

課題1

① 対象国を「イギリス」とし、イギリスにおける訪日外客数の月次推移を状態空間モデルで解析します。

基本構造時系列モデルのパターン

- ローカルレベルモデル + 固定季節変動
- ローカルレベルモデル + 可変季節変動
- 平滑化トレンドモデル + 固定季節変動
- 平滑化トレンドモデル + 可変季節変動
- ローカル線形トレンドモデル + 固定季節変動
- ローカル線形トレンドモデル + 可変季節変動
- ローカルレベルモデル + 固定季節変動 (対数変換あり)
- ローカルレベルモデル + 可変季節変動 (対数変換あり)
- 平滑化トレンドモデル + 固定季節変動 (対数変換あり)
- 平滑化トレンドモデル + 可変季節変動 (対数変換あり)
- ローカル線形トレンドモデル + 固定季節変動 (対数変換あり)
- ローカル線形トレンドモデル + 可変季節変動 (対数変換あり)

```
irow <- "英国"
label_list <- c()
# ローカルレベルモデル + 固定季節変動
mod1 <- SSMModel(jnto[,irow] ~ SSMtrend(1, Q=NA) + SSMseasonal(12, Q=0), H=NA)
fit1 <- fitSSM(mod1, numeric(2), method = "SANN")
kfs1 <- KFS(fit1$model, smoothing=c("state", "mean", "disturbance"))
logLik1 <- kfs1$logLik - sum(kfs1$Finf > 0) * log(2*pi)/2
AIC1 <- -2*logLik1 + 2*( 2 + 12 )
label_list <- "ローカル_固定"
```

```
# ローカルレベルモデル + 可変季節変動
mod2 <- SSMModel(jnto[,irow] ~ SSMtrend(1, Q=NA) + SSMseasonal(12, Q=NA), H=NA)
fit2 <- fitSSM(mod2, numeric(3), method = "SANN")
kfs2 <- KFS(fit2$model, smoothing=c("state", "mean", "disturbance"))
logLik2 <- kfs2$logLik - sum(kfs2$Finf > 0) * log(2*pi)/2
AIC2 <- -2*logLik2 + 2*( 3 + 12 )
label_list <- c(label_list, "ローカル_可変")
```

```
# 平滑化トレンドモデル + 固定季節変動
mod3 <- SSMModel(jnto[,irow] ~ SSMtrend(2, Q=list(0, NA)) + SSMseasonal(12, Q=0), H=NA)
fit3 <- fitSSM(mod3, numeric(2), method = "SANN")
kfs3 <- KFS(fit3$model, smoothing=c("state", "mean", "disturbance"))
logLik3 <- kfs3$logLik - sum(kfs3$Finf > 0) * log(2*pi)/2
AIC3 <- -2*logLik3 + 2*( 2 + 13 )
label_list <- c(label_list, "平滑_固定")
```

```
# 平滑化トレンドモデル+可変季節変動
mod4 <- SSMModel(jnto[,irow] ~ SSMtrend(2, Q=list(0, NA)) + SSMseasonal(12, Q=NA), H=NA)
fit4 <- fitSSM(mod4, numeric(3), method = "SANN")
kfs4 <- KFS(fit4$model, smoothing=c("state", "mean", "disturbance"))
logLik4 <- kfs4$logLik - sum(kfs4$Finf > 0) * log(2*pi)/2
AIC4 <- -2*logLik4 + 2*( 3 + 13 )
label_list <- c(label_list, "平滑_可変")
```

```
# ローカル線形トレンドモデル+固定季節変動
mod5 <- SSMModel(jnto[,irow] ~ SSMtrend(2, Q=list(NA, NA)) + SSMseasonal(12, Q=0), H=NA)
fit5 <- fitSSM(mod5, numeric(3), method = "SANN")
kfs5 <- KFS(fit5$model, smoothing=c("state", "mean", "disturbance"))
logLik5 <- kfs5$logLik - sum(kfs5$Finf > 0) * log(2*pi)/2
AIC5 <- -2*logLik5 + 2*( 3 + 13 )
label_list <- c(label_list, "ローカル線形_固定")
```

```
# ローカル線形トレンドモデル+可変季節変動
mod6 <- SSMModel(jnto[,irow] ~ SSMtrend(2, Q=list(NA, NA)) + SSMseasonal(12, Q=NA), H=NA)
fit6 <- fitSSM(mod6, numeric(4), method = "SANN")
kfs6 <- KFS(fit6$model, smoothing=c("state", "mean", "disturbance"))
logLik6 <- kfs6$logLik - sum(kfs6$Finf > 0) * log(2*pi)/2
