Фильтры Гаусса

3.1 Введение

**3**

В этой главе описывается важное семейство рекурсивных методов оценки состояния под общим названием «фильтры Гаусса». Исторически, гауссовские фильтры образуют самые ранние управляемые реализации байесовских фильтров для непрерывных пространств. Они также, безусловно, наиболее популярное семейство среди сегодняшних методов – несмотря на имеющиеся недостатки.

все гауссовские методы разделяют основную идею представления оценок в виде многомерных нормальных распределений. Мы уже сталкивались с определением многомерного нормального распределения в Равенстве (2.4), но, для удобства, приведем его еще раз

**КАНОНИЧЕСКАЯ ПАРАМЕТРИЗАЦИЯN**

image1

Эта плотность по переменной *x* характеризуется двумя наборами параметров: математическим ожиданием µ и ковариационной матрицей ∑. Математическое ожидание µ - это вектор, имеющий аналогичную размерность, что и вектор состояний *x*. Ковариационная матрица – это квадратная симметричная неотрицательно определенная матрица. Ее размерность равна размерности вектора состояний *x*, возведенной в квадрат. В силу этого число элементов в ковариационной матрице квадратично зависит от числа элементов в векторе состояний.

Стремление представить апостериорные вероятности в виде нормального распределения имеет важные последствия. Важнее всего то, что гауссовские распределения одномодальны – они имеют единственный максимум. Такая апостериорнная вероятность характерна для многих проблем отслеживания состояния в робототехнике, где она концентрируется вокруг настоящего состояния с небольшим диапазоном неопределённости. Нормальные апостериорные распределения плохо подходят для многих глобальных проблем оценки, где существует множество отдельных гипотез, каждая из которых образует собственную моду в апостериорном распределении.

Задание параметров нормального распределения в виде математического ожидания и ковариационной матрицы называется параметризацией моментов (moments parameterization)., поскольку математическое ожидание и ковариация – это моменты первого и второго порядка вероятностного распределения (для нормальных распределений все остальные моменты равны нулю).

В данной главе мы также обсудим альтернативный способ задания параметров, называемый канонической параметризацией, или, иногда, обычной параметризацией. Оба вида параметризации, каноническая и в виде моментов, функционально эквиваленты и образуют взаимно однозначные выражения, которые могут трансформировать друг в друга. Однако, это ведет к несколько различным вычислительным характеристикам. Как мы увидим, каноническое и натуральное задание параметров лучше всего воспринимать как дуализм: то, что выглядит вычислительно легким в одном виде параметризации, используется в другом, и наоборот.

В этой главе вводятся два основных алгоритма фильтра Гаусса. В Разделе 3.2 описан фильтр Калмана, реализующий байесовский фильтр, используя задание параметров в виде моментов для ограниченного класса задач с линейной динамикой и функциями измерения.

Фильтр Калмана расширяется для нелинейных задач в Главе 3.3, где описан расширенный фильтр Калмана.

В Разделе 3.4 описан другой нелинейный фильтр Калмана, известный как un­scented Kalman filter.

В Главе 3.5 описан информационный фильтр, который дублирует фильтр Калмана, используя каноническую гауссовскую параметризацию.

3.2 Фильтр Калмана

3.2.1 Линейные гауссовские системы

Возможно, наиболее изученный метод использования байесовских фильтров – это фильтр Калмана или ФК (KF). Фильтр Калмана был открыт Сверлингом (Swerling,1958) и Калманом (Kalman,1960) как способ фильтрации и предсказания в линейных гауссовских системах, которые сейчас будут определены. Фильтр Калмана выполняет вычисление оценок для непрерывных состояний. Он не применим для дискретных или гибридных пространств состояний.

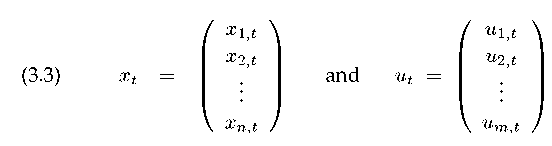
Фильтр Калмана отображает оценки для параметров в виде моментов: для момента времени t, оценка представлена математическим ожиданием pt и ковариационной матрицей Et. Апостериорные вероятности представлены функциями Гаусса если соблюдаются 3 последующих свойства, вдобавок к марковским допущениям для байесовских фильтров.

**GAUSSIAN POSTERIOR**

Аргументы вероятности перехода состояния p(xt | ut,xt\_1) должны быть линейными функциями с добавленным гауссовским шумом. Это выражается следующим равенством:

image2

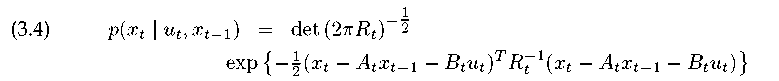
Здесь xt и xt-1 - векторы состояния, а ut – вектор управления в момент времени t. В данной записи оба вектора вертикальны. Они имеют вид



At и Bt – это матрицы. At – квадратная матрица размера n x n, где n размерность вектора состояний xt. Bt имеет размер n x m, где m –размерность вектора управления ut. После перемножения векторов состояния и управления с матрицами At и Bt, соответственно, аргументы функции перехода состояний становятся линейными. Таким образом в фильтрах Калмана достигается линейная динамика системы.

Случайная переменная et в (3.2) это случайный гауссовский вектор, который моделирует неопределенность, привносимую переходом состояний. Он имеет ту же размерность, что и вектор состояний. Математическое ожидание равно нулю, а ковариацию будем обозначать как Rt. Вероятность перехода состояний вида (3.2) называется линейной гауссовской, отражая тот факт, что она линейна в аргументах с добавлением гауссовского шума. Технически, возможно добавить и дополнительную константу в (3.2), но здесь она будет опущена, поскольку в излагаемом материале никакой роли не играет.

Выражение (3.2) определяет вероятность перехода состояний p(xt | ut, xt\_i). Эта вероятность получается после вставки Равенства (3.2) в определение многомерного нормального распределения (3.1). математическое ожидание а постериорного состояния выражается как Atxt\_1 + Btut , а ковариация Rt:



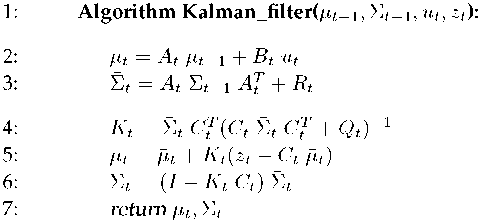


Таблица 3.1 Алгоритм фильтра Калмана для линейных гауссовских переходов состояний и измерений.

2. Вероятность измерения p(zt | xt) также должна иметь линейные аргументы с добавленным гауссовским шумом:

image6

Здесь Ct это матрица размера k x n, где k – размерность вектора измерений zt. Вектор St описывает шум измерений. Распределение St представляет собой многомерную нормальную функцию с нулевым математическим ожиданием и ковариацией Qt. Отсюда, вероятность измерения задается следующим многомерным нормальным распределением:

image7

3. Наконец, первоначальная оценка bel(x0) должна быть нормально распределена. Обозначим математическое ожидание этой величины by и ковариацию - £q:

image8

Этих трех условий достаточно, чтобы гарантировать, что апостериорная вероятность bel(xt) всегда будет гауссовской функций для любого момента времени t. доказательство этого нетривиального утверждения можно найти ниже, в математическом выводе для фильтра Калмана (Глава 3.2.4).

3.2.2 Алгоритм фильтра Калмана

Алгоритм фильтра Калмана приеден в Таблице 3.1. Фильтры Калмана выражают оценку bel(xt) в момент времени t через математическое ожидание jt и ковариацию 2t. На вход фильтр Калмана принимает оценку для момента времени t — 1, выраженную через jt-1 и 2t-1. Для обновления этих параметров, фильтрам Калмана требуются управляющее воздействие ut и измерение zt. Выводом является оценка в момент времени t, выраженная через jt и 2t.

В строках 2 и 3 вычисляются предсказанные значения j и 2 , отображая оценку bel(xt) на один такт позже, но перед учетом измерений zt. Эта оценка получается путем учета управляющего воздействия ut. Математическое ожидание обновляется, используя детерминированную версию функции перехода состояния (3.2), с математическим ожиданием jt-1, замененным на состояние xt-1. Обновление ковариации учитывает тот факт, что состояния зависят от линейной матрицы At. Эта матрица дважды умножается на ковариацию, поскольку ковариация - это квадратичная матрица.

Оценка bel(xt) последовательно преобразуется в требуемую оценку bel(xt) в строках с 4 по 6, учитывая измерения zt. Переменная Kt, вычисленная в строке 4 называется усилением Калмана. Она определяет степень, до которой измерение учитывается в новой оценке состояния, а как именно – станет понятнее в Главе 3.2.4. Строка 5 управляет математическим ожиданием, меняя его пропорционально усилению Калмана Kt и отклонению текущего измерения, zt, а также измерения, предсказанного согласно вероятности измерения (3.5). Ключевой концепцией здесь является коррекция, которая представляет собой разницу между реальным измерением zt и ожидаемым измерением Ct jt в строке 5. Наконец, в строке 6 вычисляется новая ковариационная матрица для апостериорной оценки, учитывающая добавочную информацию полученную измерением.

**КОРРЕКЦИЯ**

**УСИЛЕНИЕ КАЛМАНА**

Фильтр Калмана довольно эффективен в вычислительном смысле. Для лучших, на сегодняшний день, алгоритмов сложность инверсии матрицы размером d x d составляет приблизительно O(d2'4). Каждая итерация фильтра Калмана, как было сказано, ограничена снизу сложностью (приблизительно) O(k24), где k – размерность вектора измерений zt. Эта (приблизительно) кубическая степень сложности получается из-за инверсии матрицы в строке 4. Даже для некоторых редких коррекций, описанных в будущих главах, она составляет, по крайней мере O(n2), где n – размерность пространства состояний, из-за перемножения в строке 6 (матрица KtCt может быть разреженной). Для многих реализаций (скажем, таких, как построение карт, обсуждаемых в более поздних главах) пространство измерений имеет намного меньшую размерность, чем пространство состояний, и такт обновления, в основном, состоит из операций сложности O(n2).

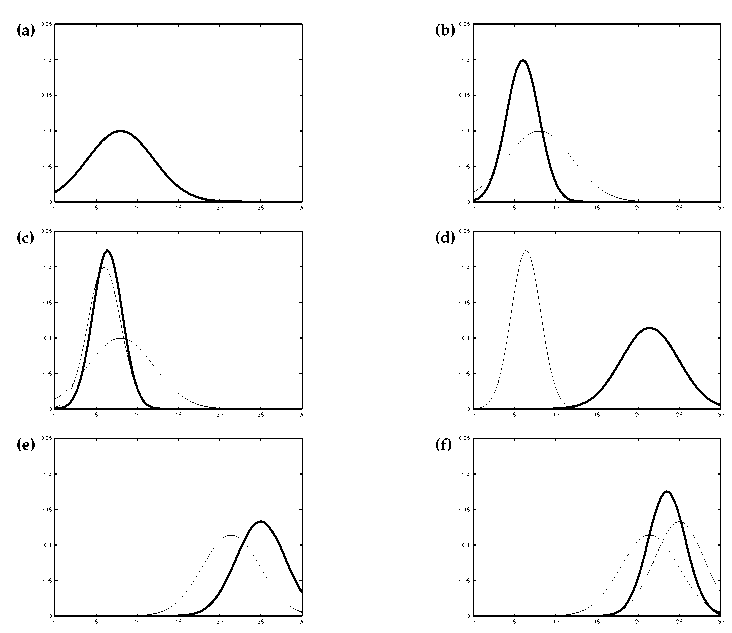


Рис. 3.2 Иллюстрация фильтров Калмана: (a) первоначальная оценка, (b) измерение (выделено жирным) с соответствующей неопределенностью, (c) оценка после учета измерения алгоритмом фильтра Калмана, (d) оценка после движения вправо (которое вносит неопределенность), (e) новое измерение с соответствующей неопределенностью, и (f) результирующая оценка.

