**C01: NONMEM Estimation 1 (FO)**

배균섭

FO는 first order approximation의 약어이다. NONMEM에서 FO method란 $ESIMATION 에서 METHOD=ZERO 또는 METHOD=0 이라고 옵션을 주거나 METHOD에 아무런 옵션을 주지 않는 경우에 default로 실행되는 추정 방법이다.

NONMEM의 FO method는 기본적으로 정규분포시의 Maximal Likelihood Estimation (MLE)와 같다.

NONMEM의 목적함수 값(Objective Function Value, OFV)은 개인별 OFV (Oi)의 단순합()이며, FO method의 개인별 목적함수(objective function)은 다음과 같다(Users Guide I p34-41).



단, 

: Weighted Residual (WRES)

여기에서

: Partial derivatives of  with respect to etas ( x )

: Partial derivatives of  with respect to epsilons ( x )

: Number of observations for i-th subject

: Number of etas

: Number of epsilons

: Prediction function, function of theta, eta, covariates

 or : function of F and epsilon

: Fixed effect, constant

: Interindividual random variability. 

: Residual variability. 

: Covariates

처음의 개인별 목적함수는 다음과 같이 써도 동일하다.



즉, 이 식은 MLE 이론에 나오는 -2LL에서 부분만 제외한 것과 같다. (MLE 이론에 대해서는 부록 참고) 따라서, 이제 여기에서 어떻게 , 가 되는 지만 이해하면 될 것이다. 하지만, 그전에 matrix의 -1/2승과, 다변량 함수의 Taylor series 전개를 이해해야 한다.

**Matrix의 -1/2승,** 

Variance-covariance matrix와 같은 symmetric positive definite (SPD) matrix의 -1/2 승은 다음과 같은 함수로 구할 수 있다.  not 

|  |  |
| --- | --- |
| |  | | --- | | SqrtInvCov <- function(M)  {  EigenResult <- eigen(as.matrix(M))  EigenVector <- EigenResult$vectors  EigenValues <- abs(EigenResult$values)  return(EigenVector %\*% diag(1/sqrt(EigenValues)) %\*% t(EigenVector))  } | |
| Fig. C01-1. R function to calculate square root inverse of a variance-covariance matrix |

NONMEM 목적함수의 Ri를 보면 이 residual (RES) 앞에 곱해져 있고, NONMEM에서는 이것을 weighted residual (WRES)로 정의하였는데, 행렬을 곱해줌으로 인해, residual(RES)과 weighted residual(WRES)의 부호가 다를 수 있다. (NONMEM OUTPUT 예를 참고)

**문제의 정의**

수치화 할 수 있는 현상(결과)을 설명하는 모형(확률적이면서 수학적인 함수)을 구성하는데 있어서, 측정과 관련된 오차가 없는 경우의 종속변수를 F라하고, 관측값을 나타내기 위한 확률변수를 Y라 하자.

예를 들어, 흔히 사용되는 1구획 정맥투여 약동학 모형에서 인구집단에서 전형적으로 관찰되는 CL를 1이라 하고, V를 2라 하면 개인들의 약동학 파라미터는 다음과 같이 분해해서 쓸 수 있다.

CL = 1 + 1 or CL=1 \* EXP(1)

V = 2 + 2 or V = 2 \* EXP(2)

여기에서 는 상수(constant), 는 확률변수(random variable)이다. 는 MVN(multi-variate normal) 분포를 따르며 그 평균은 0 vector, 분산은 행렬을 따른다고 가정한다.

투약용량(dosing amount), 채혈시간(sampling time), 체중(body weight) 등을 covariate(여기에서는 라 쓰기로 한다)라 하며, 이것이 혈중농도(F)에 영향을 미치는 경우에 우리는 다음과 같은 함수로 나타낼 수 있다.



 or 

여기에서 는 inter-individual random variability를 나타내는 확률변수이고, 는 intra-individual (or residual) random variability를 나타내는 확률변수이다. 는 인구전체 집단에서는 확률변수이지만, 개인내에서 보면 ‘어느 정도’ 고정된 상수(constant)이며 이의 추정치 를 Empirical Bayes Estimate (EBE), Maximum A posteriori (MAP), realized eta, posthoc eta 라고도 한다.