AIC6 <- -2*logLik6 + 2*( 4 + 13 )
label_list <- c(label_list, "ローカル線形_可変")
```

```
# ローカルレベルモデル+固定季節変動（対数変換あり）
mod7 <- SSMModel(log(jnto[,irow]) ~ SSMtrend(1, Q=NA) + SSMseasonal(12, Q=0), H=NA)
fit7 <- fitSSM(mod7, numeric(2), method = "SANN")
kfs7 <- KFS(fit7$model)
logLik7 <- kfs7$logLik - sum(kfs7$Finf>0) * log(2*pi)/2
AIC7 <- -2*logLik7 + 2*( 2 + 12 ) + 2*sum(log(jnto[,irow]))
label_list <- c(label_list, "ローカル_固定_log")
```

```
# mod8 ローカルレベルモデル+可変季節変動（対数変換あり）
mod8 <- SSMModel(log(jnto[,irow]) ~ SSMtrend(1, Q=NA) + SSMseasonal(12, Q=NA), H=NA)
fit8 <- fitSSM(mod8, numeric(3), method = "SANN")
kfs8 <- KFS(fit8$model, smoothing=c("state", "mean", "disturbance"))
logLik8 <- kfs8$logLik - sum(kfs8$Finf > 0) * log(2*pi)/2
AIC8 <- -2*logLik8 + 2*( 3 + 12 ) + 2*sum(log(jnto[,irow]))
label_list <- c(label_list, "ローカル_可変_log")
```

```
# mod9 平滑化トレンドモデル+固定季節変動（対数変換あり）
mod9 <- SSMModel(log(jnto[,irow]) ~ SSMtrend(2, Q=list(0, NA)) + SSMseasonal(12, Q=0), H=NA)
fit9 <- fitSSM(mod9, numeric(2), method = "SANN")
kfs9 <- KFS(fit9$model, smoothing=c("state", "mean", "disturbance"))
logLik9 <- kfs9$logLik - sum(kfs9$Finf > 0) * log(2*pi)/2
AIC9 <- -2*logLik9 + 2*( 2 + 13 ) + 2*sum(log(jnto[,irow]))
label_list <- c(label_list, "平滑_固定_log")
```

```
# mod10 平滑化トレンドモデル+可変季節変動（対数変換あり）
mod10 <- SSMModel(log(jnto[, irow]) ~ SSMtrend(2, Q=list(0, NA)) + SSMseasonal(12, Q=NA), H=NA)
fit10 <- fitSSM(mod10, numeric(3), method = "SANN")
kfs10 <- KFS(fit10$model, smoothing=c("state", "mean", "disturbance"))
logLik10 <- kfs10$logLik - sum(kfs10$Finf > 0) * log(2*pi)/2
AIC10 <- -2*logLik10 + 2*( 3 + 13 ) + 2*sum(log(jnto[, irow]))
label_list <- c(label_list, "平滑_可変_log")
```

```
# mod11 ローカル線形トレンドモデル+固定季節変動（対数変換あり）
mod11 <- SSMModel(log(jnto[, irow]) ~ SSMtrend(2, Q=list(NA, NA)) + SSMseasonal(12, Q=0), H=NA)
fit11 <- fitSSM(mod11, numeric(3), method = "SANN")
kfs11 <- KFS(fit11$model, smoothing=c("state", "mean", "disturbance"))
logLik11 <- kfs11$logLik - sum(kfs11$Finf > 0) * log(2*pi)/2
AIC11 <- -2*logLik11 + 2*( 3 + 13 ) + 2*sum(log(jnto[, irow]))
label_list <- c(label_list, "ローカル線形_固定_log")
```

```
# mod12 ローカル線形トレンドモデル+可変季節変動（対数変換あり）
mod12 <- SSMModel(log(jnto[, irow]) ~ SSMtrend(2, Q=list(NA, NA)) + SSMseasonal(12, Q=NA), H=NA)
fit12 <- fitSSM(mod12, numeric(4), method = "SANN")
kfs12 <- KFS(fit12$model, smoothing=c("state", "mean", "disturbance"))
logLik12 <- kfs12$logLik - sum(kfs12$Finf > 0) * log(2*pi)/2
AIC12 <- -2*logLik12 + 2*( 4 + 13 ) + 2*sum(log(jnto[, irow]))
label_list <- c(label_list, "ローカル線形_可変_log")
```

