

Estatística Espacial (MI418) / Geoestatística (ME907)

Guilherme Ludwig

2019-01-22

Dados

Estatísticas descritivas espaço-temporais

Dados

- ▶ Instituto Nacional de Meteorologia:
<http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesAutomaticas>
- ▶ Site do C. Wikle: <http://faculty.missouri.edu/~wiklec/datasets.html>
- ▶ Curso de Geoestatística do Y. Zhukov: <https://sites.lsa.umich.edu/zhukov>, em particular
 - ▶ Áreas e elevações no globo: https://www.usgs.gov/centers/eros#/Find_Data
 - ▶ Censo americano: <http://lakshmi.calit2.uci.edu/census2000/>
 - ▶ Geopolítica: <https://www.prio.org/Data/CSCW-Replication-Data/>
 - ▶ Conflitos internacionais: <https://www.acleddata.com/>
- ▶ Kaggle

Tropical Pacific Sea Surface Temperature (SST) Anomalies

These data represent gridded monthly SST anomalies for 399 consecutive months from January 1970 through March 2003. The data were obtained from the IRI/LDEO Climate Data Library at Columbia University (<http://iridl.ldeo.columbia.edu/>). The data are gridded at a 2 degree by 2 degree resolution and represent anomalies from a January 1970 - December 1985 monthly (average) climatology. A more complete description can be found at <http://iridl.ldeo.columbia.edu/SOURCES/.CAC/>.

Fonte: <http://faculty.missouri.edu/~wiklec/datasets.html> e Cressie and Wikle (2011)

Tropical Pacific Sea Surface Temperature (SST) Anomalies

Preparando os dados:

```
loc <- read.table("SSTlonlat.dat", header = FALSE)
Z <- read.table("SST011970_032003.dat", header = FALSE)
landmask <- read.table("SSTlandmask.dat", header = FALSE)
colnames(loc) <- c("x", "y")
temp <- outer(1:12, # Month
              1970:2003, # Year
              function(x, y) paste0(x, "-", y, "-1"))[1:399]
colnames(Z) <- as.character(temp)
colnames(landmask) <- "mask"
```

Tropical Pacific Sea Surface Temperature (SST) Anomalies

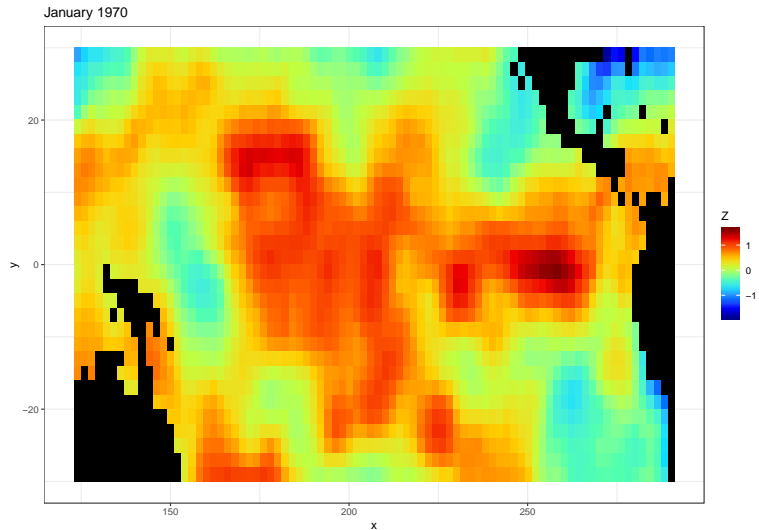
```
library(tidyverse)
SST <- cbind(loc, Z, landmask) %>%
  as.tibble %>%
  gather(date, Z, contains("-")) %>%
  mutate(date = as.Date(date, format = "%m-%Y-%d")) %>%
  mutate(mask = factor(mask, levels = 0:1,
                       label = c("sea", "land")))
```

