Channel Estimation

TEQIP-II Sponsored Short Term Training Program on
Wireless Network and Mobile Computing
organized by Computer Engineering Department
Sardar Vallabhbhai National Institute of Technology, Surat
16-20 May 2016

M. A. Zaveri
Computer Engineering Department
Sardar Vallabhbhai National Institute of Technology, Surat
mazaveri@coed.svnit.ac.in

Channel Estimation

- channel state information (CSI) channel properties of a communication link
- describes how a singal propagates from transmitter to receiver
- makes it possible to adapt transmissions to current channel conditions
- crucial for achieving reliable communication with high data rates
- channel estimation classified into three classes:
- training based, blind and semi-blind

Channel Estimation

- channel model as mathematical representation of transfer characteristics of physical medium
- channel esitmation as the process of characterising the effect of physical channel on the input sequence
- receiver to approximate impulse response of the channel
- once model established, its parameter need to be updated
- estimated in order to minimize the error as the channel changes
- minimize MMSE $E[e^2(n)]$
- if receiver has a-priori knowledge of information being sent over the channel
- it can utilize this knowledge to obtain an accurate estiamte of impulse response of the channel
- this method is called training sequence based channel estimation

Channel Estimation

- wasteful of bandwidth - training sequence transmitted for channel estimation
- most systems send information lumped frames, after receipt of frame
- channel estimate can be extracted from the embedded training sequence
- for fast fading channels not adequate - since coherence time of channel might be shorter that the frame time
- blind methods - no training sequence
- utilize underlying mathematical information about the kind of data being transmitted
- bandwidth efficient, slow to converge (more than 1000 symbols may be required for an FIR channel with 10 coefficient),
- computationally intensive - impractical to implement in real time system

Channel Estimation

- training based
  - long training for reliable channel estimate
  - reduces bandwidth efficeincy
- blind methods
  - no training
  - CSI acquired by relying on the received signal statistics
  - using statistical information - solve convergence problem
  - achieves high system throughput
  - high computational complexity
- semi-blind
  - combination of two procedures
  - few training symbols along with blind

Channel Estimation: Pilot based estimation

- pilot based - training based channel estimation
- pilot symbols are used with information symbols in the transmission frame
- pilots are fixed set of symbols which are known at the receiver
- from the received output of pilot symbols, estimation of channel can be performed
- employed for detection of the information symbols transmitted subsequently
- benefit of robust estimate and low computational complexity
- drawback - pilot symbols carry no information - overhead on communication system
- results in wastage of bandwidth - bandwidth inefficient
- least square, minimum mean square error, maximum likelihood, maximum a posteriori can be employed

Channel Estimation: Blind estimation

- does not require pilot symbols - bandwidth efficient
- channel can be estimated using statistical knowledge of received output symbols
- if transmitter employs a symmetric transmit constellation with equal priori probabilities
- then the received symbol stream has a statistical mean of zero
- with knowledge of covariance of the input information symbol
- the computed covariance of output information symbols can be employed to estimate at least part of channel
- computationally complex and having convergence problem
- not attractive where robustness of estimate and computational complexity are critical
- widely known techqunice is subspace method using second order statistics (SOS)

Channel Estimation

- in subspace method, autocorrelation matrix of received signal is decomposed into the signal and noise subspaces
- due to orthogonally of the noise and signal subspace,
- the channel estimates can be calculated based on the noise subsapce
- decomposition of autocorrelation fucntion via eigen value decomposition or singular value decomposition used
- QR decomposition restricts direct matrix inversion and convert full rank channel matrix into simple form - low complexity
- Semi-blind estimation
- combination of both pilot and blind channel estimation
- low complexity with robustness by using limited number of pilot symbol
- and bandwidth efficiency by using statistical blind informtation

Channel Estimation

- quality of channel estimate is enhanced by employing statistical information to aid estimation process or
- minimize the number of pilot symbols transmitted by employing statstical information
- to improve the nature of channel estiammte, increasing bandwidth efficiency
- CSI generated by channel estimation, sent into detection block or fed back to transmitter side to construct beam forming weight vector
- different pilot symbol arrangements:
- estimator with block type pilot (training based)
- estimator with comb type pilot (pilot symbol aided modulation)