- 最も良いモデルとそのAICは以下の通り。

```
AIC_list <- c(AIC1, AIC2, AIC3, AIC4, AIC5, AIC6, AIC7, AIC8, AIC9, AIC10, AIC11, AIC12)
AIC_df <- data.frame(AIC_list, row.names = label_list)
Mod_name <- paste("最良モデル:", rownames(AIC_df)[which.min(AIC_df$AIC_list)])
min_AIC <- paste("AIC:", AIC_df[which.min(AIC_df$AIC_list),])
print(paste(Mod_name, min_AIC))
```

```
## [1] "最良モデル: ローカル線形_可変 , AIC: 3338.56698180584"
```

- なお、それぞれのモデルのAICは下表の通り。

```
print(AIC_df)
```

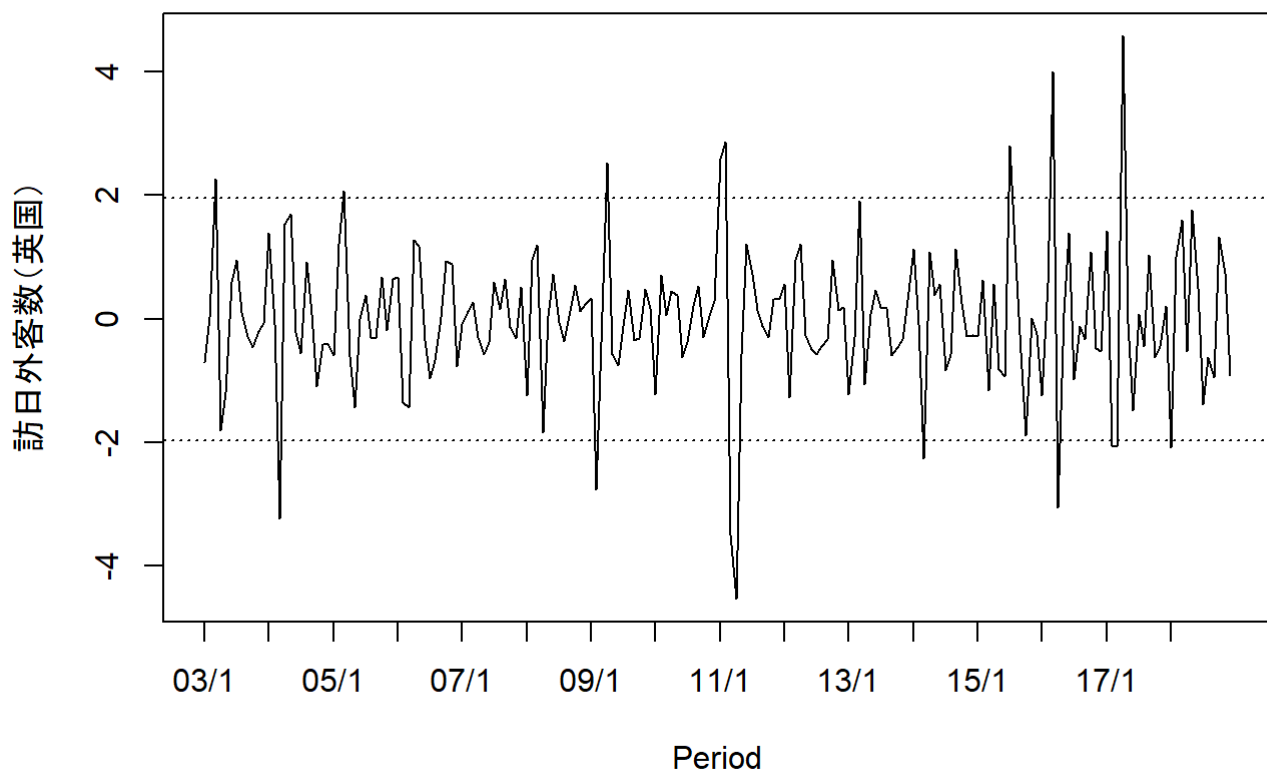
```
##           AIC_list
## ローカル_固定      3726.663
## ローカル_可変      3363.105
## 平滑_固定          3371.380
## 平滑_可変          3345.256
## ローカル線形_固定  3372.949
## ローカル線形_可変  3338.567
## ローカル_固定_log  3568.722
## ローカル_可変_log  3558.723
## 平滑_固定_log     3603.288
## 平滑_可変_log     3600.480
## ローカル線形_固定_log 3583.510
## ローカル線形_可変_log 3573.410
```