MLE에서 한 개인에서의 -2LL은 다음과 같다.



여기에서 첫 번째 항은 과는 무관함으로 이를 제외한 것을 목적함수(objective function, OBJ)로 하면 다음과 같이 된다.



Total Objective Function Value (OFV)는 개인별 OFV(Oi)의 단순합과 같으며 개인별 OFV를 보는 NONMEM 코딩은 다음과 같다.

|  |  |
| --- | --- |
| |  | | --- | | $INFN  IF (ICALL.EQ.3) THEN  WRITE (55,\*) IIDX, CNTID  CLOSE (55)  ENDIF | |
| Fig. C01- 2. 개인별 OFV를 fort.55 파일에 저장해주는 NONMEM control code |

위와 같이 코딩하는 경우에 fort.55 파일에 원하는 내용이 저장되게 된다. 여기에서 주의할 점은 IIDX는 피험자 ID가 아니라, 피험자 index로써 dataset에서 몇 번째 피험자인지를 나타낸다.

한 개인에서의 관측치 vector Y를 다음과 같은 함수라 할 때, 그 개인의 eta true value (t)와 epsilon=0 근처에서 Taylor series expansion하는 경우 다음과 같이 표현할 수 있다.



: fixed effect, constant

: interindividual random variability. , 

: residual variability., 



 





 if 

여기에서 정규분포를 따르는 두 확률변수(eta and epsilon)의 합(Y)은, 또 다른 정규분포를 따르는 새로운 확률변수가 되므로, MLE이론에 그대로 넣을 수가 있다.

이제 E(Y)와 V(Y)를 OBJi의 와 에 대입한다. 이때 t를 0으로 고정하면 FO 방법이 된다.

FO 방법(METHOD=ZERO)에서는 eta를 0으로 고정한 채로 최적화(theta, omega, sigma를 구함)하므로



V(Y)를 계산하는 방법은 INTERACTION 옵션에 따라 달라진다. 좀 더 자세히는 H를 계산하는 방법이 달라진다.

i) INTERACTION option을 사용하는 경우



ii) INTERACTION option이 없는 경우 (default)



위에서 보듯이 INTERACTION 옵션은 FO의 경우에는 분산의 계산에만 영향을 미친다.

FO method에서는 최종 theta, omega, sigma와 최종 목적함수값을 산출하기까지는 eta를 항상 0으로 고정시키며, 개인별로 현실화된 eta (realized eta, posthoc eta, EBE, MAP)는 $ESTIMATION에서 POSTHOC이라는 option을 주어야만 구할 수 있다.

POSTHOC option에 의한 EBE의 추정은 다음과 같은 별도의 objective function을 이용한다.





(여기에서의 log는 모두 자연로그이다.)

상기 목적함수에서 eta를 구하려면, theta, omega, sigma를 알고 있어야 하는데, 이는 이전에 eta를 0으로 fix한 채로 구해 놓은 theta, omega, sigma의 추정치를 사용한다.

|  |
| --- |
| $PROB THEOPHYLLINE ORAL  $INPUT ID AMT=DROP TIME DV BWT  $DATA THEO  $PRED  DOSE = 320  V = THETA(1) \* EXP(ETA(1))  K = THETA(2) \* EXP(ETA(2))  KA = THETA(3) \* EXP(ETA(3))  F = DOSE / V \* KA / (KA - K) \* (EXP(-K\*TIME) - EXP(-KA\*TIME))  Y = F + F \* EPS(1) + EPS(2)  IF (ICALL.EQ.3) THEN  WRITE (53,\*) OMEGA(BLOCK)  WRITE (54,\*) SIGMA(BLOCK)  WRITE (55,\*) IIDX, CNTID  ENDIF  $THETA (0, 50, 500) (0, 1, 10) (0, 5, 50)  $OMEGA BLOCK(3)  0.2  0.1 0.2  0.1 0.1 0.2  $SIGMA 0.1 0.1  $EST MAX=9999 METHOD=ZERO  $TAB ID TIME **G11 G21 G31 H11 H21** FILE=sdtab NOPRINT ONEHEADER |
| Fig. C01-3. NONMEM에 포함된 THEO 자료를 대상으로 한 기본 control file example |

