변분 베이지안 방법은 베이지안 추론과 기계 학습에서 발생하는 다루기 힘든 적분을 근사화하는 기술 계열입니다. 일반적으로 “관측 변수” (일반적으로 "데이터"라고 함)와 “알 수 없는 매개 변수” 및 “잠재 변수”로 구성된 복잡한 통계 모델에 사용되며 그래픽 모델에서 설명하는 것처럼 3 가지 유형의 랜덤 변수 간의 다양한 종류의 관계가 있습니다.

베이지안 추론에서 매개 변수와 잠재 변수는 "관찰되지 않은 변수"와 함께 그룹화됩니다. 변분 베이지안 방법은 주로 두 가지 목적으로 사용됩니다.

1.관찰되지 않은 변수의 사후 확률에 대한 분석적인 근사를 제공하여 이러한 변수에 대한 통계적 추론을 수행합니다.

2.관측 된 데이터.( 즉 관찰되지 않은 변수에 대해 수행 된 marginalization으로 주어진 데이터의 marginal 확률)의 marginal likelihood(때로는 "증거"라고 부름)에 대한 하한을 유도합니다.

이것은 전형적으로 모델을 선택하는데 사용된다

일반적으로 주어진 모델에 대한 더 높은 marginal likelihood은 그 모델에 의한 데이터의 적합(fit)이 더 높다는 것을 의미하고, 문제에서 해당 모델이 데이터를 생성 한 모델 일 확률이 높다는 것이 일반적이다. (베이즈 (Bayes) 기사 참조).

이전의 목적 (사후 확률을 근사화하는 것)에서, 변분 베이즈는 몬테카를로 샘플링 방법에 대한 대안이다. 특히 깁스 샘플링과 같은 마르코프 체인 몬테카를로 방법은 복잡한 분포에 대한 통계적 추론에 베이지안 방식을 완전히 사용하기위한 것이다.

복잡한 분포는 직접 평가하거나 샘플링하기가 어렵습니다. 특히, Monte Carlo 기법은 샘플 세트를 사용하여 정확한 사후(posterior)에 수치 근사를 제공하는 반면, 변분 베이즈는 사후의 근사에 대해 국부적으로 최적의 정확한 분석 솔루션을 제공합니다.

변분 베이즈는 각 매개 변수의 가장 가능성이 높은 값의 최대 사후 추정 (MAP 추정)로부터 매개변수와 잠재 변수의 전체 사후 분포를 계산(근사화)하는 완전 베이지 추정까지의 EM (expectation-maximization) 알고리즘의 확장으로 볼 수 있습니다

EM 에서 처럼 최적의 매개 변수 값 집합을 찾고 분석적으로 해결할 수 없는 상호 연동(interlock) 된 (상호 종속적인) 방정식 집합을 기반으로 EM과 동일한 대체 구조를 갖습니다.

많은 어플리케이션에서 변분 베이즈는 Gibbs 샘플링에 비해 더 빠른 속도의 솔루션을 제공합니다. 그러나 매개 변수를 반복적으로 업데이트하는 데 사용 된 방정식 집합을 파생하려면 종종 유사한 Gibbs 샘플링 방정식을 유도하는 것과 비교하여 많은 양의 작업이 필요합니다. 이는 오직 두 개의 매개 변수와 잠정적 변수가 없는 기본적인 비 계층 적 모델의 경우 아래에 설명 된 것처럼 개념적으로 매우 단순한 많은 모델에 대해서도 마찬가지입니다.

변분 추론에서 관측되지 않은 변수의 집합에 대한 사후분포 와 몇몇 데이터 X 는 변분 분포 Q(Z)으로 근사화한ㄴ다.

분포 Q(Z) 는, 실제 사후 분포 P(Z|X) 와 유사한 Q(Z)을 만들기 위한 의도로 선택된 P(Z|X) 보다 단순한 형태의 분포의 계열에 속하도록 제한된다.

유사성 의 부족은 비유사 함수 d(Q;P) 의 조건으로 측정되고 그러므로 d(Q;P)을 최소화하는 분포 Q(Z)을 선택하여 추론이 된다.

변분 베이즈의 가장 흔한 형태는 비유사 함수의 선택으로 Q로부터 P의 KL발산을 사용한다.

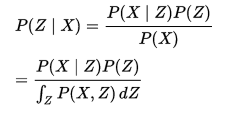
이 선택은 최소화를 다루기 쉽도록 한다.

KL 발산은 아래와 같이 정의된다



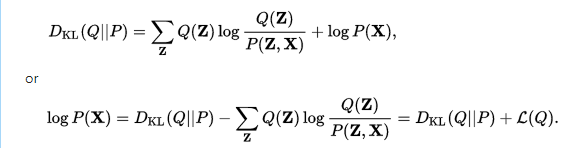
Q와 P는 예상했던 것과는 반대이다. 역 KL- 발산의 사용은 개념적으로 기대 – 최대화(EM) 알고리즘과 유사하다. (다른 방식으로 KL- 발산을 사용하면 기대 전파(Expectation Propagation) 알고리즘이 생성된다

변분 기술은 일반적으로 아래와 같은 근사를 형성하는 데 사용됩니다.

{\displaystyle P(\mathbf {Z} \mid \mathbf {X} )}. 

분모의P (X)를 계산하기 위하여 Z에 대하여 marginalization는 일반적으로 다루기가 어렵다. 예를 들어, Z의 검색 공간이 조합 적으로 크기 때문에. 우리는 Q (Z) ~ P (Z | X) 을 사용하여 근사값을 구합니다.

KL- 발산은 아래와 같이 쓸 수 있습니다.



Q에 대한 로그 증거 ( log P (X))가 고정됨에 따라 마지막 성분  Q을 최대화하여 Q로부터 P의 KL 발산을 최소화한다.

Q의 적절한 선택에 의해, 는 계산되고 최대화 되기 싶다.

따라서 사후분포 P(Z|X) 에 대하여 두개의 분석적인 근사 Q 와 증거 log P(X)에 대하여 하한 경계 를 알고 있다

하한 경계 

hbf {Z}} \ mid {\ mathbf { {\ displaystyle \ log P (\ mathbf {X})}에 대한 하한 {\ displaystyle {\ mathcal {L}} {\ mathcal {L}} (Q) \ log P ({\ mathbf {X}}). "에너지"{\ displaystyle {\ mathcal {L}} (Q) {\ mathcal {L}} (Q) \ 연산자 {E} \_ {{Q}} [\ log P ({\ mathbf {Z}, \ mathbf {X} }}, {\ mathbf {X}})]와 {\ displaystyle Q}의 엔트로피 Q.