	0	0.577
バグを報告した人数	追加した行数	3	0	2	-0.860
バグを報告した人数	各メソッドの複数度の平均	3	0	0	0.541
バグを報告した人数	関数宣言の数	3	3	0	-0.583
バグを報告した人数	空白の行数	3	3	0	-0.421
バグを報告した人数	総コード行数	3	3	1	-0.505
バグを報告した人数	宣言部のコード行数	3	3	2	-0.603
バグを報告した人数	処理を行うコード行数	3	3	0	-0.547
バグを報告した人数	宣言部のステートメント数	3	3	2	-0.620
バグを報告した人数	処理を行うステートメント数	3	3	0	-0.532
バグを報告した人数	ファイルを改定した回数	3	3	2	-0.713
バグを修正した人数	各メソッドの複数度の平均	3	3	2	0.490
バグを修正した人数	宣言部のコード行数	2	3	0	0.454
バグを修正した人数	宣言部のステートメント数	2	3	0	0.482
バグを修正した人数	ファイルを改定した回数	2	0	0	0.741

開発者→アーティファクトでの結果

アーティファクトと開発者の関係

- ソースコードに追加した行が増加すると, バグ報告者とバグ修正者が増加.
 - … 新しく追加されたソースコードにバグが混入
- コミット数やバグ修正者が増加すると, 宣言部の行数, 宣言部のステートメント数, 複雑度が増加.
 - … 独自に宣言した関数や変数が複雑となる

開発者とコミュニティの関係

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
バグを報告した人数	Eclipse Community Forums に投稿された回数	3	3	3	0.482
バグを報告した人数	Eclipse Community Forums に投稿した人数	3	0	2	-0.514
バグを修正した人数	Eclipse Community Forums に投稿された回数	2	3	0	-0.572

開発者→コミュニティでの結果

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
Eclipse Community Forums に投稿された回数	ソースコードを書き換えた人数	1	3	1	0.495
Eclipse Community Forums に投稿された回数	バグを報告した人数	3	3	2	0.419
Eclipse Community Forums に投稿された回数	バグを修正した人数	3	2	1	0.648
Eclipse Community Forums に投稿した人数	ソースコードを書き換えた人数	1	3	1	0.421
Eclipse Community Forums に投稿した人数	バグを修正した人数	1	0	2	0.443

コミュニティ→開発者での結果

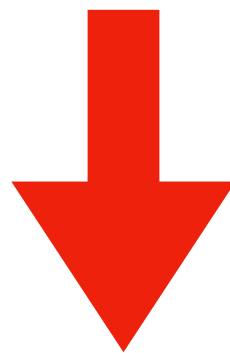
コミュニティとアーティファクトの関係

説明変数	目的変数	加算係数	差分係数	遅延係数	相関係数
Eclipse Community Forums に投稿された回数	追加した行数	3	0	2	0.592
Eclipse Community Forums に投稿された回数	各メソッドの複数 度の平均	3	0	2	-0.577
Eclipse Community Forums に投稿された回数	関数宣言の数	3	2	2	0.507
Eclipse Community Forums に投稿された回数	空白の行数	1	1	3	0.485
Eclipse Community Forums に投稿された回数	総コード行数	3	3	2	0.616
Eclipse Community Forums に投稿された回数	宣言部のコード行 数	3	2	2	0.694
Eclipse Community Forums に投稿された回数	処理を行うコード 行数	3	2	2	0.514
Eclipse Community Forums に投稿された回数	コメントの行数	3	1	3	0.526
Eclipse Community Forums に投稿された回数	宣言部のステート メント数	3	2	2	0.689
Eclipse Community Forums に投稿された回数	処理を行うステート メント数	3	2	2	0.537
Eclipse Community Forums に投稿された回数	ファイルを改定し た回数	3	2	2	0.621

コミュニティとアーティファクトでの結果

コミュニティの進化の影響

- Eclipse News Forums に投稿された回数が増加すると、コミット数, バグ修正者, バグ報告者, ソースコードの行数が増加.
 - Eclipse News Forumsの発展が2, 3ヶ月の遅延を持って、アーティファクトや開発者に影響を与える.



Eclipse News Forums を利用しやすいものにすることで、開発者やソースコードの規模が増加するのではないか.

まとめ

- OSSの進化を捉えるためには、OSSシステムとコミュニティの両方に着目する必要がある。
- 遅延相関分析を用いることで、相関が発生する時間的差分を考慮した分析を行うことができる。
- メトリクスを増やすことによって、OSSの進化をより明確にすることができる。