3.2.3 Иллюстрация

На Рис. 3.2 показан алгоритм фильтра Калмана для простейшего сценария локализации с одним измерением. Допустим, на каждой диаграмме Рис. 3.2 робот двигается вдоль горизонтальной оси. Пусть априорная оценка местоположения робота задается нормальным распределением, показанным на Рис. 3.2a. Робот запрашивает данные своего местоположения с датчиков (например, системы GPS), и возвращает результат измерения в центре выделенного жирным пика функции Гаусса на Рис. 3.2b. Выделенный жирным график гауссовой функции иллюстрирует это измерение: пик обозначает значение, предсказанное датчиками, а ширина (отклонение) связана с неопределенностью измерения. Комбинация априорной оценки и измерения в строках с 4 по 6 алгоритма фильтра Калмана в Таблице 3.1 дает жирную гауссиану на Рис. 3.2c. Значение математического ожидания этой оценки лежит между двух первичных пиков, а неопределенность - меньше, чем у обоих предыдущих гауссовских функций. Факт того, что остаточная неопределенность меньше, чем у исходных гауссиан, может показаться контринтуитивным, но он является общей характеристикой интеграции информации в калмановских фильтрах.

Далее, допустим, робот движется вправо. Степень неопределенности возрастает в силу стохастического перехода состояния. Строки 2 и 3 калмановского фильтра дают гауссовскую кривую, показанную жирным на Рис. 3.2d. Эта кривая сдвинута на величину перемещения робота, и, в силу только что озвученных причин, шире. Робот получает второе измерение, показанное жирным графиком гауссовской функции на Рис. 3.2e, которое затем ведет к апостериорной оценке, показанной на Рис. 3.2f.

Как показывает пример, фильтр Калмана последовательно чередует такт обновления измерений (строки 5-7), в которых данные датчиков интегрируются в текущую оценку, с тактом прогноза (или тактом обновления управления), который модифицирует оценку в соответствии с действием. Такт обновления уменьшает, а такт прогнозирования – увеличивает неопределенность оценки робота.

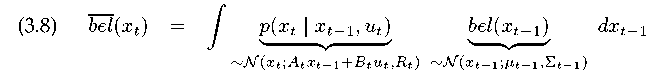
Математический вывод KF

В этом разделе выполняется вывод алгоритмов фильтрации Калмана из Таблицы 3.1. Этот раздел можно смело пропускать при первом чтении и включен из соображений полноты изложения.

Во-первых, вывод KF, по большей части, это упражнение по манипулированию квадратичными выражениями. Экспоненты добавляются, например, при перемножении двух гауссовских функций. Поскольку обе первичные экспоненты квадратичные, результирующая сумма тоже. Оставшиеся усилия состоят в представлении результата в форму, которая делает возможным прочтение необходимых параметров.

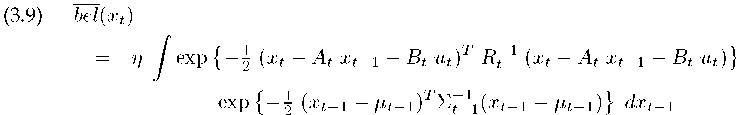
Часть 1: Прогноз

Вывод начинается со строк 2 и 3 алгоритма, в которых оценка bel(xt) вычисляется на основе оценки, полученной на предыдущем шаге, bel(xt\_i). В строках 2 и 3 выполняется такт обновления, описанный Равенством (2.41), и приведенного здесь для удобства читателя:

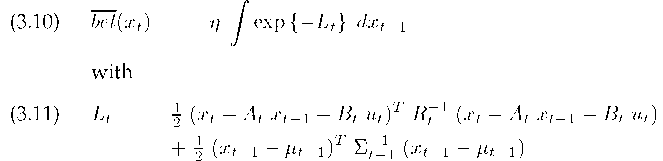


Оценка bel(xt\_ 1) представлена математическим ожиданием (или «средним», встречается и так и так – прим перев) pt\_ 1 и ковариацией 2t\_ 1. Вероятность перехода состояния p(xt | xt-1, ut) была дана в (3.4) в виде нормального распределения по xt с математическим ожиданием Atxt-1 + Btut и ковариацией Rt. Как сейчас будет продемонстрировано, результат (3.8) тоже представляет собой гауссовскую функцию с математическим ожиданием p,t и ковариацией 21, как указано в Таблице 3.1.

Начнем с записи (3.8) в гауссовской форме:



Короче, мы имеем



Заметим, что Lt квадратично как для xt-1, так и для xt.

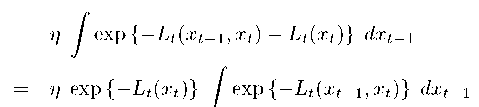
Выражение (3.10) содержит интеграл. Решение интеграла требует переопределения членов на этом интервале способом, который поначалу кажется контринтуитивным. В частности, мы выполним разложение Lt на две функции, Lt(xt-1, xt) и Lt(xt):

image13

Это разложение будет просто результатом перегруппировки членов в Lt. Главной целью этапа разложения должно стать разделение переменных в Lt на два множества, из которых только одно будет зависеть от переменной xt\_1. Другое, Lt(xt), от xt-1 зависеть не будет. В результате, мы сможем вынести второй набор переменных из-под интеграла по переменной xt-1.

Это иллюстрируется следующим преобразованием:

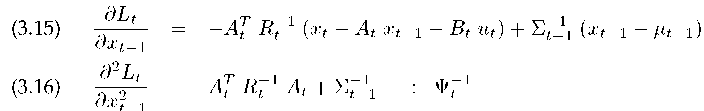
image14



Конечно, существует множество способов разложить Lt на два множества, которые удовлетворяют этому критерию. Ключевой мыслью будет выбор Lt(xt-1,xt) таким образом, что значение интеграла в (3.13) не зависит от xt. Если нам удастся определить такую функцию Lt(xt-1 ,xt), весь интеграл по Lt(xt-1,xt) для проблемы определения распределения по xt станет просто константой. Константы обычно учитываются постоянным нормализационным членом n, поэтому в нашем разложении мы сможем включить константу в n (для другого действительного значения n, как указано выше):

image16Таким образом, разложение делает возможным исключение интеграла из оценки (3.10). Результатом является нормализованная экспонента квадратичной функции, которая представляет собой нормальное распределение.

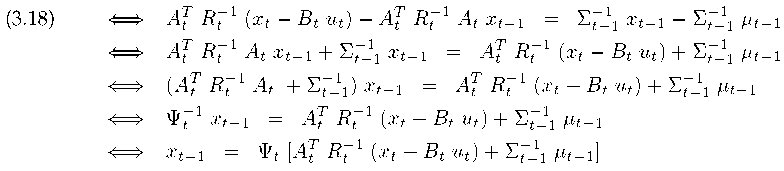
Давайте выполним это разложение. Мы ищем функцию Lt(xt-1,xt), квадратичную в xt-1. (Эта функция будет также зависеть от xt, но сейчас это нас не волнует.) Для определения коэффициентов квадратичной функции вычислим первые две производные Lt:



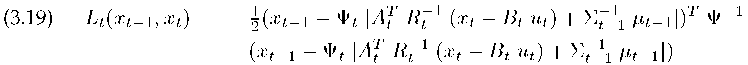
Определяет кривизну Lt(xt-1,xt). Приняв первую производную Lt за 0 получим математическое ожидание:

**image18**

Сейчас выражение решено для xt-1



Итак, теперь имеется квадратичная функция Lt(xt−1, xt), определенная следующим образом:

****

Вообще-то, это не единственная квадратичная функция, удовлетворяющая разложению в (3.12). Однако, Lt(xt-1, xt) – это общая квадратичная форма отрицательной экспоненты нормального распределения. Фактически, функция

image21

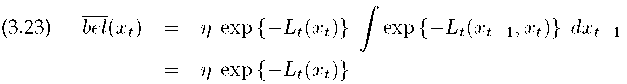
Представляет собой верную функцию плотности вероятности (ФПВ) для переменной xt-1. Как читатель может легко убедиться, форма этой функции определена в (3.1). Из (2.5) известно, что ФПВ интегрируются до 1. Отсюда, получаем

image22

Из чего следует

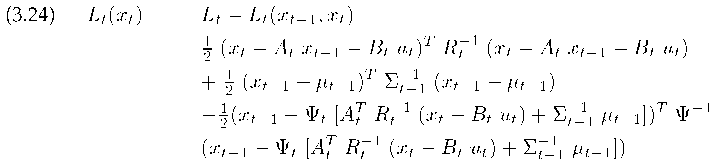
image23

Важно заметить, что значение этого интеграла не зависит от xt, нашей целевой переменной. Поэтому, в нашей задаче вычисления распределения по xt, значение интеграла будет константой. Включая константу в нормализационный член П, получаем следующее выражение для Равенства (3.13):

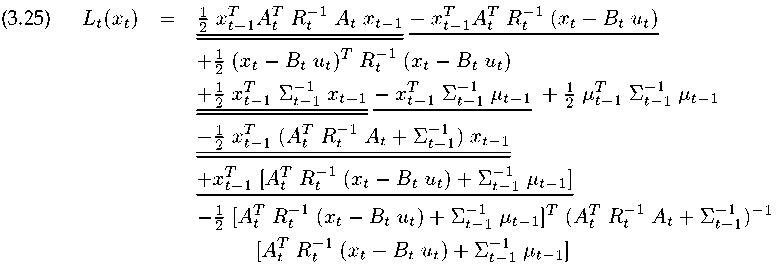


Это разложение подтверждает правильность (3.14). снова заметим, что нормализационные члены n для обоих строк не одинаковы.

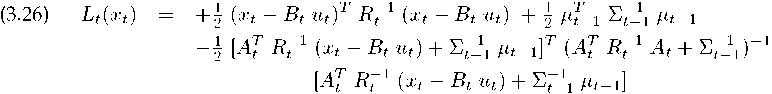
Остается определить функцию Lt(xt), которая является разницей Lt, определенной в (3.11), и Lt(xt-1,xt), определенной в (3.19):



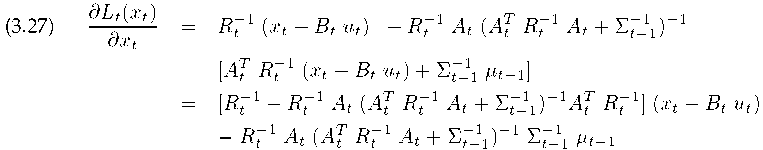
Давайте быстро убедимся, что Lt(xt) действительно не зависит от xt-1. Для этого произведем обратную замену ^t = (Af R-1 At + E-\_11)\_1, и перемножим части, указанные выше. Для удобства читателя, члены, содержащие xt-1 выделены подчеркиванием (двойным, если они квадратичны в xt\_i).



Несложно заметить, что все члены, содержащие xt-1, сокращаются. Это неудивительно, поскольку это условие нашей конструкции функции Lt(xt-i,xt).



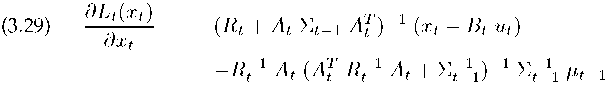
Далее, Lt(xt) квадратично на xt. Это наблюдение означает, что bel(xt) – действительно нормальное распределено. Математическое ожидание и ковариация этого распределения, конечно, минимум и кривизна Lt(xt), которые мы теперь легко можем получить вычислением первой и второй производных Lt(xt) по отношению к xt:



*Лемма об обращении матриц,* указанная (и показанная) в Таблице 3.2 позволяет выразить первый множитель следующим образом:

**image29**

Следовательно, искомая производная задается следующим выражением:



Лемма об обращении матриц. Для любых инвертируемых квадратных матриц R и Q и любой матрицы P соответствующих размерностей, следующее утверждение истинно

image31

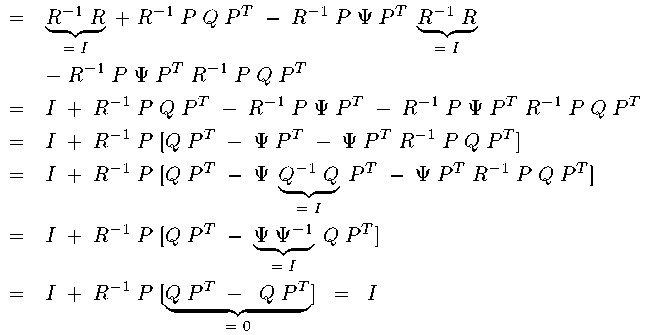
При условии, что все три указанные матрицы обратимы указанным образом.