$EST부분에 INTER를 포함시킨 경우와 아닌 경우에 C01-4의 R script가 어떻게 다른 결과를 내는지 비교해 보는 것이 좋다.

|  |
| --- |
| source("SqrtInvCov.R") #Fig. C01-1의 R function file  Data <- as.matrix(read.table("sdtab", skip=1, header=T))  OM <- as.matrix(read.table(“fort.53"))  SG <- as.matrix(read.table(“fort.54"))  # 윗 부분까지는 자신의 환경에 맞게 조정해 주어서, C01-3의 control file을 수행 결과인 sdtab, fort.53, fort.54을 읽을 수 있어야 한다.  IDs <- unique(Data[,"ID"])  nID <- length(IDs)  Oi <- cbind(IDs, matrix(nrow=nID, ncol=3))  for(i in 1:nID) {  Gi <- Data[Data[,"ID"]==IDs[i], c("G11","G21","G31")] # sdtab 내에 저장된 G 값들  Hi <- Data[Data[,"ID"]==IDs[i], c("H11","H21")] # sdtab내에 저장된 H 값들  Ci <- Gi %\*% OM %\*% t(Gi) + diag(diag(Hi %\*% SG %\*% t(Hi))) # G\*Omega\*G + H\*Sigma\*H  # Ci 개개 관측값에서의 분산  Ri <- Data[Data[,"ID"]==IDs[i], "WRES"]  Oi[i,2] <- log(det(Ci)) + t(Ri) %\*% Ri # sdtab의 WRES를 이용    Yi <- Data[Data[,"ID"]==IDs[i], "DV"]  Fi <- Data[Data[,"ID"]==IDs[i], "PRED"]  Oi[i,3] <- log(det(Ci)) + t(Yi - Fi) %\*% solve(Ci) %\*% (Yi - Fi) # MLE 공식을 이용  R0i <- SqrtInvCov(Ci) %\*% (Yi - Fi) # matrix -1/2승으로 WRES 계산  Oi[i,4] <- log(det(Ci)) + t(R0i) %\*% R0i # 계산된 WRES를 이용  }  Oi # 세 가지 방식으로 계산한 개인별 OFV(Oi) 들  sum(Oi[,2]) # sdtab의 WRES를 이용한 경우의 total OFV  sum(Oi[,3]) # MLE이론의 공식을 이용한 경우의 total OFV  sum(Oi[,4]) # Ci matrix의 -1/2승을 이용하여 WRES를 계산한 경우의 total OFV |
| Fig. C01-4. Fig. C01- 3의 control file 결과를 이용하여 R로 OFV를 계산하는 예 |

부록1 기본 통계 용어와 MLE 설명

*기본 용어 정의*

**확률(probability):** 절대 일어나지 않으면 0, 항상 일어나면 1로 표현하고, 위 두 가지가 아니면 0과 1사이의 값으로 일어날 가능성을 표현할 수 있고, 결과는 (적어도 관찰자가 보기에) 우연에 의해 지배되는 그 무엇. 독립적으로 무수히 시행한다면 (해당사건이 나오는 횟수)/(전체 시도횟수)로 정의할 수 있다.

**확률변수(random variable):** 우연에 의해 (확률적으로) 결정되는 사건(event)을 수(number)로 대응시킨 사상(mapping, 넓은 의미의 함수 function)

확률변수의 보통 함수(ordinary function of random variable)는 새로운 확률변수가 된다. 확률변수에는 연속형(continuous)과 이산형(discrete)이 있다. 확률변수는 주로 영어 대문자(X, Y)로 나타내고, 많이 쓰이는 일부에 대해서는 그리스 소문자(, )로 나타낸다. 그리스 대문자  등과 소문자 등은 상수를 나타낸다. 영어 소문자(x, y)는 해당 대문자 확률변수(X, Y)의 관측치(관측값, observation value)를 나타낸다. 표본평균  도 확률변수이다.