② 続いて、①で最も良いAICを示したローカルレベルモデル+可変季節変動の残差分析を行う

- 観測値攪乱項の残差プロットはこちら

```
plot(rstandard(kfs2, "pearson"), xaxt="n", xlab="Period", ylab="訪日外客数（英国）", main="観測値攪乱項の標準化残差")
abline(h=c(-1.96, 1.96), lty=3)
axis(side=1, at=1+0:15*12, labels=c("03/1", "04/1", "05/1", "06/1", "07/1", "08/1", "09/1", "10/1", "11/1", "12/1", "13/1", "14/1", "15/1", "16/1", "17/1", "18/1"))
```

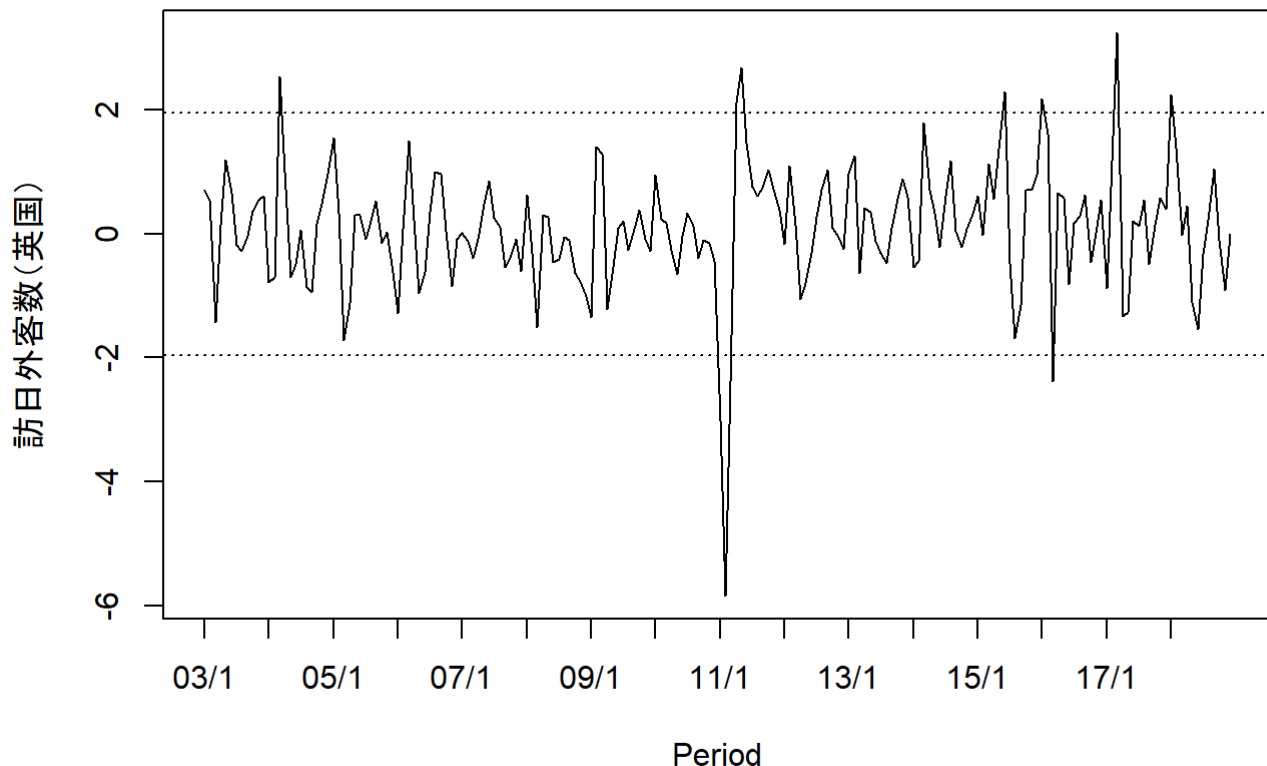
観測値攪乱項の標準化残差



- 状態攪乱項の残差プロットはこちら

```
plot(rstandard(kfs2, "state")[, 1], xaxt="n", xlab="Period", ylab="訪日外客数（英国）", main="状態攪乱項の標準化残差")
abline(h=c(-1.96, 1.96), lty=3)
axis(side=1, at=1+0:15*12, labels=c("03/1", "04/1", "05/1", "06/1", "07/1", "08/1", "09/1", "10/1", "11/1", "12/1", "13/1", "14/1", "15/1", "16/1", "17/1", "18/1"))
```

状態攪乱項の標準化残差



残差分析の結果から分かること

- 観測値攪乱項について、2011年3月付近に負の異常残差がみられる。
 - これは、東日本大震災による一時的な減少と推測する。
 - 打ち手としては以下の2点を考案する。
 - 異常値として除外する
 - 短期変動として切り出す
- 観測値攪乱項において、2015年9月あたりを境に定期的に正の異常残差がみられる。
 - 状態攪乱項の残差の推移など総合的に検討すると、訪日外客数の上昇トレンドが発生していると考えられる。
 - 打ち手としては以下を考案する。
 - 判断基準（※）を用いて期間を設定し、それ以降の期間で上昇トレンドの成分を加える（干渉変数で構造を変える） ※判断基準について：昨対比、前月比あたりの伸び率を算出し、可視化。違いが現れるようならそれを基準にしたい。