**Доказательство.** Определимimage32. Этого достаточно, чтобы показать, что

image33

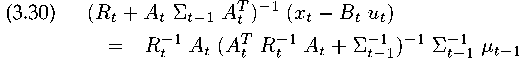
Это можно показать путем серии преобразований

**It suffices to show that**



**Таблица 3.2** (Специализированная) лемма об обращении матриц, иногда называемая *формулой Шермана/Моррисона*.

Минимум Lt(xt) достигается, когда первая производная равна нулю.



Решение для целевой переменной xt дает, на удивление, компактный результат

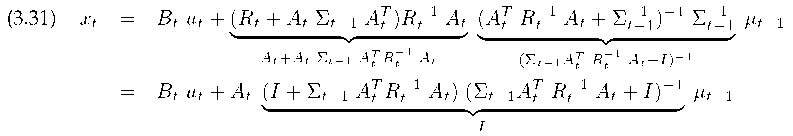


image37

Таким образом, математическое ожидание оценки bel(xt) после учет команды на движение ut в виде Bt ut + At jt-1. Это доказывает правильность строки 2 в алгоритме фильтра Калмана Таблицы 3.1.

Сейчас строка 3 получается путем вычисления второй производной от Lt(xt):

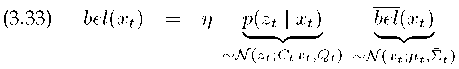
image38

Это кривизна квадратичной функции Lt(xt), инверсия которой представляет собой ковариацию оценки bel(xt).

Подведем итог. Мы показали, что такты прогноза в строках 2 и 3 алгоритма фильтра Калмана в действительности реализуют такт прогноза байесовского фильтра. Чтобы это сделать, сначала на две функции, Lt(xt-1,xt) и Lt(xt), была разделена экспонента оценки bel(xt). Затем мы показали, что Lt(xt-1, xt) изменяет прогнозируемую оценку bel(xt) только на постоянный множитель, который будет включен в нормализационный член p. Наконец, мы определили функцию Lt(xt), и показали, что ее результат представляет собой математическое ожидание j2t и ковариацию 21 прогноза фильтра Калмана bel(xt).

Часть 2: Обновление измерения

Сделаем вывод обновления измерения в строках 4,5, и 6 (Таблица 3.1) нашего алгоритма фильтра Калмана. Начнем с общего механизма учета измерения в байесовском фильтре, определенном в Равенстве (2.38) и приведенном здесь в аннотированном виде:



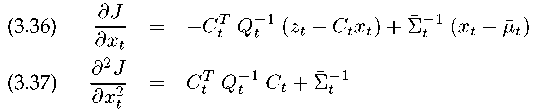
Математическое ожидание и ковариация bel(xt), очевидно, задаются j2t и 21. Вероятность измерения p(zt | xt) была определена в (3.6и так же является нормальным распределением, с математическим ожиданием Ct xt и ковариацией Qt. В силу этого, произведение дает экспоненту

**image40**

С

image41

Для xt эта функция квадратична, следовательно, bel(xt) – это нормальное распределение. Чтобы вычислить его параметры, снова вычислим первые две производные Jt по отношению к xt:



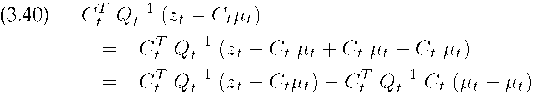
Второй член – это инверсия ковариации bel(xt):

image43

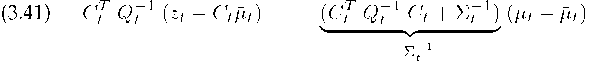
Математическое ожидание bel(xt) представляет собой минимум квадратичной функции, который вычисляем путем приравнивания Jt к нулю (и замены ^t на xt):

image44

Выражение слева от знака равенства может быть преобразовано следующим образом:

****

Производим обратную замену для (3.39), что дает

****

Откуда получается

**image47**

Теперь можем определить усиление Калмана как

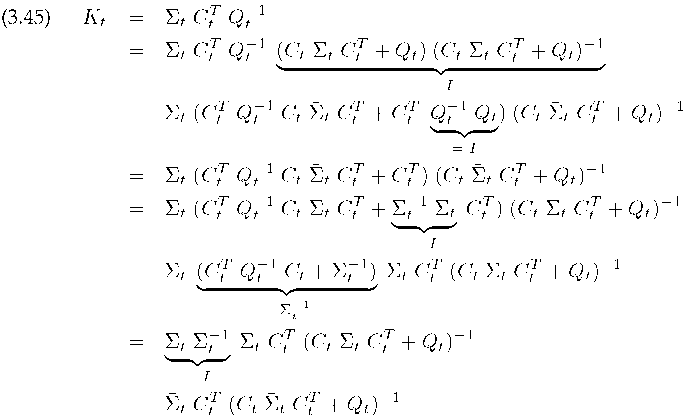
**image48**

Aи получить

image49

Это доказывает правильность строки 5 в алгоритме фильтра Калмана в Таблице 3.1

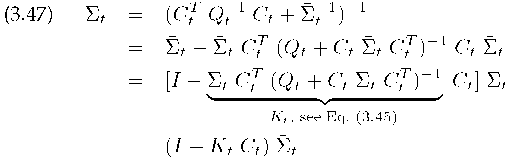
Усиление Калмана, как определено (3.43), является функцией Et. Это расходится с фактом использования Kt для вычисления Et в строке 6 алгоритма. Следующее преобразование демонстрирует, каким образом выразить Kt в терминах ковариации, отличных от Et. Оно начиается с определения Kt в (3.43):



Это выражение доказывает правильность строки 4 нашего алгоритма фильтра Калмана.

Строка 6 получается в результате выражения ковариации, используя калмановское усиление Kt. Преимущество вычисления в Таблице 3.1 над определением в Равенстве (3.38) заключается в том, что мы можем избежать инверсии ковариационной матрицы состояния. Это важно для использования калмановских фильтров в пространствах состояний высокой размерности

Снова трансформация производится с использованием леммы об обращении матриц, которая уже была указана в Таблице 3.2. Заново приведем ее, используя нотацию Равенства (3.38): **image51**

Это позволяет прийти к следующему выражению для ковариационной матрицы: 

На этом доказательство корректности завершается, поскольку показана правильность строки 6 нашего алгоритма калмановского фильтра.

3.3 Расширенный фильтр Калмана

3.3.1 Зачем линеаризовывать?

Допущения о том, что наблюдения являются линейными функциями состояния и что следующее состояние является линейной функцией предыдущего, являются важнейшими для работы калмановского фильтра. Наблюдение о том, что любое линейное преобразование случайной гауссовской переменной дает в результате другую гауссовскую случайную переменную играет важную роль в выводе алгоритма фильтра Калмана. Эффективность фильтра Калмана зависит от того, возможно ли выразить параметры результирующей гауссовской функции в закрытой форме.

В этой и последующих главах мы проиллюстрируем свойства различных представлений плотности, используя преобразование одномерной случайной гауссовой переменной. На Рис. 3.3a показано линейное преобразование такой случайной переменной. На графике справа внизу изображена плотность случайной переменной X ~ N(x; y, n2). Допустим, X проходит по линейной функции y = ax + b, показанной на графике справа вверху. Результирующая случайная переменная, Y, распределяется нормально, с математическим ожиданием ay,+b и среднеквадартичным отклонением a2a2. Эта гауссова функция показана в виде серой зоны вверху слева графика на Рис. 3.3a. Читатель может заметить тесную связь данного примера с обновлением следующего состояния фильтра Калмана, где X = xt-i , а Y = xt, но без добавочной переменной зашумления. См также Равенство (3.2).

К сожалению, на практике переходы состояний и измерения редко линейны. Например, робот, который движется с постоянной поступательной и вращательной скоростью, обычно движется по круговой траектории, которую невозможно описать линейными переходами состояния. Это наблюдение, вкупе с допущением одномодовых оценок, делает простые калмановские, такие, как обсуждалось, неприменимыми для всех робототехнических задач, кроме самых тривиальных.

Расширенный фильтр Калмана, или EKF (РФК?), снимает одно из допущений: линейность. Вместо этого принимается допущение о том, что вероятности перехода состояний и измерений задаются нелинейными функциями g и h, соответственно:

**РАСШИРЕННЫЙ ФИЛЬТР КАЛМАНА**

image53

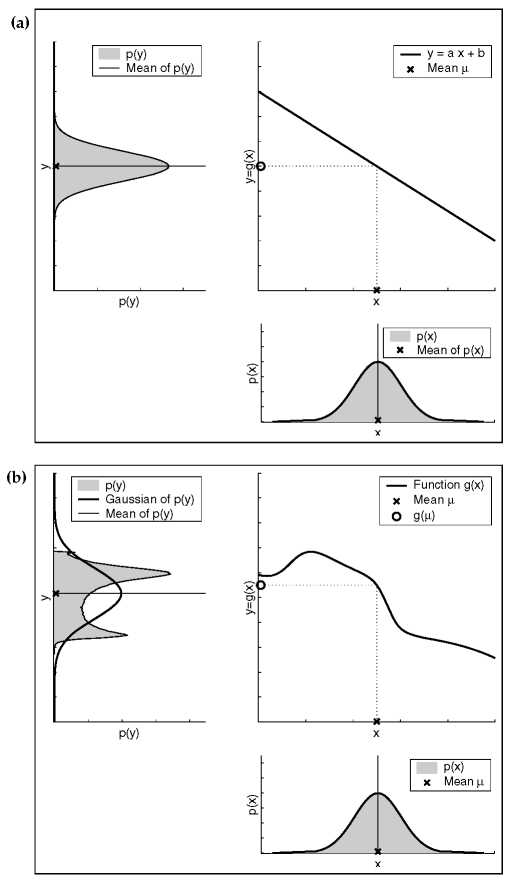


Рис 3.3 Линейное (a) и нелинейное (b) преобразование случайной гауссовской переменной. На нижнем правом графике показана плотность оригинальной случайной переменной, X. Эта случайная переменная проходит через функцию, показанную на верхних правых графиках (пунктиром показано математическое ожидание). Плотность результирующей случайной переменной Y показано на верхних графиках слева.

Эта модель генерализует линейную гауссовскую модель, лежащую в основе фильтров Калмана, как допускается Равенствами (3.2) и (3.5). Функция g заменяет матрицы At и Bt в (3.2), а h заменяет матрицу Ct в (3.5). К сожалению, при произвольных функциях g и h, оценка больше не представляет собой гауссовское распределение. Фактически, выполнение обновления оценки обычным образом обычно невозможно для нелинейных функций g и h, а байесовский фильтр не дает решения в закрытой форме.

На Рис. 3.3b показано воздействие нелинейного преобразования на случайную гауссовскую переменную. На графиках снизу справа и сверху справа показаны случайная переменная X и нелинейная функция g, соответственно. Плотность преобразованной случайной переменной, Y = g(X), обозначена серой зоной на верхнем левом графике Рис. 3.3b. поскольку эта плотность е может быть вычислена в закрытой форме, оценочно возьмем 500000 значений по p(x), передадим через функцию g, а затем отразим на гистограмме в диапазоне значений g. Как видно, Y не гауссовская функция, в силу того, что нелинейности в функции g искажают вероятность X, исказив форму нормального распределения.