Channel Estimation: LS and MMSE

- estimator takes measurement data as inputs and produces estimated values of parameters
$$Y = XH + \eta$$
rewriting
$$Y = Xh + \eta$$
- H and h are unknown vectors, X is known matrix, Y is measurement matrix
- LS channel estimation
- channel estimation $\hat{h}$ for equation $X\hat{h} \approx Y$
- in LS minimization of Euclidean norm squared to the residual $X\hat{h} - Y$

$$\arg_h \min \parallel X\hat{h} - Y \parallel^2$$

Channel Estimation: LS and MMSE

$$\parallel X\hat{h} - Y \parallel^2 = \left(X\hat{h} - Y\right)^H (X\hat{h} - Y)$$
$$= \left(X\hat{h}\right)^H \left(X\hat{h}\right) - Y^H X\hat{h} - \left(X\hat{h}\right)^H Y + Y^H Y$$

- minimum is found at the zero of derivative with respect to $\hat{h}$

$$2X^H X\hat{h} - 2X^H Y = 0 \Rightarrow X^H X\hat{h} = X^H Y$$

$$\hat{h} = \left(X^H X\right)^{-1} X^H Y$$

Channel Estimation: LS and MMSE

- MMSE channel estimation
- optimal result by exploiting statistical dependence between measured data and estimated parameters
- signal source $\rightarrow$ mutlipath channel $\rightarrow$ + noise $\rightarrow$
- receiver filter $\rightarrow$ channel estimator $\rightarrow$ MMSE detector
- minimizing $E[(h - \hat{h}_{MMSE})^2]$

$$\hat{h} = R_{hY} R_{YY}^{-1} Y$$

- R_{hY} and R_{YY} are cross covariance matrices between h and Y and autocovariance matrix of Y respectively

Channel Estimation: LS and MMSE

$$R_{hY} = E[hY^H] = E[h(Xh+\eta)^H] = R_{hh}X^H$$
$$R_{YY} = E[YY^H] = E[(Xh+\eta)(Xh+\eta)^H]$$
$$= E[Xh(Xh)^H + Xh\eta^H + \eta(Xh)^H + \eta\eta^H]$$
$$= XR_{hh}X^H + \sigma_n^2 I$$

- $R_{hh} = E[hh^H]$ is autocovariance matrix of h
- σ_n^2 noise covariance $E[\eta\eta^H]$
- these two quantities are assumed to be known at the estimator
- channel estimate can be written as

$$\hat{h} = R_{hh}X^H(XR_{hh}X^H + \sigma_n^2 I)^{-1}Y$$

Channel Estimation: Maximum a posteriori (MAP)

- requires knowledge of the training sequence, the channel covariance, and the noise covariance at the receiver
- system model described for LS estimation applies to MAP estimation
- maximizes $p(H|Y,X)$ with respect to H
- MAP estimate for H satisfy

$$\frac{\partial \ln(P(H|Y,X))}{\partial H}|H = \hat{H}_{MAP} = 0$$

using Bayes' rule

$$P(H|Y,X) = \frac{p(Y|H,X)p(H,X)}{p(Y|X)}$$

$$\hat{H}_{MAP} = (X^H C_n^{-1} X + C_H)^{-1} X^H C_n^{-1} Y$$

Channel Estimation

- noise covariance $C_n = R_{\eta\eta} = E[\eta\eta^H]$
- channel covariance $C_H = R_{HH} = E[HH^H]$
- for independent Rayleigh fading channels, C_H can be approximated as an identity matrix
- block type pilot - continuous pilot blocks to obtain channel impulse resposne on all sub-carriers
- the length of training block is fixed to the number of sub-carriers in the block
- comb-tye pilot - channel changes even from one block to subsequent one