2011年3月付近のデータを異常値として除外したモデルを構築する。

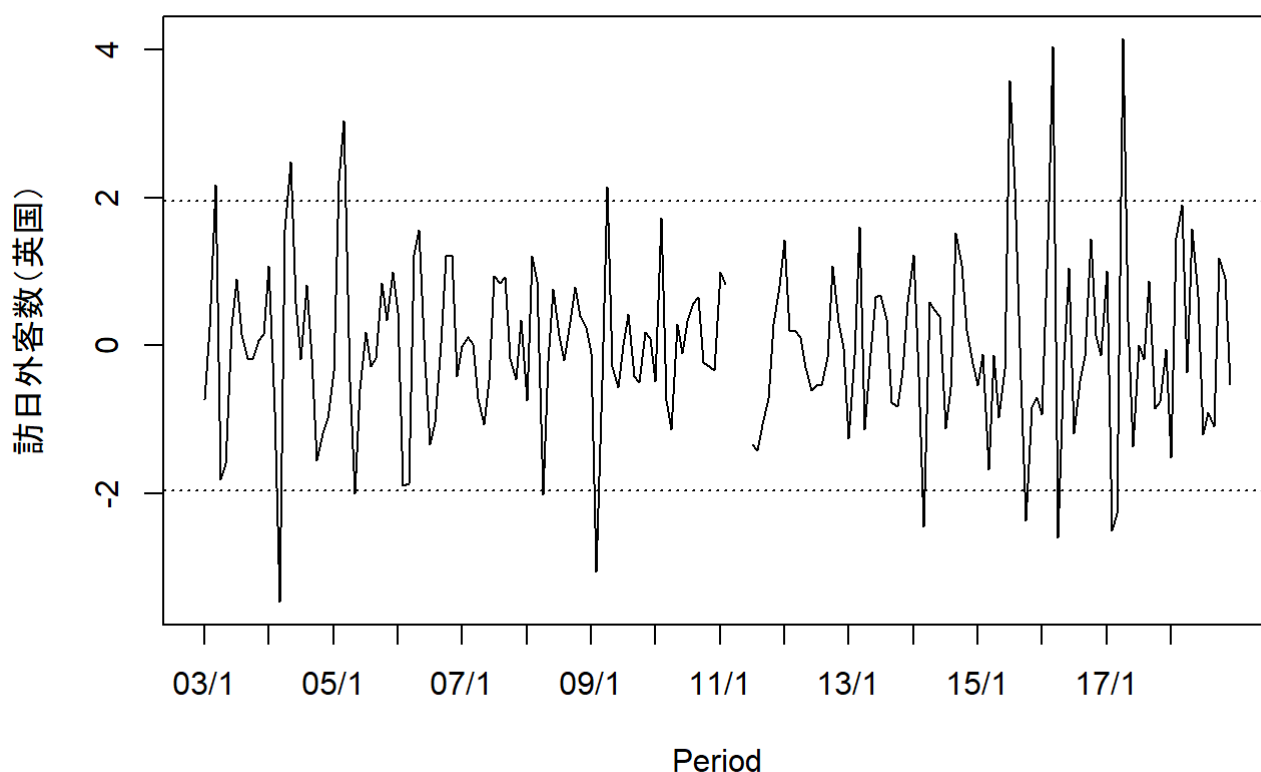
```
# 2011年3-5月のデータを除外
jntoNA <- jnto
jntoNA[jntoNA[, "年月"] %in% c("2011年3月", "2011年4月", "2011年5月", "2011年6月")==TRUE, 2:22]
<- NA

mod_NA2 <- SSMModel(jntoNA[, irow] ~ SSMtrend(1, Q=NA) + SSMseasonal(12, Q=NA), H=NA)
fit_NA2 <- fitSSM(mod_NA2, numeric(3), method = "SANN")
kfs_NA2 <- KFS(fit_NA2$model, smoothing=c("state", "mean", "disturbance"))
logLik_NA2 <- kfs_NA2$logLik - sum(kfs_NA2$Finf > 0) * log(2*pi)/2
AIC_NA2 <- -2*logLik_NA2 + 2*( 3 + 12 )
```

- 観測値攪乱項の残差プロットはこちら

```
plot(rstandard(kfs_NA2, "pearson"), xaxt="n", xlab="Period", ylab="訪日外客数（英国）", main="観測値攪乱項の標準化残差")
abline(h=c(-1.96, 1.96), lty=3)
axis(side=1, at=1+0:15*12, labels=c("03/1", "04/1", "05/1", "06/1", "07/1", "08/1", "09/1", "10/1", "11/1", "12/1", "13/1", "14/1", "15/1", "16/1", "17/1", "18/1"))
```

