Forecasting

- The main criteria to produce forecasts with AR or ARMA models is Mean Square Error (MSE).
- For a stationary AR model, we can write model as and infinite MA process

$$X_t = \Phi(B)^{-1}\omega_t = \sum_{j=0}^{\infty} a_j \omega_{t-j}$$

where a_j (ψ_j book notation) are obtained by inversion of the characteristic polynomial $\Phi(B)$, $a_0 = 1$.

- For $t = n + k$, we have

$$X_{n+k} = \sum_{j=0}^{\infty} a_j \omega_{n+k-j}; \quad k = 1, \ldots, m$$

- The idea is to predict X_{n+k} as a linear combination of past error terms ω_n, ω_{n-1}, ω_{n-2}, ..., ω_1.
- We wish to obtain a forecast $\hat{X}_n(k)$ such that

$$\hat{X}_n(k) = a_k^* \omega_n + a_{k+1}^* \omega_{n-1} + a_{k+2}^* \omega_{n-2} + \ldots$$

  where the a^* constants are determined with this MSE criteria.
- The mean square error of such forecast is

$$E(X_{n+k} - \hat{X}_n(k))^2 = \sigma^2 \sum_{j=0}^{k-1} a_j^2 + \sigma^2 \sum_{j=0}^{\infty} (a_{k+j} - a_{k+j}^*)^2$$

- The MSE reaches its minimum if we set $a_{k+j}^* = a_{k+j}$ for $j = 0, \ldots, \infty$.

- Additionally, we have that $E(\omega_{n+k}|X_n, X_{n-1}, \ldots) = 0$, if $k > 0$ or the observed model residual ω_{n+k} if $k \leq 0$.
- Then using the MA representation of the AR model we get,

$$E(X_{n+k}|X_n, X_{n-1}, \ldots) = a_k\omega_n + a_{k+1}\omega_{n-1} + a_{k+2}\omega_{n-2} + \ldots$$

- The minimum Mean Square Error forecast for X_{n+k} is its conditional expectation given $X_n, X_{n-1}, \ldots, X_1$.
- This conditional expectation, denoted by $f_n(k)$, is a function of k and its know as the forecast function of the process.
- In other words,

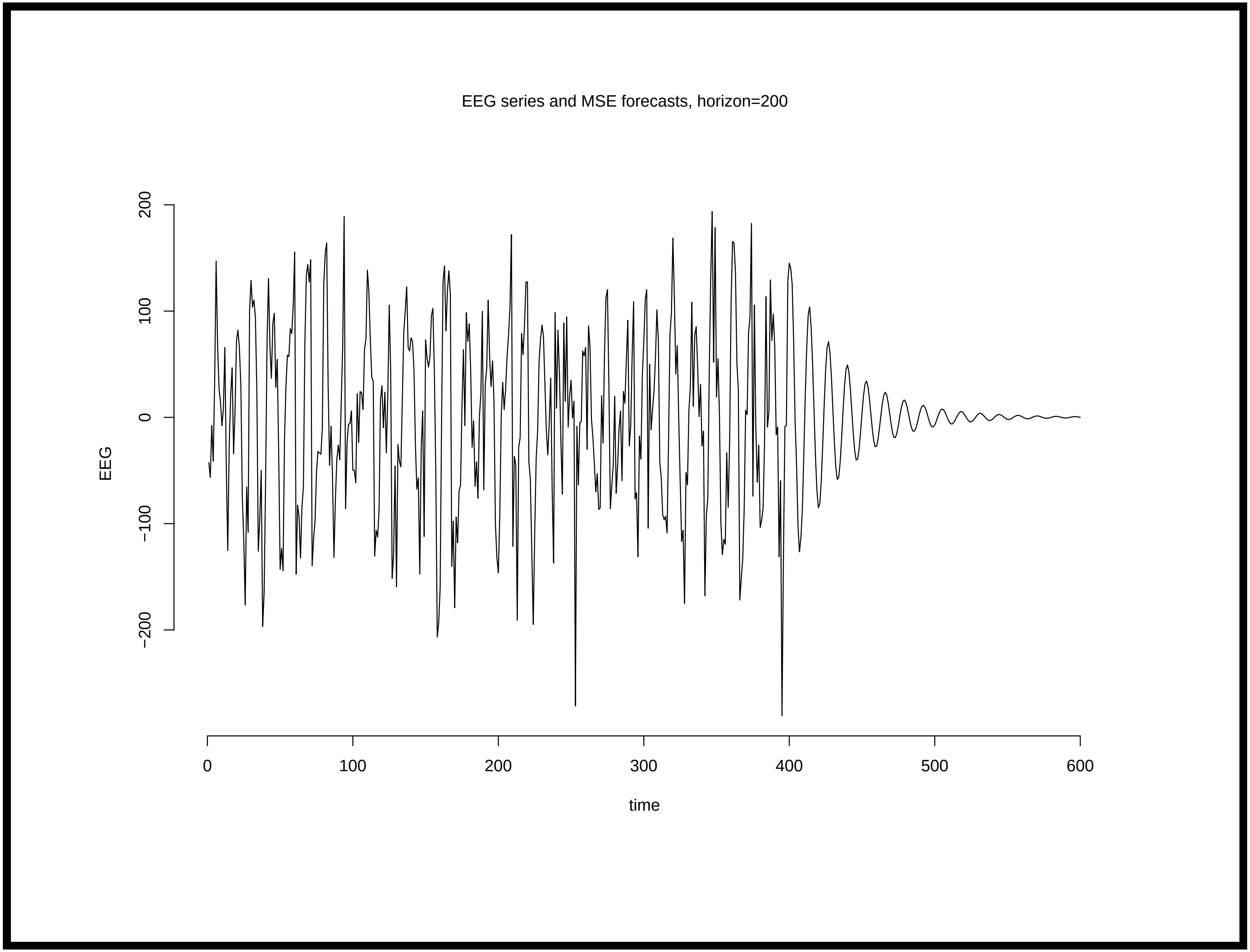
$$f_n(k) \equiv E(X_{n+k}|X_n, X_{n-1}, \ldots) = \hat{X}_n(k)$$