Channel Estimation using Pilot

- $y = Mh + n$
- channel impluse response $h = [h_0 h_1 \dots h_L]^T$
- within each transmission burst the transmitter sends a unique training sequence which divided into
- a reference length of P and guard period of L bits
- $m = [m_0 m_1 \dots m_{P+L-1}]^T$ bipolar elements $m_i \in \{-1, +1\}$
- circulant training sequence matrix M is formed as

$$M = \begin{bmatrix} m_L & \dots & m_1 & m_0 \\ m_{L+1} & \dots & m_2 & m_1 \\ \ddots & \ddots & \ddots & \ddots \\ m_{L+P-1} & \dots & m_P & m_{P-1} \end{bmatrix}$$

- LS channel estimates $\hat{h} = \arg\min_h \parallel y - Mh \parallel^2$
- assuming white Gaussian noise

$$\hat{h}_{LS} = (M^H M)^{-1} M^H y$$

Channel Estimation

- signal multipath - multi propagation paths, separate phase, attenuation, delay and
- doppler frequency - they add up destructively - called fading
- $y(t) = \sum_{i=1}^{N} \alpha_i s(t - \tau_i(t))$
- N paths arriving at receiver, α and τ attenuation and delay

$$s(t) = \text{real part of } \{\tilde{s}(t)e^{j2\pi f_c t}\}$$

$$\tilde{y}(t) = \sum_{i=1}^{N} \tilde{\alpha}_i \tilde{s}(t - \tau_i(t))$$

- f_c carrier frequency, $\tilde{\alpha}_i = \alpha_i e^{j2\pi f_c t}$ time varying complex attenuation of each path
- time varying discrete multipath channel by time varying complex impulse response

$$\tilde{h}(\tau; t) = \sum_{i=1}^{N} \tilde{\alpha}_i \delta(t - \tau_i(t))$$

Channel Estimation

- modelling channel tap gain as an auto regressive process
- complex gaussian random process can be represented by a general auto regressive model
- any stationay random process can be represented as an infinite tap AR process
- infinite tap AR process model is impractical, truncated to N-tap form
- AR process represented by a difference equation

$$S(n) = \sum_{i=1}^{N} \phi_i S(n - i) + w(n)$$

- $S(n)$ complex gaussian process, ϕ_i parameters of the model
- N number of delays in the autoregressive model
- $w(n)$ sequence of identically distributed zero-mean complex gaussian random variables

Channel Estimation

- $E\{w(n)\} = 0$

$$E\{w(n)w(j)\} = \left\{ \begin{array}{cc} s_n^2 & \text{for } n = j \\ 0 & n \neq j \end{array} \right\}$$

$$f_{w(n)}(\lambda) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{\frac{-\lambda^2}{2\sigma_n^2}}$$

- sequence $w(n)$ white gaussian noise because its spectrum is broad and uniform over an infinite frequency range
- AR process is another name for a linear difference equation model driven by gaussian noise
- Nth order difference equation can be reduced to a state model in the vector form

$$\bar{S}(n) = F\bar{S}(n-1) + \bar{W}(n)$$

- $\bar{S}$ and $\bar{W}(n)$ column vectors of size $N \times 1$ and F is $N \times N$ matrix

Channel Estimation

- mean

$$\mu_s = E[S(n)] = E\left[\sum_{i=1}^{N} \phi_i S(n-i) + w(n)\right] = 0$$

- variance

$$\sigma_S^2 = E\{S(n)S(n)\} = E\left\{S(n)\left(\sum_{i=1}^{N} \phi_i S(n-i) + w(n)\right)\right\}$$

$$= \sum_{i=1}^{N} \phi_i R_{SS}(i) + \sigma_n^2$$

- autocorrelation

$$R_{SS}(m) = E\{S(n-m)S(n)\} = E\left\{\left[\sum_{i=1}^{N} \phi_i S(n-i) + w(n)\right] S(n-m)\right\}$$