Расширенный фильтр Калмана (EKF) вычисляет гауссовскую аппроксимацию настоящей оценки. Пунктиром на графике в верхнем левом углу на Рис. 3.3b показано гауссовское приближение случайной переменной Y. Соответственно, EKF представляет оценку bel(xt) в момент времени t в виде математического ожидания gt и ковариационной матрицы Et. Таким образом, EKF наследуют от фильтра Калмана общее отображение оценки, но отличается в том, что эта оценка лишь приближенная, а не точная, как в случае калмановских фильтров. Таким образом, цель EKF смещается с вычисления точной апостериорной вероятности на эффективную приближенную оценку ее математического ожидания и ковариации. Но, поскольку эти статистики не могут быть вычислены в закрытой форме, в случае EKF необходимо прибегать к дополнительной аппроксимации.

Линеаризация разложением в ряд Тейлора

Ключевая идея в основе EKF называется линеаризация. На Рис. 3.4 показана общая концепция. Линеаризация выполняет приближение нелинейной функции g с помощью линейной функции, касательной к g в точке математического ожидания гауссиана (пунктирная линия на верхнем правом графике). Опускается проекция гауссовой функции с помощью результатов линейной аппроксимации плотности, как показано пунктирной линией на верхнем левом графике. Сплошной линией на левом верхнем графике показывает математическое ожидание и ковариацию аппроксимации Монте-Карло. Разница между этими двумя гауссианами показывает ошибку, вызванную линейной аппроксимацией функции g.

Однако, ключевое преимущество линеаризации лежит в ее эффективности.

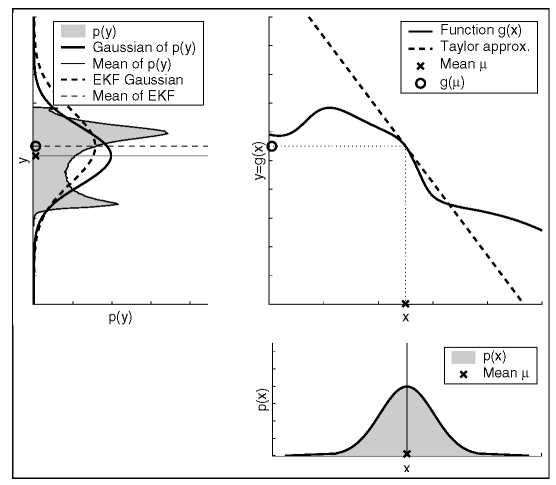


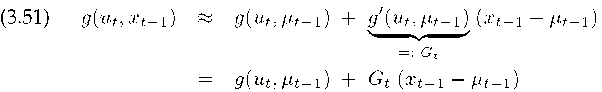
Рис. 3.4 Иллюстрация линеаризации, применяемой для EKF. Вместо представления гауссовой функции через нелинейную функцию g, он строится методом линейной аппроксимации g. Линейная функция касательная к g в точке математического ожидания настоящего гауссиана. Результирующая гауссовская функция показана пунктиром на левом верхнем графике. Линеаризация влечет за собой потери на аппроксимацию, что видно в виде разницы между линеаризованной гауссовской функцией (пунктирной) и гауссианом, вычисленным с помощью очень точной оценки методом Монте-Карло (сплошная линия).

Оценка гауссовской функции методом Монте-Карло была получена заданием 500000 на функцию g с последующим вычислением математического ожидания и ковариации. Линеаризация, применяемая в EKF, с другой стороны, требует лишь определения линейной аппроксимации с последующим вычислением результирующей гауссовской функции в закрытой форме. Фактически, после линеаризации g механика распространения оценки в EKF эквивалентно таковой в фильтрах Калмана.

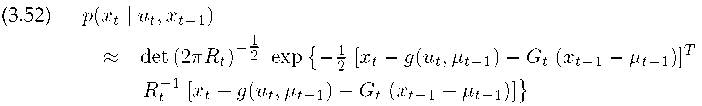
Этот метод применим также при перемножении гауссовских функций, когда участвует функция измерения h. И снова EKF выполняет приближение h с помощью касательной линейной функции, сохраняя, таким образом, гауссовскую природу апостериорной оценки.

РАЗЛОЖЕНИЕ В РЯД ТЕЙЛОРА

Существует множество методов для линеаризации нелинейных функций. В EKF используется способ, называемый разложением в ряд Тейлора (первого порядка). Разложение в ряд Тейлора создает линейное приближение функции g на основании значения g и наклона. Наклон дается частной производнойimage56

Разумеется, и значение g и наклон зависят от аргументов g. Логичным способом выбрать аргумент является выбор состояния, которое, вероятнее всего, будет на момент линеаризации. Для гауссовских функций наиболее вероятное состояние, это среднее апостериорной вероятности gt-1. Другими словами, функция g аппроксимируется по значению в gt-1 (и в ut), а линейная экстраполяция достигается с помощью члена, пропорционального градиенту g в gt-1 и ut: 

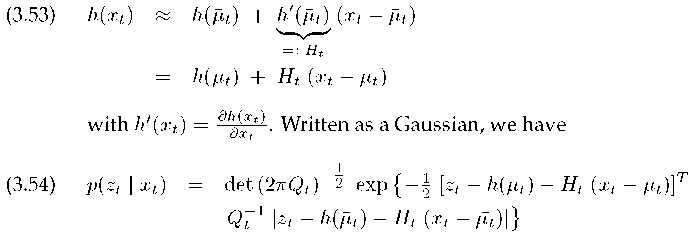
Записанная в виде гауссвских функций, вероятность перехода состояния в приближенном виде выглядит следующим образом:

****

ЯКОБИАН

Заметим, что Gt – это матрица размера n x n, где n выражает размерность состояния. Эта матрица часто называется якобиан. Значение якобиана зависит от ut и gt-1, и, в силу этого, различается для разных моментов времени.

В EKF используется точно такая же линеаризация для функции измерения h. Здесь разложение в ряд Тейлора происходит по gt, наиболее вероятное состояние определяется роботом для момента времени, который был линеаризован по h:



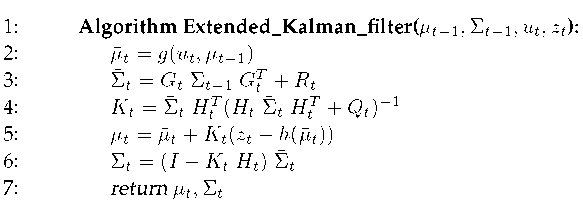
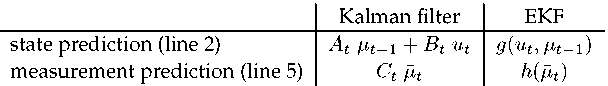


Таблица 3.3 Алгоритм расширенного фильтра Калмана.

3.3.3 Алгоритм EKF

В Таблице 3.3 приводится алгоритм расширенного фильтра Калмана. Во многих отношениях, этот алгоритм похож на алгоритм калмановского фильтра, приведенный в Таблице 3.1. Наиболее важные различия сведены в следующую таблицу:

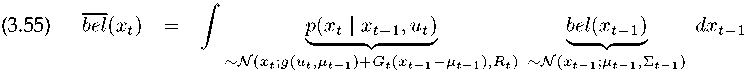


|  |  |  |
| --- | --- | --- |
|  | Фильтр Калмана | EKF |
| Прогноз состояния (строка 2) |  |  |
| Прогноз измерения (строка 5) |  |  |

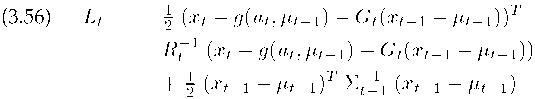
Таким образом, линейные прогнозы в калмановских фильтрах заменяются в EKF нелинейными генерализациями. Более того, в EKF используются матриц Якоби для Gt и Ht вместо соответствующих линейных систем гауссовских функций At, Bt, и Ct в калмановских фильтрах. Якобиан Gt соответствует матрицам At и Bt, а якобиан Ht - соответсвует Ct. Детальный пример расширенных калмановских фильтров будет дан в Главе 7.

3.3.4 Математический вывод для EKF

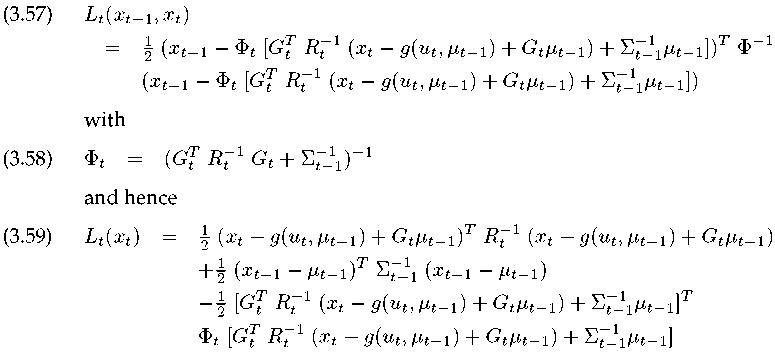
Математический вывод EKF аналогичен выводу для фильтра Калмана в Главе 3.2.4, поэтому, приводится лишь эскизно. Прогноз вычисляется следующим образом (c.f. (3.8)):



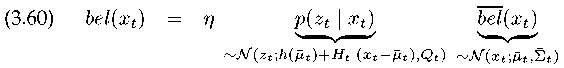
Это распределение для EKF является аналогом прогнозного распределения в калмановском фильтре, приведенном в (3.8). Гауссовская функция p(xt | xt-l,ut) приводится в Равенстве (3.52). Функция Lt задана с помощью (c.f. (3.11))



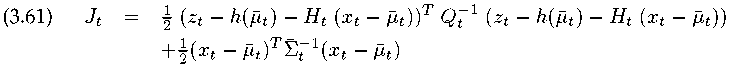
Которая является квадратичной как для xt-1, так и для xt, как указано выше. Как и в (3.12), разложим Lt на Lt(xt\_i,xt) и Lt(xt):

Как пользователь может с легкость удостовериться, установка первой производной Lt(xt) в нуль дает обновление ^t = g(ut,gt-1), по аналогии с выводом Равенства от (3.27) до (3.31). Вторая производная задана (Rt + Gt Et\_i GT)-1 (см (3.32)).

Обновление измерения также выводится так же, как и для фильтра Калмана в Главе 3.2.4. По аналогии с (3.33), для EKF получается



Используем линеаризованную функцию перехода состояния из (3.53). Это дает следующую экспоненту (см (3.35))

****

Получившиеся в результате математическое ожидание и ковариация заданы с помощью

image67

Калмановское усиление равно

image68

Вывод этих равенств аналогичен Равенствам от (3.36) до (3.47).

3.3.5 Практические соображения

EKF стал чуть ли не самым популярным инструментом для оценки состояния в робототехнике. Его сильными сторонами является простота и вычислительная эффективность. Как и в случае фильтров Калмана, каждое обновление требует времени O(k2 4 + n2), где k – размерность вектора измерений zt, а n – размерность вектора состояний xt. Другие алгоритмы, например, обсуждаемые ниже многочастичные фильтры, могут потребовать времени, экспоненциального к n.

Вычислительная эффективность EKF основана на факте представления оценки с помощью многомерного гауссовского распределения. Гауссовская функция представляет собой одномодальное распределение, которое можно считать единственной догадкой, дополненной неопределенностью в виде эллипса. Гауссиана является надежной функцией оценки для многих практических задач. Использование калмановского фильтра на пространствах состояний с 1000 и более измерений будет обсуждаться далее в книге. EKF с большим успехом были использованы в ряде задач оценки состояния, которые нарушают указанные допущения.