Tropical Pacific Sea Surface Temperature (SST) Anomalies

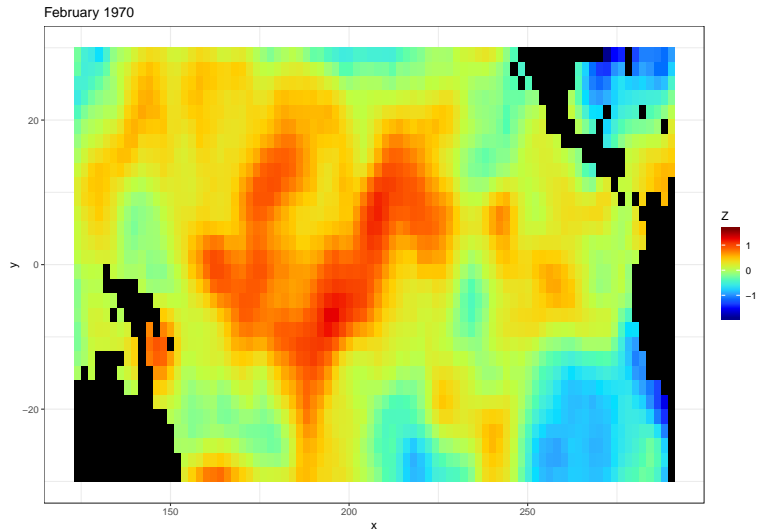
SST

```
## # A tibble: 1,005,480 x 5
##       x     y mask  date      Z
##   <dbl> <dbl> <fct> <date>    <dbl>
## 1   124   -29 land  1970-01-01 -0.363
## 2   126   -29 land  1970-01-01 -0.285
## 3   128   -29 land  1970-01-01 -0.192
## 4   130   -29 land  1970-01-01 -0.157
## 5   132   -29 land  1970-01-01 -0.124
## 6   134   -29 land  1970-01-01 -0.185
## 7   136   -29 land  1970-01-01 -0.353
## 8   138   -29 land  1970-01-01 -0.481
## 9   140   -29 land  1970-01-01 -0.601
## 10  142   -29 land  1970-01-01 -0.658
## # ... with 1,005,470 more rows
```

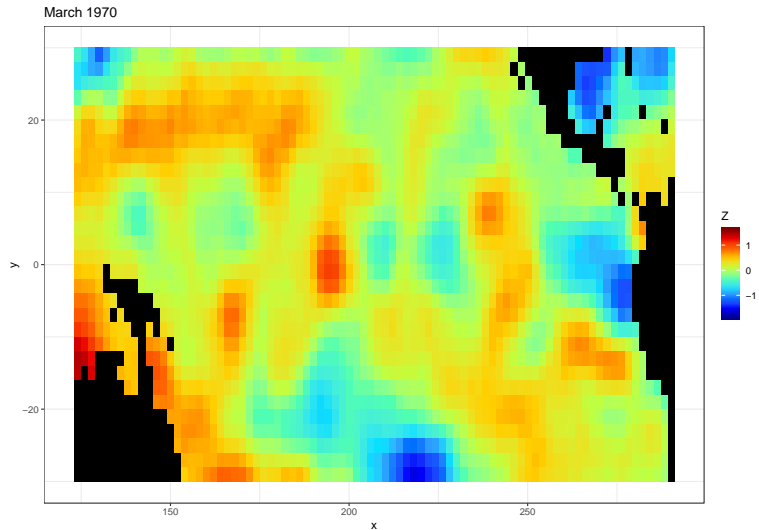
Tropical Pacific Sea Surface Temperature (SST) Anomalies