$$= \sum_{i=1}^{N} \phi_i R_{SS}(m-i)$$

Channel Estimation

- autocorrelation coefficeint

$$r_{SS}(m) = \frac{R_{SS}(m)}{s_X^2} = \sum_{i=1}^{N} \phi_i r_{SS}(m-i)$$

$m \geq 1$

- N th order difference equation can be solved for desired AR coefficients

$$\begin{bmatrix} 1 & r_{SS}(1) & \dots & r_{SS}(N-1) \\ r_{SS}(1) & 1 & \dots & r_{SS}(N-2) \\ \ddots & \ddots & \dots & \ddots \\ r_{SS}(N-1) & \dots & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} = \begin{bmatrix} r_{SS}(1) \\ r_{SS}(2) \\ \vdots \\ r_{SS}(N) \end{bmatrix}$$

matrix equation is known as Yule-Walker equation

- $\bar{R}\phi = \bar{r}_{SS}$ $\phi = \bar{R}^{-1}\bar{r}_{SS}$
- matrix of AR coefficeints that models the complex gaussian process
- given autocorrelation of the process using Yule-Walker equation calculate AR coeffcients

Channel Estimation

- data based channel estimator uses training sequences sent over the channel to estimate the impulse response of the channel
- channel estimation using correlation method
- say, training seuqnece of length M known to receiver is sent over the channel
- assumed that the channel does not change over the span of data sent

$$\bar{x} = [x_0 x_1 \dots x_{M-1}]^T$$

- this bit sequence mapped to unit energy symbols bit 0 $\rightarrow +1$ symobl and bit 1 $\rightarrow$ -1 symbol to simulate BPSK modulation
- channel impulse response when the training sequence is sent over the channel

$$\bar{h} = [\tilde{h}_0 \tilde{h}_1 \tilde{h}_2 \dots \tilde{h}_{L-1}]^T$$

- L is channel impulse response length or the number of processes to be tracked

Channel Estimation

- signal received is the convolution sum of the signal sent and the impulse response received

$$\bar{y} = \bar{x} * \bar{h} + \bar{n}_c$$

$$y(n) = \sum_{m=0}^{L-1} h(m)x(n-m) + n_c$$

matrix form

$$\bar{y} = \begin{bmatrix} x_0 & 0 & . & 0 \\ x_1 & x_0 & . & . \\ . & x_1 & . & 0 \\ x_{M-1} & . & . & x_1 \\ 0 & x_{M-1} & . & . \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & . & x_{M-1} \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{L-1} \end{bmatrix} + \begin{bmatrix} n_{c_0} \\ n_{c_1} \\ \vdots \\ n_{c_{L+M-1}} \end{bmatrix}$$

$$\bar{Y} = \bar{X}\bar{h} + \bar{n}_c$$

Channel Estimation

- $\bar{X}$ Toeplitz matrix containing delayed versions of the training sequence sent
- gaussian channel noise variance σ_c^2 of $\bar{n}_c$
- $E_b = 1$ SNR of channel given by $\frac{E_b}{N_0} = \frac{1}{2\sigma_c^2}$
- following general linear regression method, the estimate of channel is given by

$$\hat{\bar{h}} = (\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{Y})$$

$$\hat{\bar{h}} = (\bar{X}^T\bar{X})^{-1}(\bar{X}^T(\bar{X}\bar{h} + \bar{n}_c))$$

$$= (\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{X})\bar{h} + (\bar{X}^T\bar{X}^{-1}(\bar{X}^T\bar{n}_c)$$

$$= \bar{h} + (\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)$$

- error $\tilde{\bar{h}} = \hat{\bar{h}} - \bar{h} = (\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)$

Channel Estimation

- expection of estimation error

$$E\left[\tilde{\bar{h}}\right] = E\left[(\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)\right] = (\bar{X}^T\bar{X})^{-1}(\bar{X}^T E[\bar{n}_c])$$