**확률밀도함수(probability density function, pdf):** 연속형 확률변수 X에 대해 다음과 같은 조건을 만족하는 함수 f(x)를 확률변수 X의 확률밀도함수(pdf)라 한다.



이산형의 경우에는 적분기호대신 summation 기호()를 사용하며, 확률질량함수(probability mass function, pmf)이라 한다.

다음 조건을 만족하는 함수 F(x)를 cumulative distribution function (cdf)라 한다.



평균이 , 분산이 인 정규분포(normal distribution)를 따르는 확률변수 X의 pdf는 다음과 같다.



이를 다음과 같이도 표현한다.



**우도(Likelihood):** 확률밀도함수가 f(y)라고 생각되는 확률변수 Y에 대해 n개의 관측값(observation value) (i=1,…,n)가 있을 때, likelihood는 다음과 같이 정의한다.



그런데, f(y)의 모양(shape)을 결정하는 parameter인 을 나타내주기 위해 다음과 같이 쓰기도 한다.

 or 

여기에서 vector인 는 **y** (볼드체)로 표현할 수도 있으나, 혼동을 줄이기 위해 위에 화살표를 표시했다. 여기에서 y는 확률변수가 아닌 관측치이므로 소문자로 표현한다. 연습으로 정규분포를 따르는 확률변수 Y의 관측치들이 yi일 때 이의 likelihood 식을 작성해 보세요.



확률은 항상 0에서 1까지의 값(실수)을 갖지만 likelihood는 이론적으로 0보다 크거나 같고 1보다 커져서 양의 무한대()에 가까운 값을 가질 수도 있다. 실제로는 대부분 0에 매우 가까워서 -2 log likelihood (-2LL) 값을 계산에 사용하게 된다. 관측을 나타내는 확률변수를 Y라하고 i번째 피험자의 j번째 관측치를 yij라 하고, 그 시점에서의 평균과 분산을 각각 와 라 했을 때 -2 log likelihood는 다음과 같이 쓸 수 있다.



를 E(Yij), 를 V(Yij)라고도 쓸 수 있으며, 많은 경우에 E(Yij)는 몇(j) 번째이냐에 따라 달라지지만, V(Yij)는 몇 번째와 상관없이 동일하게 어떤 상수()로 가정(homoscedastic)하는 경우도 많다. 이런 점에 유의해서 식을 봐야 한다. 만약에 를 로 하고, 를 으로 단순화시켜 가정하고 이 likelihood를 최대화하는 문제를 푸는 경우(와 에 대해 편미분하여 기울기가 0이 되는 값을 찾는다.)에는 다음과 같은 해를 얻을 수 있고 이를 최대우도 추정량(Maximum Likelihood Estimator, MLE)이라 한다..

, 

여기에서 특이한 것은 모분산(population variance)의 추청량(estimator)으로 흔히 사용되는 표본분산 S2과 분모가 약간 다르다는 점이다.



의 추정량으로 흔히 사용되는 표본분산 S2에 비해 MLE는 편의된(biased) 추정량이긴 하나 일치추정량(consistent estimator)이면서 효율적(efficient)이고, 기존 MLE를 다른 함수에 대입해도 일치성이 유지되는 등, 여러 가지 좋은 특성 때문에 계속 사용되고 있다. 특히 MLE의 분산을  또는 (여기서 는 또는)로 사용할 수 있다는 점이 특이하며, 여기에서 주의할 점은 -2LL이 아닌 그냥 log likelihood이므로 흔히 사용하는 목적함수(objective function)로부터 계산할 때는 이를 보정해 주어야 한다.

불편성(unbiasedness)

E(A) = 

일치성(Consistency)



최소분산성(Minimum Variance)/효율성(Efficiency)

Var(A) < Var(B)