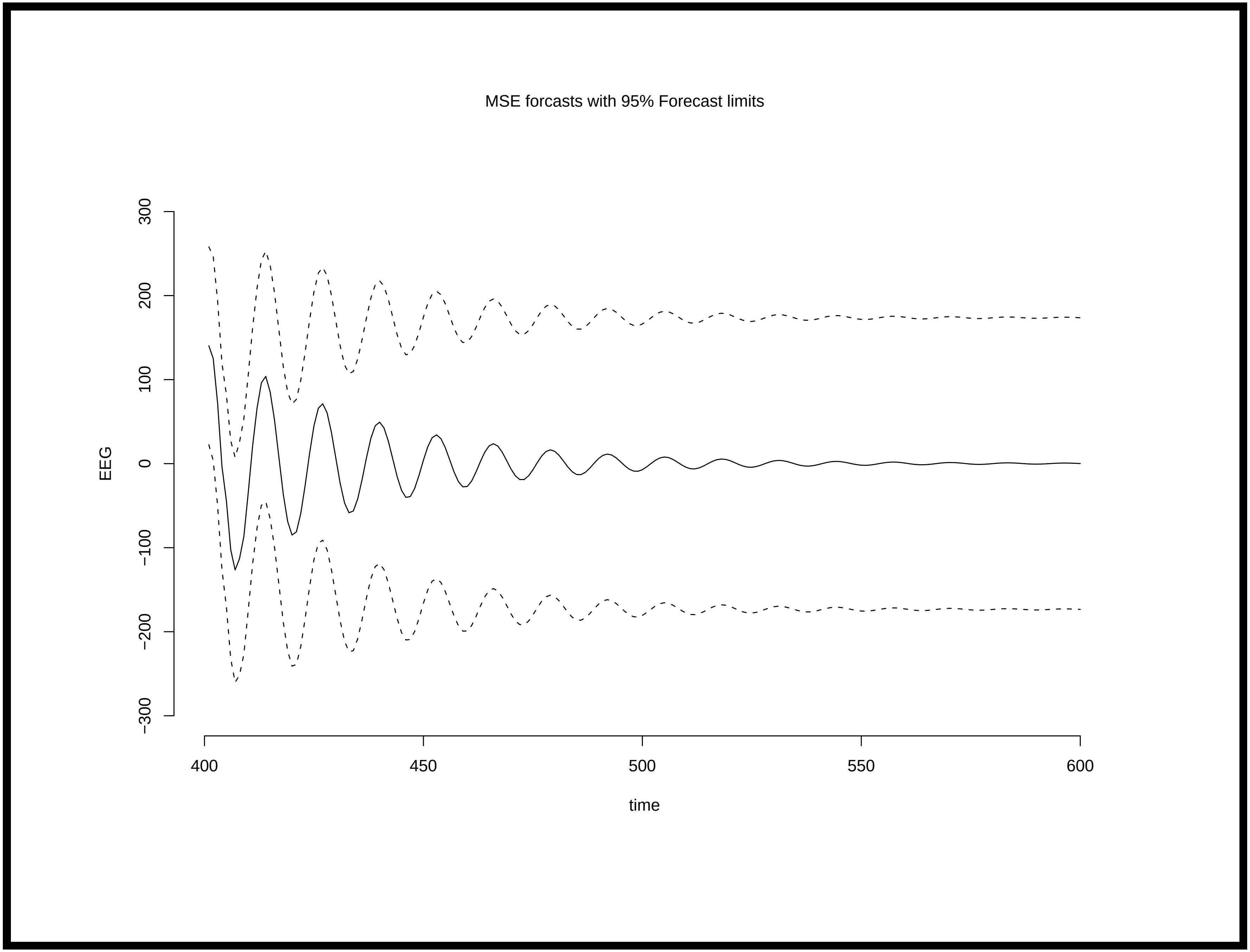
- The forecast error for the minimum MSE forecast is

$$\omega_n(k) = X_{n+k} - \hat{X}_n(k) = \sum_{j=0}^{k-1} a_j \omega_{n+k-j}$$

- Notice that $E(\omega_n(k)) = 0$ or $E(X_{n+k} - \hat{X}_n(k)) = 0$
- The variance of this error is $Var(\omega_n(k)) = \sigma^2 \sum_{j=0}^{k-1} {a_j}^2$
- Under the assumption that the $\omega's$ are Normally distributed, if we wish a 95% predictive interval for X_{n+k} we could use

$$\hat{X}_n(k) \pm 1.96\hat{\sigma}\sqrt{\sum_{j=0}^{k-1} {a_j}^2}$$


EEG series and MSE forecasts, horizon=200
EEG
200
100
0
−100
−200
0
100
200
300
400
500
600
time


MSE forcasts with 95% Forecast limits
EEG
300
200
100
0
−100
−200
−300
400
450
500
550
600
time

```
eeg=scan("eeg")
# Computing MSE forecasts for AR(10)
fr=predict(arima(eeg,order=c(10,0,0)),n.ahead=200)
plot(c(eeg,fr$pred),type='l',xlab="time",
ylab="EEG")
# Plotting mean and 95 % forecast limits.
plot(fr$pred,type='l',xlab="time",
ylab="EEG",ylim=c(-300,300))
lines(fr$pred+1.96*fr$se,lty=2)
lines(fr$pred-1.96*fr$se,lty=2)
```