Важное ограничение EKF возникает в силу аппроксимации перехода состояний и измерений с использованием линейных разложений в ряд Тейлора. Для большинства задач робототехники переходы состояния и измерения нелинейны. Качество линейной аппроксимации, применяемой в EKF, зависит от двух основных факторов: Степени неопределенности и степени локальной нелинейности аппроксимируемых функций. На двух графиках на Рис. 3.5 показана зависимость от неопределенности. На них две случайные гауссовские переменные пропускаются через одну и ту же нелинейную функцию (так же см. Рис. 3.4). Хотя обе гауссовские функции имеют одинаковое математическое ожидание, переменная, показанная на графике (a) имеет более высокую степень неопределенности, чем (b). Поскольку разложение в ряд Тейлора зависит только от математического ожидания, обе гауссовские функции проходят через одинаковую линейную аппроксимацию. Серые зоны на графике слева сверху обозначают плотности результирующей случайной переменной, вычисленной с помощью оценки Монте-Карло. Вероятность, выраженная более широкой гауссовской кривой, значительно более искажена по сравнению с более узким и менее неопределенным гауссианом. Нормальные аппроксимации этих плотностей показаны на рисунках сплошными линиями. Пунктиром на графике обозначены линеаризации

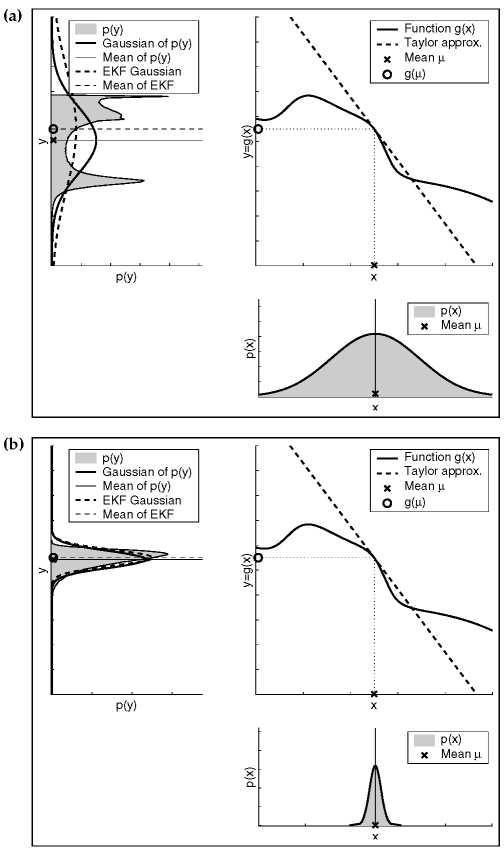


Рис. 3.5 Зависимость качества аппроксимации от неопределенности. Обе гауссовские функции (справа снизу) имеют одинаковое математическое ожидание и подаются на одну и ту же нелинейную функцию (сверху справа). Большая неопределенность левой гауссовской функции дает в результате более искаженную плотность результирующей случайной переменной (серая область слева сверху). Сплошными линиями на графиках слева сверху показаны гауссовские функции, основанные на плотностях. Пунктиром обозначены гауссовские функции, полученные в результате линеаризации в EKF.

Сравнение гауссовских функций, полученных в результате аппроксимаций методом Монте-Карло, показывает, что большая степень неопределенности обычно ведет к менее точной оценке математического ожидания и ковариации результирующей случайной переменной.

Второй фактор, влияющий на качество линейного приближения гауссовских функций, это локальная нелинейность функции g, что показано на Рис. 3.6. Показанные две гауссовские функции с одинаковым среднеквадратичным отклонением проходят через одну и ту же нелинейную функцию. На панели (a), математическое ожидание гауссовской функции попадает в диапазон функции g с большей степенью нелинейности по сравнению с изображенной на панели (b). Разница между точной оценкой гауссовской функции методом Монте-Карло (сплошная линия, сверху слева) и полученной в результате линейной аппроксимации (пунктирная линия) демонстрирует, что более высокая нелинейность влечет увеличение ошибок аппроксимации. Совершенно очевидно, что в гауссовской функции EKF недооценивается размах получившейся плотности.

Иногда необходимо иметь дело с несколькими отдельными гипотезами. Например, у робота может быть две отдельные гипотезы относительно того, где он находится, но среднее арифметическое этих гипотез не выглядит правдоподобным. Для таких ситуаций требуется многомодальные представления апостериорной оценки. EKF, в описанной форме, неспособны отобразить такие многомодальные оценки. Обычным расширением EKF является отображение апостериорных вероятностей, используя смеси, или суммы гауссовых функций. Сумма гауссовых функций может иметь формуimage70

**СМЕСЬ ГАУССОВЫХ ФУНКЦИЙ**

Здесь фг,1 – это параметры суммирования фt,i > 0. Эти параметры служат весами компонентов смеси. Они оцениваются на основе подобия наблюдений для соответствующих гауссовых функций. EKF, в которых используются такие представления в виде смеси называются multi-hypothesis (ex­tended) Kalman filters, or MHEKF.

**EKFС НЕСКОЛЬКИМИ ГИПОТЕЗАМИ**

Подводя итог, в случае, если нелинейные функции приблизительно линейны в области среднего оценки, то аппроксимация EKF будет, в общем, хорошим выбором для вычисления апостериорной оценки с удовлетворительной точностью. Более того, чем меньше определенность робота, тем шире будет гауссова оценка, и тем сильнее она подвержена нелинейности функций перехода состояний и измерения. На практике, когда применяются EKF, важно следить за тем, чтобы неопределенность оценки состояния оставалась мала.

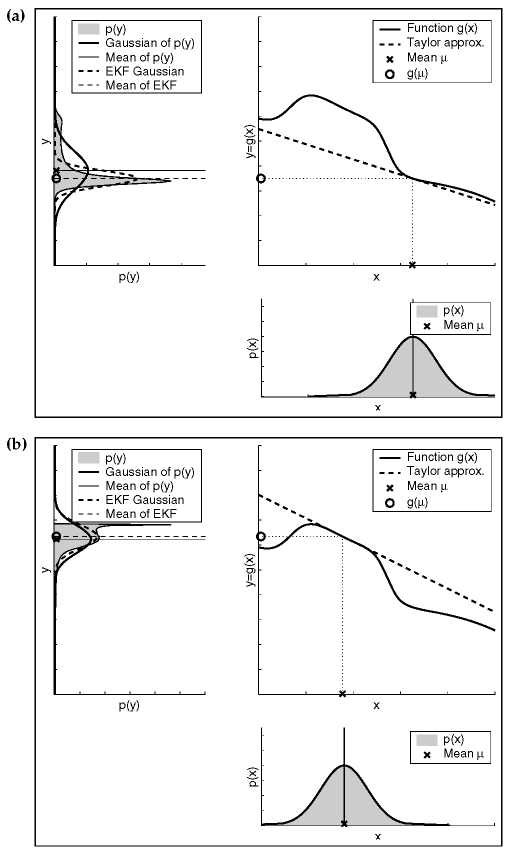


Рис. 3.6 Зависимость качества аппроксимации от локальной нелинейности функции g. Обе гауссовы функции (изображены внизу справа на обоих панелях) имеют одинаковую ковариацию и передаются через одну и ту же функцию (вверху справа). Линейная аппроксимация, применяемая в EKF, показана пунктиром на графиках справа сверху. Сплошные линии на графиках слева сверху обозначают гауссовы функции, полученные из высокоточных оценок методом Монте-Карло. Пунктирными линиями обозначены гауссовы функции, сгенерированные в результате линеаризации EKF.

3.4 Unscented Kalman Filter

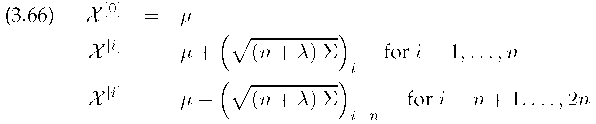
Разложение в ряд Тейлора, применяемое для EKF, является не единственным способом линеаризации преобразования гауссовой функции. Были найдены два других подхода, которые дают лучшие результаты. Один известен, как как выравнивание моментов (moments matching) (а фильтр известен под названием фильтр предполагаемой плотности (assumed density filter - ADF)), в котором линеаризация вычислена способом, сохраняющим настоящее значение математического ожидания и ковариации апостериорного распределения (чего не происходит в случае EKF). Другой метод линеаризации использован в unscented Kalman filter (нет русского термина – при перев), или UKF, который выполняет стохастическую линеаризацию, используя процесс взвешенной статистической линейной регрессии. Обсудим UKF без математического вывода. Мы призываем читателя детально ознакомиться с темой в литературе, указанной в библиографическом примечании.

**UNSCENTED KALMAN FILTER**

3.4.1 Линеаризация с помощью Unscented преобразования

На Рис. 3.7 показана линеаризация, применяемая в UKF, под названием unscented преобразование. Вместо приближения функции g разложением в ряд Тейлора

СИГМА ТОЧКА

UKF детерминированно извлекает так называет сигма-точки гауссовой кривой и передает их на функции g. В общем случае, эти сигма-точки расположены на среднем значении и симметричны вдоль главных осей ковариации (две на измерение). Для n-мерного гауссиана с математическим ожиданием μ и ковариацией Σ, в результате получается 2n + 1 сигма- точка X [i] , выбранные согласно следующему правилу: ****

Здесь A = a2 (n+к) — n, где a и к –параметры масштабирования, которые определяют, как далеко сигма точки расположены от среднего значения. У каждой сигма точки X имеет два связанных с нею веса. Первый вес, w^j, используется для вычисления математического ожидания, а другой вес, wC, применяется при восстановлении ковариации гауссовой функции.

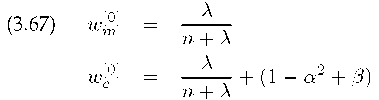
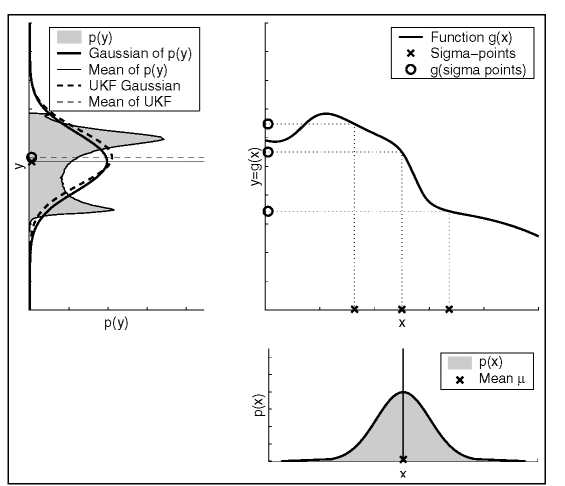


image75

****

**Рис. 3.7** Иллюстрация линеаризации, применяемой вUKF. Алгоритм фильтра сначала извлекает 2n + 1 взвешенных сигма точек из n-мерного гауссиана (в данном примере n = 1). Эти сигма точки передаются нелинейной функции g. На основании помеченных сигма точек извлекается гауссиан (маленькие круги на правом верхнем графике). Что до EKF, линеаризация вызывает ошибку аппроксимации, обозначенную разницей между линеаризованной гауссовой функцией (пунктирная линия) и гауссианом, вычисленном высокоточной оценкой методом Монте-Карло (сплошная линия).

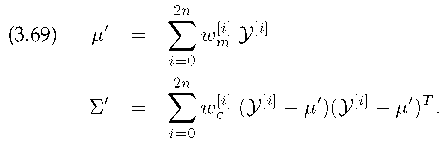
Параметр в можно подобрать таким образом, чтобы закодировать дополнительные (более высокого порядка) знания о распределении, лежащем в основе гауссова представления. Если распределение представляет собой точную гауссову функцию, в = 2 будет оптимальным выбором.

Затем сигма точки передаются функции g, проверяя, таким образом, как g меняет форму гауссиана.

image76

Параметры (p' £') результирующей гауссовой функции извлекаются из

Помеченных сигма-точек Y[i] в соответствии с

****

На Рис. 3.8 показана зависимость unscented transform от неопределенности оригинальной гауссовой функции. Для сравнения, результаты использования разложения в ряд Тейлора EKF помещены на графике рядом с результатами UKF.

На Рис. 3.9 показано дополнительное сравнение между UKF и аппроксимацией EKF, в зависимости от локальной нелинейности функции g. Как легко заметить, unscented transform дает большую точность по сравнению с разложением в ряд Тейлора первого порядка, используемое EKF. Фактически, можно показать, что

unscented transform в первых двух членах разложения Тейлора, а EKF охватывает только член первого порядка. (Однако, следует заметить, что и EKF, и UKF можно модифицировать для охвата членов высоких порядков.)