- channel noise is zero mean $E\left[\tilde{\bar{h}}\right] = 0$
- estimator is unbiased
- error covariance

$$P_D = E\left[\tilde{\bar{h}}\left(\tilde{\bar{h}}\right)^H\right]$$

$$= E\left\{\left[(\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)\right]\left[(\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)\right]^H\right\}$$

$$P_D = E\left\{\left[(\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)\right]\left[(\bar{X}^T\bar{n}_c)^H((\bar{X}^T\bar{X})^{-1})^H\right]\right\}$$

$$= E\left\{(\bar{X}^T\bar{X})^{-1}(\bar{X}^T\bar{n}_c)(\bar{n}_c^H\bar{X})(\bar{X}^T\bar{X})^{-1}\right\}$$

$$= (\bar{X}^T\bar{X})^{-1}\bar{X}^T E\left(\bar{n}_c\bar{n}_c^H\right)\bar{X}(\bar{X}^T\bar{X})^{-1}$$

Channel Estimation

$$= (\bar{X}^T\bar{X})^{-1}\bar{X}^T\{\sigma_c^2 I\}\bar{X}(\bar{X}^T\bar{X})^{-1}$$

$$= \sigma_c^2\left[(\bar{X}^T\bar{X})^{-1}\left(\bar{X}^T\bar{X}\right)(\bar{X}^T\bar{X})^{-1}\right]$$

$$= \sigma_c^2(\bar{X}^T\bar{X})^{-1}$$

H is hermetian transpose of a matrix defined as complex conjugate of standard transpose

$(\bar{X}^T\bar{X}) =$

$$\begin{bmatrix} x_0 & x_1 & . & . & x_{M-1} & 0 & . & 0 \\ 0 & x_0 & x_1 & . & . & x_{M-1} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & . & 0 & x_0 & x_1 & . & . & x_{M-1} \end{bmatrix} \begin{bmatrix} x_0 & 0 & . & 0 \\ x_1 & x_0 & . & . \\ . & x_1 & . & 0 \\ x_{M-1} & . & . & x_1 \\ 0 & x_{M-1} & . & . \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & . & x_{M-1} \end{bmatrix}$$

Channel Estimation

$$= \begin{bmatrix} \sum_{i=0}^{M-1} x_i^2 & \sum_{i=0}^{M-1} x_i x_{i-1} & . & \sum_{i=0}^{M-1} x_i x_{i-L+1} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=0}^{M-1} x_i x_{i-L+1} & . & \sum_{i=0}^{M-1} x_i x_{i-1} & \sum_{i=0}^{M-1} x_i^2 \end{bmatrix}$$

$x_i = \pm 1$, $\sum_{i=0}^{M-1} x_i^2 = M$ and $(\bar{X}^T\bar{X}) =$

$$\begin{bmatrix} M & \sum_{i=0}^{M-1} x_i x_{i-1} & . & \sum_{i=0}^{M-1} x_i x_{i-L+1} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=0}^{M-1} x_i x_{i-L+1} & . & \sum_{i=0}^{M-1} x_i x_{i-1} & M \end{bmatrix}$$

$$= M\begin{bmatrix} 1 & \sum_{i=0}^{M-1} x_i x_{i-1} & . & \sum_{i=0}^{M-1} x_i x_{i-L+1} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=0}^{M-1} x_i x_{i-L+1} & . & \sum_{i=0}^{M-1} x_i x_{i-1} & 1 \end{bmatrix}$$

$$r_{xx}(k) = \frac{1}{R_{xx}(0)}\sum_{i=0}^{M-1} x_i x_{i-k} = \frac{1}{M}\sum_{i=0}^{M-1} x_i x_{i-k}$$

Channel Estimation

- $(\bar{X}^T\bar{X})$ $L \times L$ matrix containing delayed versions of training sequence autocorrelation

$$(\bar{X}^T\bar{X}) = M\begin{bmatrix} 1 & r_{xx}(1) & . & r_{xx}(L-1) \\ \vdots & \vdots & \vdots & \vdots \\ r_{xx}(L-1) & . & r_{xx}(1) & 1 \end{bmatrix}$$

- $r_{xx}(\tau)$ normazized training sequence autocorrelation
- for ideal auto correlation $P_D = \frac{\sigma_c^2}{M}[I]$
- for a single process estimate $L = 1$ error covariance is $P_D = \frac{\sigma_c^2}{M}$
- inverse relationship between length of training sequence and the covariance of the data estimate
- data estimate worsens as noise in the channel increases

Thank You