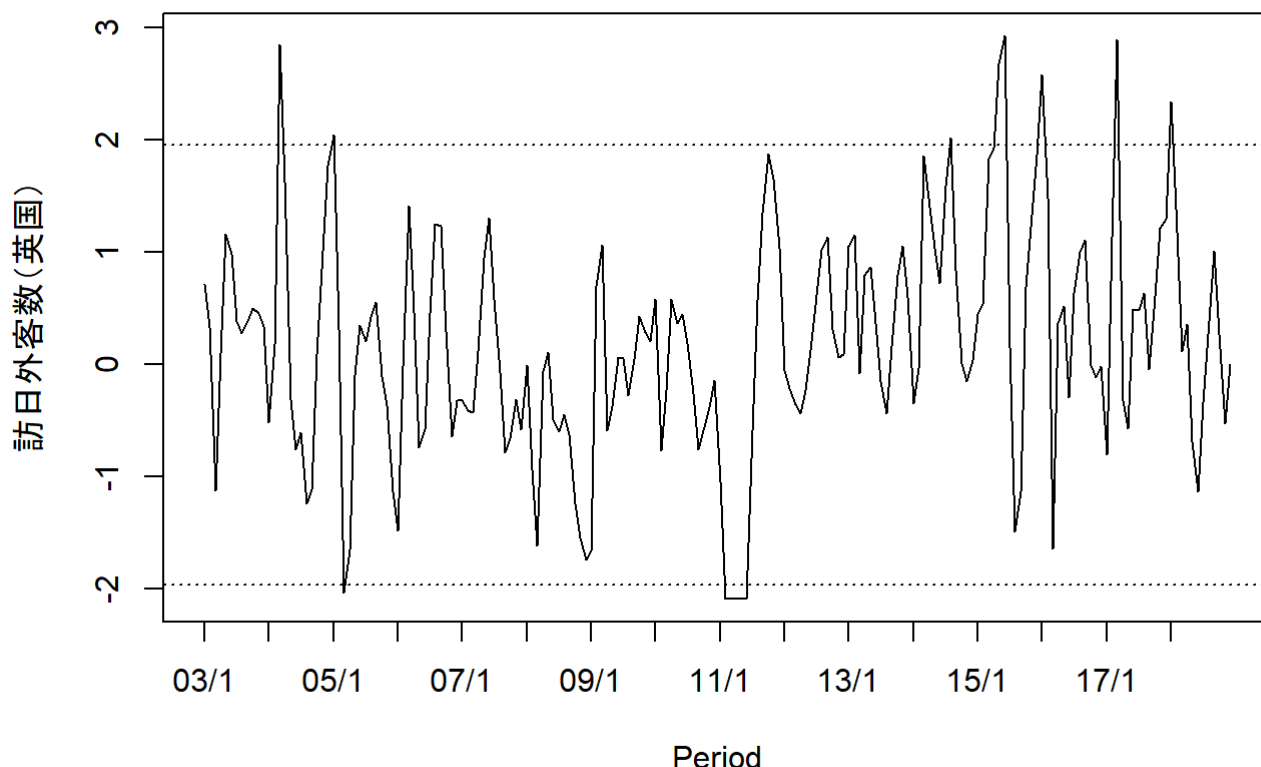
観測値攪乱項の標準化残差



- 状態攪乱項の残差プロットはこちら

```
plot(rstandard(kfs_NA2, "state")[, 1], xaxt="n", xlab="Period", ylab="訪日外客数（英国）", main="状態攪乱項の標準化残差")
abline(h=c(-1.96, 1.96), lty=3)
axis(side=1, at=1+0:15*12, labels=c("03/1", "04/1", "05/1", "06/1", "07/1", "08/1", "09/1", "10/1", "11/1", "12/1", "13/1", "14/1", "15/1", "16/1", "17/1", "18/1"))
```

状態攪乱項の標準化残差



- 2つの残差プロットを見ると、観測値攪乱項の残差において、2010年代前半ごろまでの極端な異常値はなくなったように見える
- しかしながら、依然として2010年代の上昇トレンド傾向が残差の異常値として現れているため、こちらに対しても対応が必要なのが伺える。

