

칼만 필터, 불확실성 속에서 최적의 해 찾기: 3x3 행렬 예제부터 EKF까지

서론: 불확실한 세상에서의 추정 문제

우리가 사는 세상은 불확실성으로 가득 차 있습니다. 자율주행 자동차가 자신의 정확한 위치를 파악하려 할 때, 로켓이 달을 향해 날아갈 때, 혹은 경제학자가 다음 분기 성장률을 예측할 때, 우리는 완벽한 정보를 가질 수 없습니다.¹ 시스템의 상태를 예측하는 수학적 모델은 현실의 모든 변수를 담지 못해 오차가 누적되고, 상태를 측정하는 센서는 언제나 노이즈라는 불청객을 동반합니다.³

이러한 상황에서 우리는 두 가지 불완전한 정보, 즉 '모델에 기반한 예측'과 '센서로부터의 측정'을 마주하게 됩니다. 예측은 시간이 지남에 따라 현실과 멀어지는 경향(drift)이 있고, 측정은 매 순간 무작위적인 오차를 포함합니다.³ 그렇다면 이 둘 중 무엇을 믿어야 할까요? 혹은, 둘을 어떻게 조합해야 가장 정확한 '진실'에 가까운 값을 얻을 수 있을까요?

이 근본적인 질문에 대한 가장 우아하고 강력한 해답 중 하나가 바로 **칼만 필터(Kalman Filter)**입니다.⁵ 칼만 필터는 단순히 두 값을 평균 내는 것이 아니라, 각 정보가 얼마나 '확실한지'를 정량적으로 따져서 최적의 가중치를 동적으로 계산하고, 이를 통해 두 정보의 장점만을 취합하여 어떤 단일 정보보다도 더 정확한 추정치를 만들어내는 재귀적(recursive) 알고리즘입니다.³ 이 보고서는 칼만 필터가 어떤 철학과 수학적 원리를 통해 이 놀라운 작업을 수행하는지, 대학교 1학년 수준의 기초부터 시작하여 가장 본질적인 '왜?'라는 질문에 답하며 차근차근 탐구해 나갈 것입니다.

Part I: 추정의 철학적, 수학적 기초

칼만 필터의 복잡한 행렬 수식을 마주하기 전에, 이 필터가 탄생하게 된 근본적인 이유와 그 작동 원리의 핵심 아이디어를 이해하는 것이 중요합니다. 이 장에서는 칼만 필터를 떠받치는 두 개의 기둥, '불확실성의 정량화'와 '정보의 융합'에 대해 알아봅니다.

섹션 1: 불확실한 세상에서 추정의 필요성

칼만 필터가 해결하고자 하는 핵심 문제는 명확합니다: **측정 불가능하거나, 측정하더라도 노이즈가 섞인 시스템의 상태를 어떻게 가장 정확하게 알아낼 것인가?**⁴

여기에 우리가 이 보고서 전반에 걸쳐 사용할 물리적 시나리오가 있습니다: 직선 레일 위를 움직이는 작은 자율주행 카트의 상태(위치, 속도 등)를 추적하는 것입니다. 이 카트의 상태를 알기 위해 우리는 두 가지 종류의 정보를 활용할 수 있습니다.

- 예측 모델 (내부적 믿음):** 우리는 뉴턴의 운동 법칙과 같은 물리 모델을 통해 카트가 다음 순간에 어디에 있을지를 예측할 수 있습니다. 이는 우리의 '사전 믿음(prior belief)'에 해당합니다.⁸ 하지만

이 모델은 완벽하지 않습니다. 예측하지 못한 마찰력의 변화나 미세한 공기 저항 같은 요인들, 즉 **프로세스 노이즈(process noise)** 때문에 모델의 예측값은 시간이 지날수록 실제 카트의 상태와 점차 멀어지게 됩니다.³

2. **센서 측정 (외부적 증거):** 카트에 장착된 GPS나 카메라 같은 센서를 통해 현재 위치를 직접 측정할 수 있습니다. 이는 우리의 '새로운 증거(new evidence)'에 해당합니다.⁸ 하지만 센서 역시 완벽하지 않습니다. 모든 센서는 내재적인 한계나 외부 환경의 영향으로 인해 무작위적인 오차, 즉 **측정 노이즈(measurement noise)**를 포함합니다.⁴