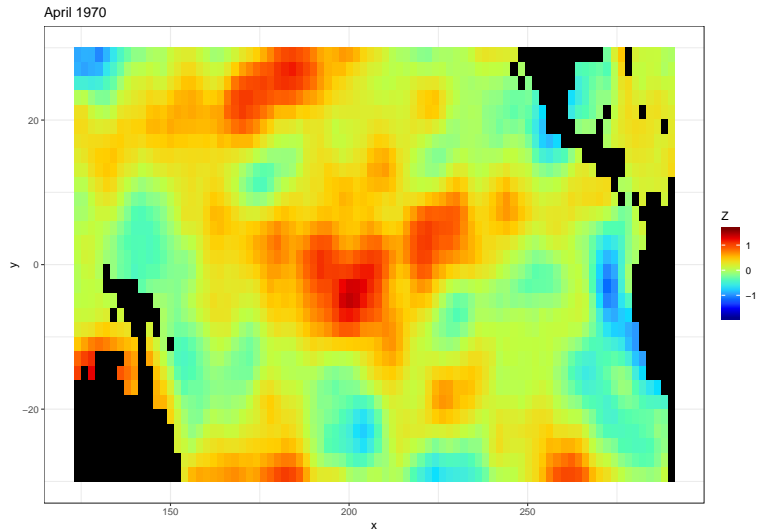
Tropical Pacific Sea Surface Temperature (SST) Anomalies



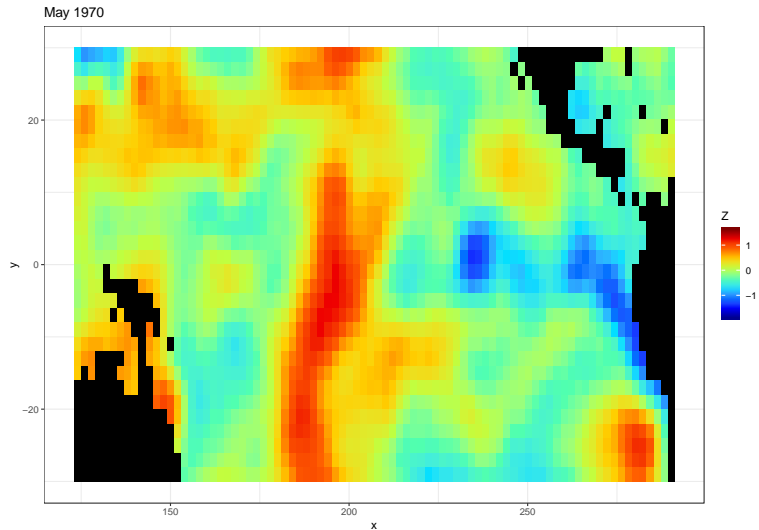
Tropical Pacific Sea Surface Temperature (SST) Anomalies



Tropical Pacific Sea Surface Temperature (SST) Anomalies



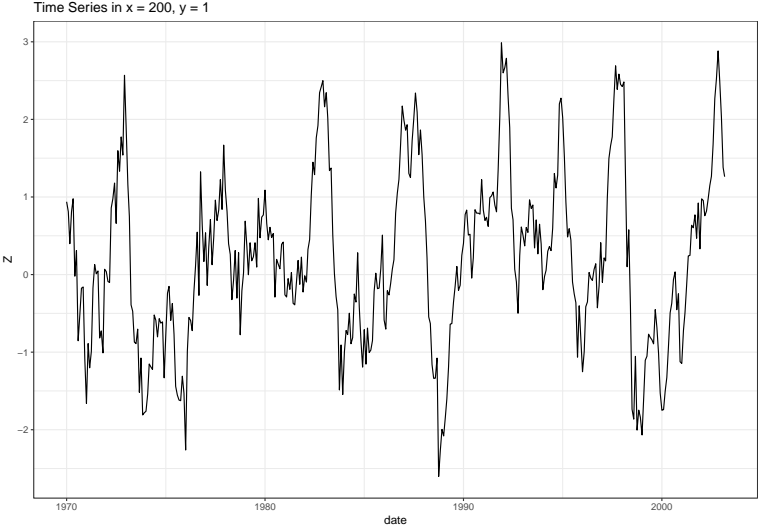
Tropical Pacific Sea Surface Temperature (SST) Anomalies