短期変動成分を加えたモデルを構築する。

- 2011年春ごろの観測値を除外したデータに、短期変動成分を加える。

```
# model2に定常AR成分を導入
mod13_NA <- SSMModel(log(jntoNA[,irow]) ~ SSMtrend(1, Q= NA) + SSMseasonal(12, Q=NA) + SSMarima
(ar = 0, Q = NA), H = NA)

updatefn <- function(pars, model){
  SSMModel(log(jntoNA[,irow]) ~ SSMtrend(1, Q = exp(pars[1]))
    + SSMseasonal(12, Q=exp(pars[2]))
    + SSMarima(ar = artransform(pars[3]), Q = exp(pars[4])),
    H = exp(pars[5]))}

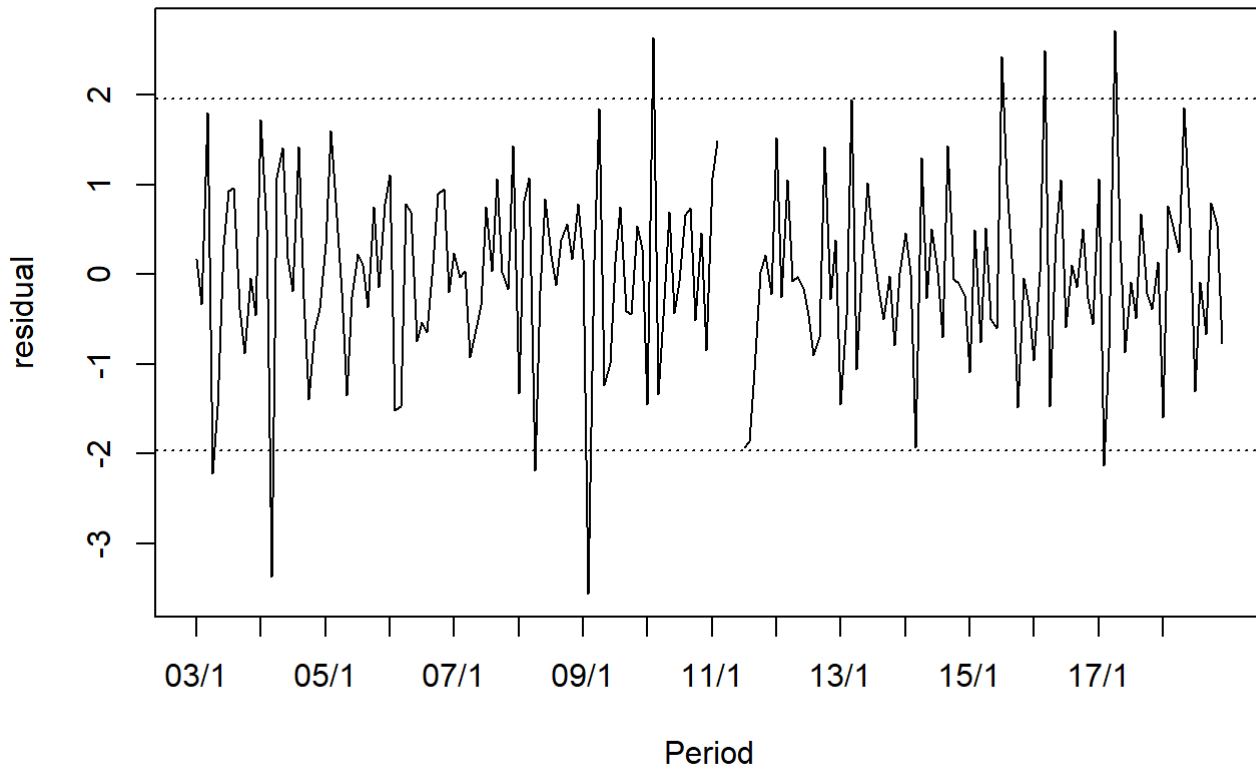
fit13_NA <- fitSSM(mod13_NA, numeric(5), updatefn, method = "BFGS")
kfs13_NA <- KFS(fit13_NA$model, smoothing=c("state", "mean", "disturbance"))
logLik13_NA <- kfs13_NA$logLik - sum(kfs13_NA$Finf > 0) * log(2*pi)/2
AIC13_NA <- -2*logLik13_NA + 2*( 3 + 12 )

# 推定されたAR係数を確認
tail(fit13_NA$model$T, 1)
```

```
## [1] 0.05661548
```

- 観測値攪乱項の残差プロットはこちら

```
plot(rstandard(kfs13_NA, "pearson"), xaxt="n", xlab="Period", ylab="residual")
abline(h=c(-1.96, 1.96), lty=3)
axis(side=1, at=1+0:15*12, labels=c("03/1", "04/1", "05/1", "06/1", "07/1", "08/1", "09/1", "10/1", "11/1", "12/1", "13/1", "14/1", "15/1", "16/1", "17/1", "18/1"))
```



- 状態攪乱項の残差プロットはこちら

```
plot(rstandard(kfs13_NA, "state")[,1], xaxt="n", xlab="Period", ylab="residual")
abline(h=c(-1.96, 1.96), lty=3)
axis(side=1, at=1+0:15*12, labels=c("03/1", "04/1", "05/1", "06/1", "07/1", "08/1", "09/1", "10/1", "11/1", "12/1", "13/1", "14/1", "15/1", "16/1", "17/1", "18/1"))
```