결국 우리는 '점점 부정확해지는 예측'과 '매번 부정확한 측정'이라는 두 개의 불완전한 정보만을 갖게 됩니다. 칼만 필터의 역할은 이 두 정보를 지능적으로 융합하여 각각의 정보보다 더 신뢰할 수 있는 단 하나의 추정치를 만들어내는 것입니다.³

여기서 칼만 필터가 '최적(optimal)'이라는 수식어가 붙는 이유가 중요합니다. 이는 단순히 좋은 성능을 낸다는 의미를 넘어, 특정 조건(선형 시스템, 가우시안 노이즈) 하에서 추정 오차의 평균 제곱(mean squared error)을 최소화하는, 수학적으로 가장 뛰어난 추정기임을 의미합니다.³ 이러한 최적성 덕분에 아폴로 달 탐사 프로젝트의 우주선 궤도 추정이라는 중차대한 임무에 사용될 수 있었고³, 오늘날 로봇 공학, 자율주행, 금융 등 다양한 첨단 분야의 핵심 기술로 자리 잡게 된 것입니다.¹

섹션 2: 불확실성의 언어: 가우시안 분포

불확실한 예측과 불확실한 측정을 수학적으로 다루려면, 먼저 '불확실성'을 표현할 언어가 필요합니다. 칼만 필터는 이 언어로 **가우시안 분포(Gaussian Distribution)**, 즉 정규 분포를 사용합니다.

우리는 카트의 위치를 '숫자 하나'로 단정하는 대신, 확률 분포로 표현합니다.

- **평균(μ):** 우리가 생각하는 가장 가능성 높은 값(best guess)입니다. 가우시안 분포의 종 모양 그래프에서 가장 높은 지점, 즉 중심에 해당합니다.¹¹
- **분산(σ^2):** 우리의 불확실성의 정도를 나타냅니다. 분산이 작으면 그래프는 키가 크고 뾰족한 모양이 되며, 이는 우리가 추정치에 대해 높은 확신을 가지고 있음을 의미합니다. 반대로 분산이 크면 그래프는 키가 작고 평평한 모양이 되며, 이는 낮은 확신, 즉 높은 불확실성을 의미합니다.⁶

그렇다면 왜 수많은 확률 분포 중에서 하필 가우시안 분포를 가정하는 것일까요? 이는 단순히 계산의 편의성을 넘어서 깊은 수학적, 철학적 이유를 가집니다.

첫째, **중심 극한 정리(Central Limit Theorem)**에 따르면, 서로 독립적인 수많은 무작위적 요인들의 합은 그 요인들이 원래 어떤 분포를 따랐는지에 관계없이 가우시안 분포에 가까워지는 경향이 있습니다.¹¹ 센서 노이즈나 프로세스 노이즈는 보통 수많은 미세하고 독립적인 물리적 현상들의 총합으로 발생하기 때문에, 이들을 가우시안 분포로 가정하는 것은 매우 합리적인 물리적 근사입니다.

둘째, 더 심오한 이유는 **최대 엔트로피 원리(Principle of Maximum Entropy)**에 있습니다. 정보 이론에서 엔트로피는 불확실성의 척도로 사용됩니다.¹¹ 주어진 평균과 분산 값에 대해, 가우시안 분포는 가능한 모든 확률 분포 중에서 엔트로피를 최대로 만드는, 즉 가장 불확실성이 큰 분포입니다.¹² 따라서 우리가 아는 정보(평균과 분산) 외에 어떠한 추가적인 가정도 하지 않고, 가능한 최대의 불확실성을 모델에

내포시키는 가장 '정직하고' '보수적인' 선택이 바로 가우시안 분포를 가정하는 것입니다. 이는 필터의 강건성(robustness)을 높여주는 중요한 특성입니다.

섹션 3: 융합의 본질: 두 가우시안의 만남

이제 칼만 필터의 핵심 마법, 즉 두 개의 불확실한 정보를 융합하여 어떻게 더 확실한 정보 하나를 만들어나는지 살펴보겠습니다. 결론부터 말하자면, 두 개의 가우시안 분포를 곱하는 것입니다.

1D 공간에서 카트의 위치를 추정하는 과정을 단계별로 따라가 보겠습니다.