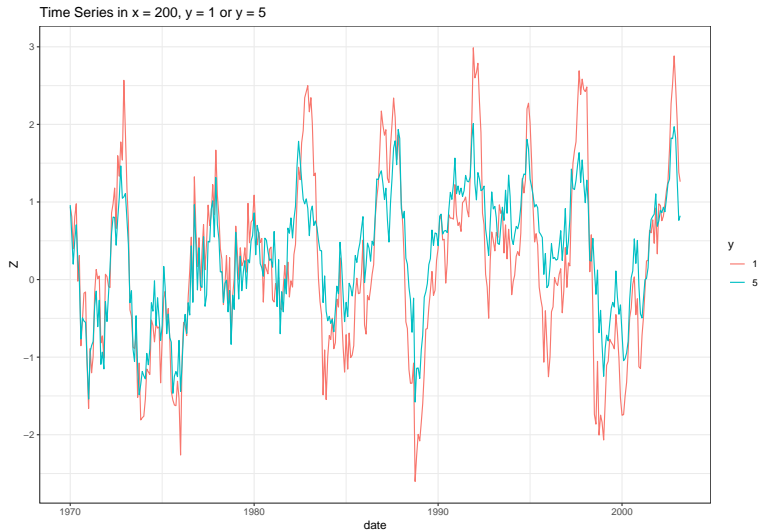
Tropical Pacific Sea Surface Temperature (SST) Anomalies

- ▶ Os valores mínimos e máximos de variação de temperatura no oceano são -4.2414875, 7.0676575, respectivamente. Eu plotei em -1.8573418, 1.6228905 porque são os limites de Janeiro a Maio de 1970.
- ▶ Lembre-se sempre de manter os dados nas mesmas escalas, em múltiplos gráficos (a menos que seja justificável – mas raramente é).

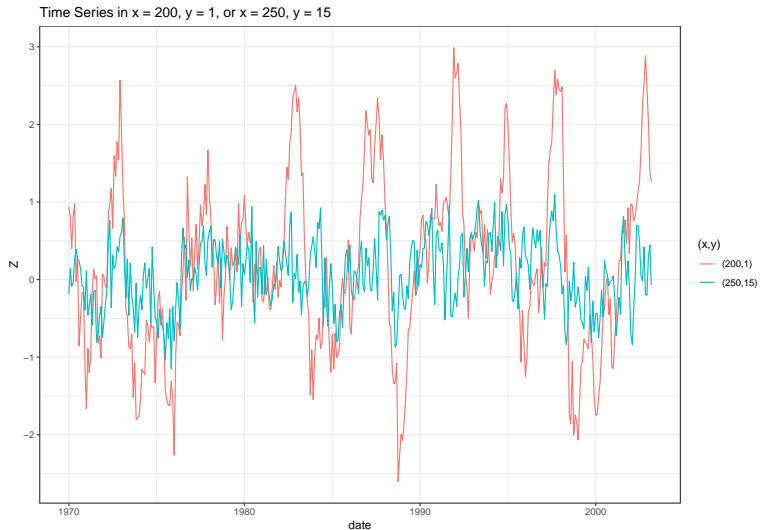
Tropical Pacific Sea Surface Temperature (SST) Anomalies