Алгоритм UKF, использующий unscented transform представлен в Таблице 3.4. Формат входных и выходных данных аналогичен алгоритму EKF. В строке 2 определены сигма точки предыдущей оценки, используя Равенство (3.66), with у short for Vn + A. Эти точки передаются через алгоритм прогнозирования в строке 3, который не учитывает шум. Прогнозируемые математическое ожидание и отклонение затем вычисляются из результирующих сигма-точек (строки 4 и 5). Rt в строке 5 добавляется к вычисленной с помощью сигма-точек ковариации с целью моделирования неопределенности дополнительного прогнозируемого шума (сравнимо со строкой 3 алгоритма EKF в Таблице 3.3). Шум прогноза Rt считается аддитивным. Позже, в Главе 7, будет представлена версия алгоритма UKF, который выполняет более точную оценку членов прогнозируемого шума и шума измерений.

Новый набор сига-точек извлекается из прогнозируемого гауссиана в строке 6. Этот набор сигма-точек Xt сейчас выражает общую неопределенность после такта прогнозирования. В строке 7 для каждой сигма-точки вычисляется прогнозируемое наблюдение. Результирующие сигма-точки наблюдения Zt используются для вычисления прогнозируемого наблюдения zt и его неопределённости, St. Матрица Qt представляет собой ковариационную матрицу аддитивного шума измерения. Заметим, что St отображает ту же неопределенность, что и Ht 21 Hf + Qt в строке 4 алгоритма EKF в

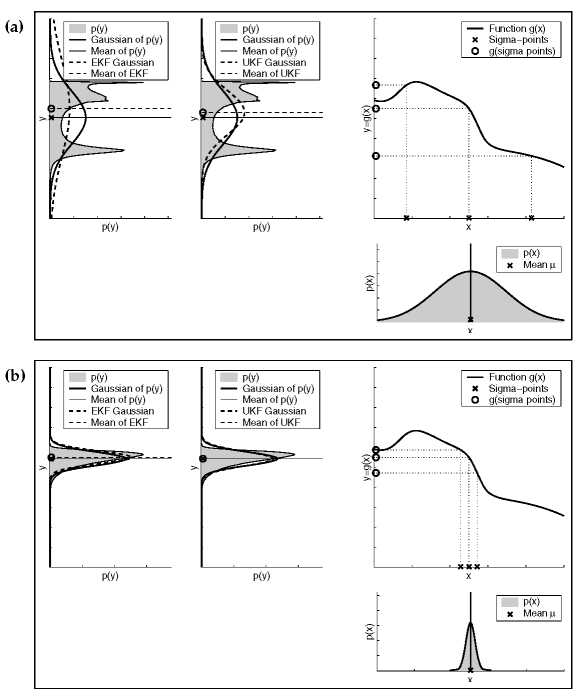


Рис. 3.8 Результаты линеаризации для UKF в зависимости от неопределенности оригинальной гауссовой функции. Результаты линеаризации EKF также показаны для сравнения (на Рис. 3.5). Unscented transform имеет меньшую ошибку аппроксимации, что видно по заметной схожести между пунктирным и сплошным графиком гауссовых функций.

В строке 10 Таблицы 3.3. определяется взаимная ковариация между состоянием и наблюдением, которая затем используется в строке 11 для вычисления усиления Калмана Kt. Взаимная ковариация Ёc соответствует члену Ёt HtT в строке 4 алгоритма EKF. Учитывая все сказанное, легко показать, что что обновление оценки, выполненное в строках 12 и 13 являются эквивалентной формой обновления, выполняемого алгоритмом EKF.

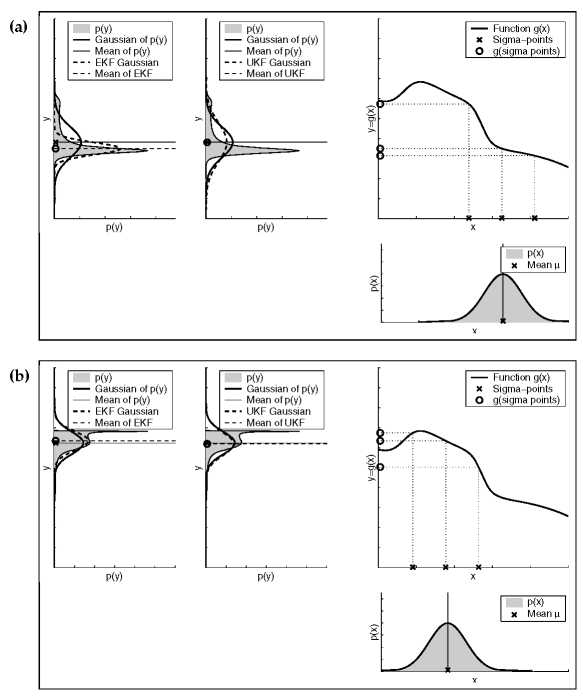
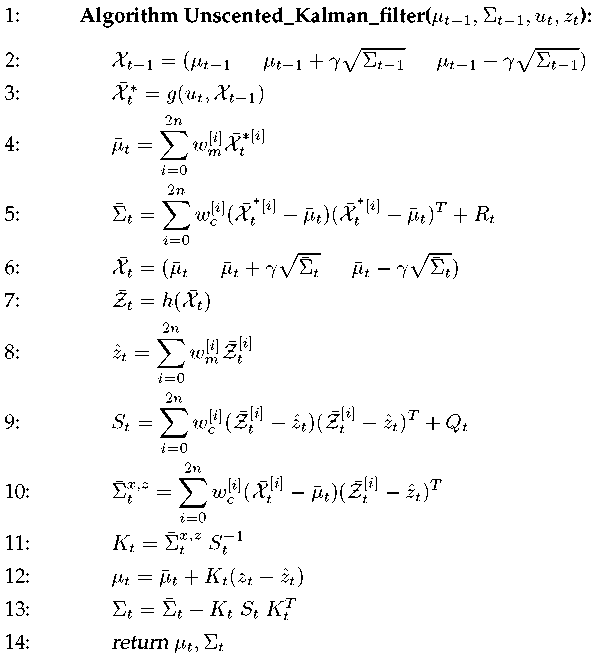


Рис. 3.9 Результаты линеаризации UKF в зависимости от математического ожидания оригинальной гауссовой функции. Результаты линеаризации EKF также показаны для сравнения (см. Рис. 3.6). Линеаризация с помощью сигма-точек что видно по заметной схожести между пунктирным и сплошным графиком гауссовых функций.

Асимптотическая сложность алгоритма UKF аналогична EKF. На практике EKF часто слегка быстрее, чем UKF. Алгоритм UKF все еще очень эффективен, даже учитывая увеличение затрат времени на постоянный коэффициент. Более того, UKF наследует преимущества unscented transform для линеаризации. Для чисто линейных систем 

будет показано, что оценки, сгенерированные с помощью UKF идентичны оценкам калмановского фильтра. Для нелинейных систем UKF дает аналогичные или лучшие результаты, чем EKF, при этом степень улучшения по отношению к EKF зависит от нелинейности и разброса неопределенности предыдущего состояния. Во многих практических реализациях разница между EKF и UKF ничтожна.

Table 3.4 The unscented Kalman filter algorithm. The variable n denotes the di­mensionality of the state vector.

Другим преимуществом UKF является факт того, что они не требуют вычисления якобианов, которые, в некоторых случаях, трудно определить. В силу этого UKF часто называют фильтром без производных.

**ФИИЛЬТР БЕЗ ПРОИЗВОДНЫХ**

Наконец, unscented transform имеет некоторое сходство с представлением на основе выборки, используемом для многочастичного фильтра, которых будет обсуждаться в следующей главе. Однако, ключевое различие состоит в том, что сигма-точки в unscented transform определены специальным образом, а в многочастичных фильтрах выборка делается случайно. Это влечет важные последствия. Если лежащее в основе распределение приближено к нормальному, тогда UKF представление значительно более эффективно по сравнению с представлением многочастичного фильтра. Если, с другой стороны, оценка очень далека от нормальной, тогда UKF представление слишком ограничено и фильтр плохо выполняет определение точек.

3.5 Информационный фильтр

Дополением калмановского фильтра является информационный фильтр (information filter- IF). Также как KF и его нелинейные версии, EKF и UKF, информационный фильтр отображает оценку в виде гауссовой функции. Поэтому стандартный информационный фильтр основан на тех же допущениях, что и фильтр Калмана. Ключевым различием между KF и IF является способ представления гауссовской оценки. Во всем семействе алгоритмов фильтров Калмана гауссовы функции представлены моментами (математическим ожиданием, ковариацией), а в информационных фильтрах – в каноническом представлении параметров, состоящем из информационной матрицы и информационного вектора. Разница в параметризации ведет к разным равенствам обновления. В частности, то, что вычислительно сложно в одном виде параметризации оказывается простым в другой (и наоборот). Каноническая параметризация и параметризация моментов часто считаются взаимодополняющими, как и IF и KF.

3.5.1 Каноническая параметризация

Каноническая параметризация многомерной гауссовой функции задается матрицей Q и вектором £. Матрица Q – это инвертированная матрица ковариации:

**КАНОНИЧЕСКАЯ ПАРАМЕТРИЗАЦИЯ**

image81

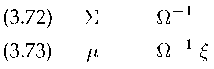
Q называется информационной матрицей, или, иногда, матрицей точности. Вектор £ называется информационным вектором. Он определяется как

**INFORMATION MATRIX**

image82

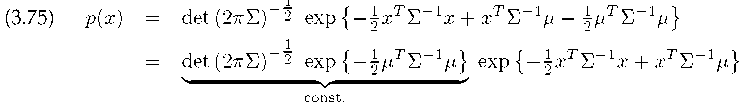
Легко заметить, что Q и £ образуют полную параметризацию гауссовой функции. В частности, математическое ожидание и ковариацию гауссовой функции можно легко получить

Из канонической параметризации путем инверсии (3.70) and (3.71):



Каноническая параметризация часто выводится путем умножения экспоненты гауссовой функции. В (3.1), определим многомерное нормальное распределение следующим образом:

image84

Очевидная последовательность преобразований ведет к следующей параметризации: ****

Член, помеченный “const.” не зависит от целевой переменной x. поэтому его можно заменить нормализующим членом η.

image86

Такая форма мотивирует использовать параметризацию в виде канонических параметров Q и £.

image87

Во многих отношениях каноническая параметризация более элегантна, чем параметризация в виде моментов. В частности, отрицательный логарифм гауссовой функции представляет собой квадратичную функцию в x, с каноническими параметрами Q и £:

image88

Здесь "const." – это константа. Читатель может заметить, что мы не можем использовать символ n для обозначения этой константы, поскольку отрицательные логарифмы вероятностей не нормализуются до 1. Отрицательный логарифм распределения p(x) квадратичен на x, его квадратичный член параметризован с помощью Q, а линейный – с помощью £. Фактически, для гауссовых функций Q должна быть неотрицательно полуопределена, отсюда — logp(x) является квадратичной функцией расстояния с математическим ожиданием p = Q-1 £. Это легко проверяется установкой первой производной (3.78) в нуль:

image89

**РАСТОЯНИЕ МАХАЛАНОБИСА**

Матрица Q определяет скорость, с которой возрастает функция расстояния в различных измерениях переменной x. Квадратичное расстояние, взвешенное матрицей Q называется расстоянием Махаланобиса.

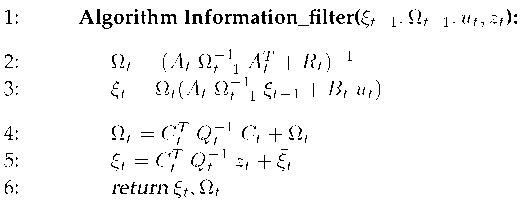


Таблица 3.5 Алгоритм информационного фильтра.

3.4.2 Алгоритм «информационного фильтра»

В Таблице 3.4 приводится модифицированный вариант алгоритма, который известен под названием «информационного фильтра». На вход подается нормальная гауссова функция в классическом представлении, заданная вектором и матрицей и определяющая оценку состояния в момент времени t-1. Аналогично всем байесовским фильтрам, входные данные включают векторы управляющего воздействия и измерения . Выходными данными фильтра являются новые значения вектора матрицы и обновленная гауссова функция, соответственно.

Этап обновления предусматривает использование матриц , , , , и , описанных в Разделе 3.2. Алгоритм информационного фильтра основан на предположении о том, что вероятности перехода состояний и измерений могут быть выражены с помощью следующих гауссовых уравнений, впервые вводимых в (3.2) и (3.5)

= + + (3.75)

= + (3.76)

В данном случае, и , – это ковариации переменных нулевого среднего зашумления, и , соответственно.

Как и классический калмановский фильтр, информационный фильтр предусматривает обновление в две фазы, экстраполяции и коррекции. Фаза экстраполяции отображена в строках 2 и 3 Таблицы 3.4. Параметры и описывают гауссовую функцию оценки по вектору состояний , после учета управляющего воздействия , но до учета измерений . Измерения принимаются во внимание в строках 4 и 5 Таблицы 3.4, где происходит обновление значений параметров оценки на основе измерений в момент времени t.

Эти две фазы обновления могут очень сильно различаться по сложности, особенно для случая, когда пространство состояний имеет большое число измерений. Как показано в Таблице 3.4, во время фазы экстраполяции происходит инверсия двух матриц размера n\*n, где n – размерность пространства состояний. Инверсия матриц, по текущим оценкам, требует затрат времени, приблизительно равных O(). Но для фильтров Калмана обновление имеет аддитивный характер и требует, максимум O(). Требуемое время еще более уменьшается, если набор переменных зависит только от управляющего воздействия или если перенос переменных выполняется независимо друг от друга. Для информационного фильтра ситуация обратная, поскольку в нем аддитивно происходит обновление измерений. Поэтому максимальные затраты времени также составят O(), а эффективность может быть более улучшена для случая, когда измерения содержат только данные о наборе всех переменных состояния для конкретного момента времени. Обновление измерений в калмановских фильтрах является очень затратным этапом, поскольку требует инверсии матриц, и вычислительная сложность, в худшем случае, достигает O(). Это хорошо иллюстрирует двойственную природу калмановских и информационных фильтров.

3.5.3 Математический вывод алгоритма информационного фильтра

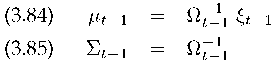
Вывод для информационного фильтра аналогичен выводу для фильтра Калмана.

Для математического вывода для такта прогнозирования (строки 2 и 3 в Таблице 3.5), начнем с соответствующих равенств обновления для калмановских фильтров которые можно найти в строках 2 и 3 алгоритма в Таблице 3.1 и повторно приводятся здесь для удобства читателя:

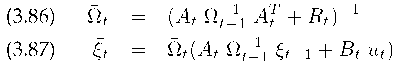
ТАКТ ПРОГНОЗИРОВАНИЯ

image92

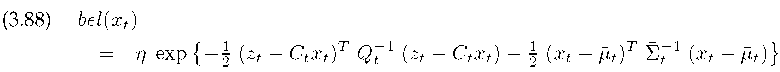
Далее в такте прогнозирования информационного фильтра производится замена моментов p и E каноническими параметрами £ и Q, в соответствии с их определениями в (3.72) и (3.73):



Подставка этих выражений в (3.82) и (3.83) дает нам набор равенств для прогнозирования



Эти равенства идентичны приведенным в Таблице 3.5. Легко заметить, что такт прогнозирования включает две вложенные инверсии потенциально больших матриц. Эти вложенные инверсии можно избежать, если только небольшое число переменных состояния затрагивается обновлением этапа движения, что мы обсудим чуть позже.

Вывод для обновления измерения еще проще. Начнем с гауссовой функции оценки в момент t, которая была приведена в Равенстве (3.35) и повторим ее еще раз здесь: 

**ОБНОВЛЕНИЕ ИЗМЕРЕНИЯ**

Для гауссовой функции, представленной в канонической форме, это распределение задается следующим образом

image96

которое, будучи переписанным в виде экспоненты, разрешается как

**image97**

Сейчас можем окончательно сформулировать равенства обновления измерений, собрав все члены в квадратные скобки:

image98

Эти равенства идентичны обновлению измерений в строках 4 и 5 Таблицы 3.5.

3.5.4 Алгоритм расширенного информационного фильтра

Расширенный информационный фильтра (extended information – EIF), расширяет информационный фильтр для нелинейного случая, практически таким же образом, как EKF представляет собой нелинейное расширение фильтра Калмана. Алгоритм EIF приведен в Таблице 3.6. Прогнозирование выполняется в строках с 2 по 4, а обновление измерения - с 5 по 7. Эти уравнения такта обновления, по большей части, аналогичны линейному информационному фильтру, где функции g и h (и их якобианы Gt и Ht) заменяют параметры линейной модели At, Bt, и Ct. Как и прежде, g и h обозначают нелинейную функцию перехода состояния и функцию измерения, соответственно. Это было определено в (3.48) and (3.49) и повторено ниже:

image99

К сожалению, обе функции, и g и h, требуют на вход состояние. Это требует восстановления оценки состояния g из канонических параметров. Восстановление

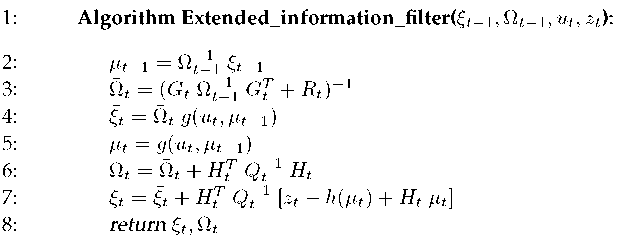


Таблица 3.6 Алгоритм расширенного информационного фильтра (EIF).

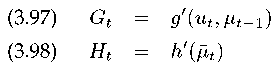
выполняется в строке 2, в которой состояние ut-1 очевидным образом вычисляется из Qt-1 и £t-i. В строке 5 вычисляется состояние p,t, используя уже знакомое по EKF выражение (строка 2 в Таблице 3.3). Кажется, что необходимость восстанавливать оценку состояния противоречит желанию представить фильтр, используя канонические параметры. Мы вернемся к этой теме при обсуждении использования расширенных информационных фильтров в контексте построения карт для роботов.

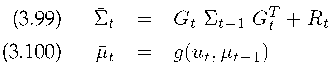
3.5.5 Математический вывод расширенного информационного фильтра

Расширенный информационный фильтр легко выводится используя, в основном, ту же самую линеаризацию, которая была ранее использована для расширенных фильтров Калмана. Как и в (3.51) и (3.53), в расширенном информационном фильтре с помощью разложения в ряд Тейлора оценивается g и h:

image101

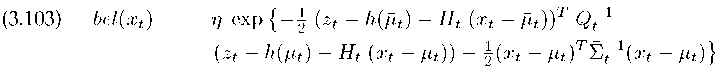
Here Gt and Ht are the Jacobians of g and h at gt-1 and gt, respectively:



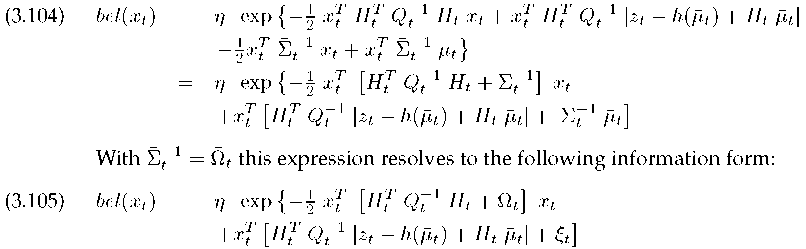
Эти определения аналогичны используемым в EKF. Этап прогнозирования выводится из строк 2 и 3 алгоритма EKF (Таблица 3.3), которые приведены ниже: ****

Замена Σt-1 на Ω- t-11 и μ¯t на Ω¯ - t 1ξ¯t дает уравнения прогнозирования для расширенного информационного фильтра:

image104

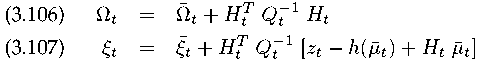
Обновление измерения выводится из уравнений (3.60) и (3.61). В частности, (3.61) определяет следующую апостериорную гауссову функцию: ****

Умножение экспоненты и переопределение членов дает следующее выражение для апостериорной вероятности:



С Σ¯ -1 t = Ω¯ t это выражение разрешается до следующего информационного вида:

Сейчас можно сформулировать уравнения обновления измерения, собрав все члены в квадратных скобках:

****

3.5.6 Практические соображения

В контексте задач робототехники информационный фильтр имеет несколько преимуществ по сравнению с фильтром Калмана. Например, глобальное представление неопределенности в информационном фильтре достаточно простое: просто представить, что Q = 0. При использовании моментов, такие величины общей неопределенности дают ковариацию бесконечной степени. Это особенно проблематично, когда измерения датчиков несут информацию о жестком наборе всех переменных состояния, что часто встречается в робототехнике. Для решения таких ситуаций средствами EKF необходимо предпринимать специальные действия. Информационный фильтр обычно ведет себя более стабильно количественно, чем калмановский фильтр во многих реализациях, которые будут обсуждаться позже.

Как мы должны увидеть в более поздних главах книги, информационный фильтр и несколько дополнений позволяют роботу интегрировать информацию без немедленного разрешения ее в виде вероятностей. Это может стать решающим преимуществом для сложных задач оценки, включающих сотни или даже миллионы переменных. Для таких больших задач попытки интеграции по типу фильтра Калмана вызывают серьезные вычислительные проблемы, поскольку для каждой новой порции информации требуется распространение ее через большую систему переменных. Для информационного фильтра с соответствующими модификациями возможно обойти эту проблему, просто добавив эту информацию в систему локально. Однако, это свойство нехарактерно для простого информационного фильтра, обсуждаемого сейчас. В Главе 12 мы расширим этот фильтр.

Другое преимущество информационного фильтра перед фильтром Калмана возникает в силу естественной пригодности для решения проблем нескольких роботов. Задачи нескольких роботов часто включают интеграцию данных датчиков, которые собираются децентрализовано. Такая интеграция обычно выполняется с помощью теоремы Байеса. Будучи представлена в логарифмической форме, теорема Байеса становится суммой. Как было сказано выше, канонические параметры информационных фильтров выражают вероятность в логарифмической форме. Поэтому интеграция информации достигается суммированием данных с нескольких роботов. Сложение коммутативно. В силу этого информационные фильтры часто могут интегрировать информацию в произвольном порядке, с произвольными задержками и в полностью децентрализованной манере. Хотя то же самое возможно сделать, используя параметризацию в виде моментов – в конце концов, они выражают одну и ту же информацию – необходимые дополнительные усилия намного больше. Несмотря на это преимущество, применение информационных фильтров в системах из нескольких роботов остается, по большей части, не исследованным. Мы вернемся к теме систем из нескольких роботов в Главе 12.

Обратной стороной преимуществ информационного фильтра являются важные ограничения. Основным недостатком расширенного информационного фильтра является необходимость восстанавливать оценку состояния на такте обновления в случае нелинейных систем. Этот шаг, если применять его указанным выше образом, требует инверсии информационной матрицы. На этапе прогнозирования информационные фильтры требуют дополнительных операций инверсии матриц. Для многих задач робототехники EKF не требует инверсии матриц сравнимого размера. Считается, что для пространств состояний с большим числом измерений информационный фильтр проигрывает фильтру Калмана в смысле вычислительной сложности. Фактически, это является одной из причин подавляющей популярности фильтров Калмана по сравнению с информационными фильтрами.

Как будет позже показано в книге, эти ограничения не обязательно применимы для задач, в которых информационная матрица имеет структуру. Во многих задачах робототехники взаимодействие переменных состояния локально. В результате, информационная матрица разрежена. Эта разреженность не соотносится с разреженностью матрицы ковариации.