Tropical Pacific Sea Surface Temperature (SST) Anomalies



Tropical Pacific Sea Surface Temperature (SST) Anomalies



Matrizes de covariância empírica

A matriz $n^2 \times n^2$ de lag- τ covariâncias empíricas é dada por

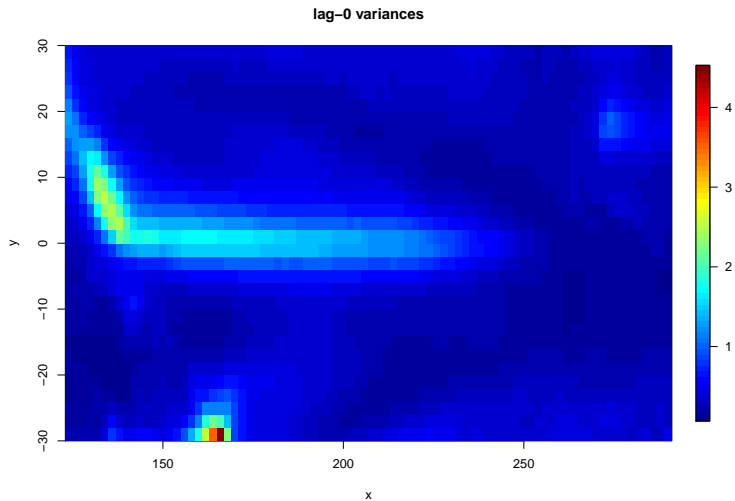
$$\hat{\mathbf{C}}_X^\tau = \frac{1}{t_m - \tau} \sum_{t=\tau+1}^{t_m} (\mathbf{x}_t - \hat{\mu}_X)(\mathbf{x}_{t-\tau} - \hat{\mu}_X)^t$$

Note: só é simétrica para $\tau = 0$!

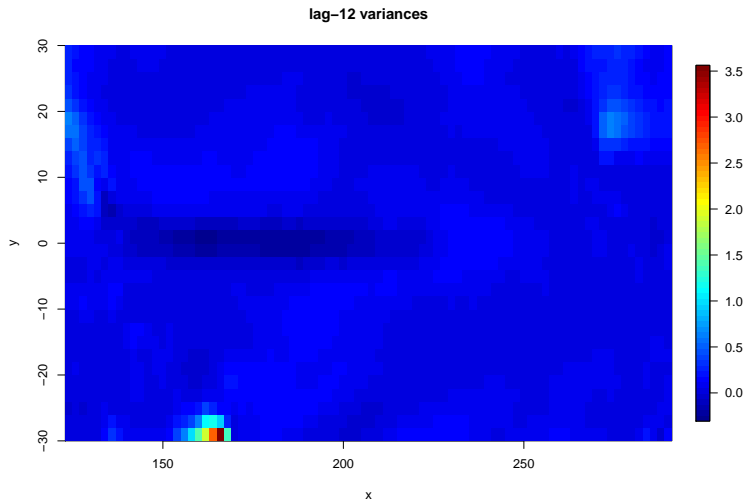
Matrizes de covariância empírica

```
# Exemplo: lag 0 e lag 12 (1 ano)
SSTex <- SST %>% filter(date == as.Date("1970-01-01"))
tau <- seq(as.Date("1970-01-01"), as.Date("2003-03-01"), by = "month")
n <- nrow(SSTex)
C0 <- C12 <- matrix(0, n, n)
mu <- rowMeans(Z)
for(i in seq_along(tau)){
  C0 <- C0 + outer(as.numeric(SST[SST$date == tau[i],]$Z) - mu,
                  as.numeric(SST[SST$date == tau[i],]$Z) - mu)
  if(i > 12){
    C12 <- C12 + outer(as.numeric(SST[SST$date == tau[i],]$Z) - mu,
                      as.numeric(SST[SST$date == tau[i-12],]$Z) - mu)
  }
}
C0 <- C0/length(tau)
C12 <- C12/(length(tau)-12)
```

Matrizes de covariância empírica



Matrizes de covariância empírica



Outras técnicas

- ▶ Análise de componentes principais de $\hat{\mathbf{C}}$, $\hat{\mathbf{C}}^\tau$.
- ▶ Análise de correlações canônicas: entre mapas, buscando correlações temporais

Referências I

Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. Wiley, New York.