Можно считать, что информационные фильтры – это графы, в которых состояния соединены, когда соответствующие элементы вне главной диагонали информационной матрицы не равны нулю. Разреженные информационные матрицы соответствуют разреженным графам. Фактически, такие графы широко известны как марковские случайные поля. Для таких полей имеется уйма алгоритмов, эффективно выполняющих операции общей оценки и обновления и имеющих названия наподобие «алгоритм распространения доверия» (loopy belief propagation). В ходе издожения мы столкнемся с проблемой картографирования, в которой информационная матрица (примерно) разрежена, и разработаем расширенный информационный фильтр, который будет существенно более эффективен как по сравнению с фильтрами Калмана, как и информационными фильтрами с неразреженными матрицами.

**Markov random field**

3.6 Вывод

В этом разделе мы представили эффективные алгоритмы байесовских фильтров, которые выражают апостериорные оценки в виде многомерных гауссовых функций. Мы отметили, что гауссовы функции могут быть представлены двумя различными способами: параметризацией в виде моментов и канонической параметризацией. Параметризация в виде моментов состоит из математического ожидания (первый момент) и ковариации (второй момент) функции нормального распределения. Каноническая или обычная параметризация состоит из информационной матрицы и информационного вектора. Оба вида параметризации дополняют друг друга, и каждую можно восстановить из другой, используя инверсию матриц.

Для обоих видов параметризации можно применять байесовские фильтры. При использовании параметризации в виде моментов получившийся фильтр называется фильтром Калмана. Дополнением фильтра Калмана является информационный фильтр, который представляет апостериорную вероятность в канонической параметризации. Обновление фильтра Калмана на основе управляющего сигнала вычислительно просто, в то время, как учет измерения более труден. Для информационного фильтра ситуация обратная, учет измерения достаточно прост, но вычислительно сложно обновить фильтр на основе управляющего сигнала.

Для вычисления правильной апостериорной вероятности в обоих фильтрах требуется соблюсти три условия. Во-первых, начальная оценка должна быть нормальным распределением. Во-вторых, вероятность перехода состояния должна выражаться функцией с линейным аргументом с добавлением независимого гауссовского шума. В третьих, то же справедливо для вероятности измерения: она должна иметь линейный аргумент с добавленным гауссовским шумом. Системы, которые отвечают этим требованиям называются линейными гауссовыми системами.

Оба фильтра можно расширить для нелинейных задач. Один из методов, описанных в этой главе, вычисляет касательную к нелинейной функции. Касательные линейны, что позволяет использовать это в обоих фильтрах. Способ нахождения касательной называется разложением в ряд Тейлора. Разложение в ряд Тейлора включает вычисление первой производной целевой функции и оценка ее в определенной точке. Результатом этой операции является матрица, известная как якобиан. Получившийся в результате фильтры называются «расширенными».

Так называемый unscented Kalman filter использует другой метод линеаризации, называемый «unscented transform». Он берет значения функции, которую необходимо линеаризовать, в нескольких выбранных точках и вычисляет линеаризованную аппроксимацию на основе полученных результатов. Этот фильтр может быть использован без якобианов, и, в силу этого, именуется «фильтром без производных». Un­scented Kalman filter эквивалентен фильтру Калмана для линейных систем, но часто дает лучшие оценки в нелинейных системах. Вычислительная сложность этого фильтра та же самая, что и для расширенного фильтра Калмана.

Точность разложений в ряд Тейлора и unscented transform зависит от двух факторов: степени нелинейности системы и ширина апостериорной вероятности. Расширенные фильтры обычно показывают хорошие результаты, если состояние системы известно с относительно высокой точностью, и оставшаяся ковариация мала. Чем больше неопределенность, чем больше ошибка, вносимая линеаризацией.

Одним из главных преимуществ гауссовых фильтров является вычислительная эффективность: обновление требует времени кратного размерности пространства состояний. Это неприменимо для некоторых методов, описанных в следующей главе. Основным недостатком является ограниченность одномодальными гауссовыми распределениями. •

Расширение гауссовых функций для многомодальных апостериорных вероятностей известно под названием фильтра Калмана с несколькими гипотезами. Этот фильтр отображает апостериорную вероятность в виде смеси гауссовых функций, которые представляют собой не что иное, как взвешенную сумму гауссовых функций.

Механика обновления этого фильтра требует механизмов для разделения и слияния или отсечения отдельных гауссианов. Фильтры Калмана с несколькими гипотезами хорошо подходят, в частности, для проблем с дискретной интеграцией данных, что обычно и происходит в робототехнике.

• В режиме многовариантных гауссовых функций оба фильтра, и фильтр Калмана, и информационный фильтр, имеют ортогональные сильные и слабые стороны. Однако, калмановский фильтр и его нелинейное расширение, расширенный калмановский фильтр, значительно более популярны, чем информационный фильтр.

Выбор материалов для этой главы основывается на самых популярных, на сегодняшний день, методах робототехники. Существует огромное количество вариантов и расширений гауссовых фильтров, которые нацелены на обход различных ограничений и недостатков конкретных фильтров.

Большое число алгоритмов в этой книге основано на гауссовых фильтрах. Множество практических проблем робототехники требует использования расширений, которые используют разреженные структуры или факторизацию апостериорной вероятности.

Библиографические примечания

Фильтры Калмана были отрыты Сверлингом (1958) и Калманом (1960). Обычно, они представляются как оптимальная функция оценки методом наименьших квадратов, и менее часто – в качестве метода вычисления апостериорных распределений, хотя, при соответствующих допущениях, обе точки зрения идентичны. Есть большое число великолепных учебников по калмановским и информационным фильтрам, включая работу Мэйбека (Maybeck, 1990) и Язвински (Jazwinsky -1970). На сегодняшний день вклад в калмановские фильтры с включением данных был внесен в работах Бар-Шалома и Фортманна (Bar-Shalom и Fortmann,1988), Бар-Шалома и Ли (Bar- Shalom и Li, 1998).

Лемма об обращении матриц можно найти в работе Голуба и Лоана (Golub и Loan,1986). Инверсия матриц может быть выполнена за время O(n2-376), согласно Копперсмиту и Винограду (Coppersmith, Winograd,1990). Самым заметным результатом в серии работ стало улучшение в области вычислительной сложности более чем O(n3) для алгоритма отбора переменных. Серия началась с основополагающей работы Страссена (Strassen, 1969), в которой он сформулировал алгоритм, требующий ng O(n2 807). Кавер и Томас (Cover и Thomas,1991) выполнили обзор теории информации, но с акцентом на дискретные системы. Unscented Kalman filter был создан Джулиером и Ульманном (Julier и Uhlmann, 1997). Сравнение UKF с EKF в контексте различных задач оценки состояния можно найти в работе ван де Мерве (van der Merwe, 2004). Минка (Minka, 2001) выполнил работу по подбору моментов moments matching и фильтрации допустимых плотностей гауссовых смесей.

Упражнения

В этом и последующих упражнениях требуется разработать фильтр Калмана для простой динамической системы: автомобиль с линейной динамикой двигается в линейной окружающей среде. Для простоты допустим, что At =1. Положение автомобиля в момент t задано xt. Его скорость xt, а ускорение - xt. Допустим, ускорение задается случайно в каждый момент времени, согласно гауссовой функции с нулевым математическим ожиданием и ковариацией a2 = 1.

Каков минимальный вектор состояний для фильтра Калмана (и какова будет результирующая система марковской)?

Для вашего вектора состояний найдите вероятность перехода состояния p(xt | ut,xt\_i). Подсказка: эта функция перехода будет состоять из линейных матриц A и B и ковариации зашумления R (см Равенство (3.4) и Таблицу 3.1).

Реализуйте такт прогнозирования состояния для фильтра. Допустим, мы знаем, что в момент времени t = 0, x0 = x0 = x0 =0. Вычислить распределение состояний для моментов времени t = 1, 2,..., 5.

Для каждого значения t, изобразить точками на графике апостериорное распределение по x и x, где x – горизонтальная, а x – вертикальная ось. Для каждой апостериорной вероятности, необходимо изобразить эллипс неопределенности, который представляет собой эллипс точек одного стандартного отклонения от среднего. Подсказка: Если нет доступа к математической библиотеке, можно найти эллипсы путем анализа характеристического значения матрицы ковариации.

**Эллипс неопределенности**

Что произойдет с корреляцией между xt и xt при 11 ж?

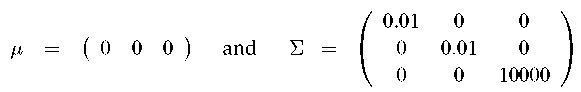
Сейчас добавим измерения к нашему фильтру Калмана. Допустим, в момент времени t, возможно полчуить зашумленное наблюдение x. Предполагается, что датчик измеряет истинное местоположение. Однако, это измерение искажено гауссовским шумом с ковариацией a2 = 10.

Определить модель измерений. Подсказка: Необходимо определить матрицу C и еще одну матрицу Q (см Равенство (3.6) и Таблицу 3.1).

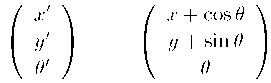
Реализовать обновление измерения. Допустим, в момент времени t = 5, мы наблюдаем значение измерения z = 5. Укажите параметры гауссовой функции до и после обновления KF. Изобразить на графике эллипс неопределенности перед и после учета измерения (в инструкции выше указано, каким образом изобразить эллипс неопределенности).

В Главе 3.2.4 был выведен такт прогнозирования в KF. Этот такт часто выводят через Z –преобразования или преобразования Фурье, используя теорему свертки. Заново вывести такт прогнозирования, используя преобразования. Примечание: Это упражнение требует знания трансформаций и свертки, что выходит за пределы материала этой книги.

4. В тексте было отмечено, что линеаризация EKF – это приближение. Для того, чтобы выяснить, насколько эта аппроксимация плоха, выполните упражнение. Допустим, имеется мобильный робот, действующий в плоской среде. Его состояние – координаты его местоположения x-y и глобальное направление ориентации в. Допустим, с высокой степенью определенности известны x и y, но неизвестна ориентация. Это отражается следующей первоначальной оценкой



Графически изобразить лучшую модель апостериорной вероятности положения робота после передвижения робота на d =1 единиц вперед. Для этого упражнения, предположим, что робот движется безукоризненно и без зашумления. Отсюда, прогнозируемое положение робота после движения будет



Для графика можно игнорировать в и изобразить только апостериорную вероятность в координатах x-y.

Теперь развейте это движение до такта обновления для EKF. Для этого необходимо определить функцию перехода состояний и линеаризовать ее. Затем необходимо сгенерировать новую гауссову оценку положения робота, используя линеаризованную модель. Необходимо дать точные математические равенства для каждого шага и оценить результирующий гауссиан.

Изобразить эллипс неопределенности гауссовой функции и сравнить ее с интуитивным решением.

А сейчас добавим измерение. Нашим измерением будет зашумленная проекция координаты x робота с ковариацией Q = 0.01. Необходимо определить модель измерения. Примените измерение к интуитивному апостериорному распределению и формально к оценке EKF, используя стандартные механики расширенных фильтров Калмана. Дайте точный результат для EKF и сравните его с результатом интуитивного анализа.

Объясните разницу между вашей оценкой апостериорной вероятности и гауссовой функцией, получившейся в результате работы EKF. Насколько существенны эти различия? Что можно изменить, чтобы сделать эти приближения более точными? Что произойдет, если будет известна начальная ориентация, а не координата y робота?

В фильтре Калмана в Таблице 3.1 отсутствует постоянный дополнительный член в моделях движения и измерения. Расширьте алгоритм так, чтобы учитывать эти члены.

Докажите (на примере) существование разреженной информационной матрицы в многомерных гауссовых распределениях (размерности d), которые связывают все d переменных с коэффициентом корреляции e-близким к 1. Назовем информационную матрицу разреженной, если все, кроме небольшого постоянного количества элементов каждой строки и столбца, равны нулю.