- 예측 (Prediction):** 이전 단계의 정보를 바탕으로, 우리는 현재 카트의 위치에 대한 '사전(prior)' 믿음을 가집니다. 이 믿음은 평균이 μ_{prior} 이고 분산이 σ_{prior}^2 인 가우시안 분포로 표현됩니다. 일반적으로 예측 단계는 불확실성을 증가시키므로, 이 분포는 비교적 펴져있는 모양일 것입니다.⁸
- 측정 (Measurement):** 이제 센서가 카트의 위치를 측정합니다. 이 측정값 역시 불확실성을 가지며, 평균이 $\mu_{measure}$ 이고 분산이 $\sigma_{measure}^2$ 인 또 다른 가우시안 분포로 표현할 수 있습니다.¹³
- 융합 (Fusion):** 가장 가능성 높은 실제 위치는 우리의 '믿음'과 '증거'가 모두 높은 확률을 가리키는 지점일 것입니다. 수학적으로 이는 두 가우시안 분포의 확률 밀도 함수를 곱하는 것과 같습니다. 놀랍게도, **두 가우시안 분포의 곱은 언제나 또 다른 새로운 가우시안 분포가 됩니다.**¹⁴
- 결과 (Update):** 이렇게 새로 얻어진 '사후(posterior)' 가우시안 분포는 두 가지 중요한 특징을 가집니다. 첫째, 새로운 평균은 원래의 두 평균($\mu_{prior}, \mu_{measure}$) 사이 어딘가에 위치합니다. 둘째, 그리고 가장 중요한 것은, **새로운 분산은 원래의 두 분산($\sigma_{prior}^2, \sigma_{measure}^2$)보다 항상 작아집니다.** 즉, 우리는 정보를 융합함으로써 카트의 위치에 대해 이전보다 더 확신하게 된 것입니다. 이것이 바로 융합이 불확실성을 줄이는 원리에 대한 수학적 증명입니다.

이 융합의 결과로 얻어지는 새로운 평균(업데이트된 추정치)은 다음과 같은 간단한 가중 평균 형태로 표현될 수 있으며, 여기서 우리는 칼만 이득(Kalman Gain)의 본질을 처음으로 엿볼 수 있습니다.

$$\hat{x}_{new} = \hat{x}_{prior} + K \cdot (z - \hat{x}_{prior})$$

여기서 \hat{x}_{prior} 는 예측값, z 는 측정값, 그리고 $(z - \hat{x}_{prior})$ 는 예측과 측정의 차이, 즉 '놀라움(surprise)' 또는 '혁신(innovation)'이라 불립니다.⁶

K 가 바로 **칼만 이득(Kalman Gain)**이며, 이 '놀라움'을 얼마나 반영할지를 결정하는 가중치입니다.

1D 스칼라의 경우, 이 칼만 이득 K 는 각 정보의 불확실성(분산)으로부터 직접 유도됩니다⁸:

$$K = \frac{\sigma_{prior}^2}{\sigma_{prior}^2 + \sigma_{measure}^2}$$

이 식은 칼만 이득의 본질을 명확하게 보여줍니다.

- 만약 우리의 예측이 매우 불확실하다면(σ_{prior}^2 가 매우 크다면), K 는 1에 가까워집니다. 이 경우 업데이트 식은 $\hat{x}_{new} \approx z$ 가 되어, 불확실한 예측을 버리고 새로운 측정값을 거의 그대로 신뢰하게 됩니다.

- 반대로, 만약 센서 측정이 매우 불확실하다면($\sigma_{measure}^2$ 가 매우 크다면), K 는 0에 가까워집니다. 이 경우 업데이트 식은 $\hat{x}_{new} \approx \hat{x}_{prior}$ 가 되어, 노이즈가 심한 측정값을 무시하고 기존의 예측을 고수하게 됩니다.

결론적으로, 칼만 이득은 두 정보원(예측과 측정)의 불확실성을 비교하여 동적으로 계산되는 '신뢰도 비율'입니다. 이는 최종 추정치의 불확실성(분산)을 최소화하는 최적의 가중치이며, 확률적 믿음을 결합하는 수학적 원리로부터 자연스럽게 도출됩니다.³

Part II: 선형 칼만 필터의 실제: 3x3 행렬 예제

이제 1차원에서의 개념적 이해를 바탕으로, 더 현실적인 다차원 문제로 확장해 보겠습니다. 복잡한 시스템을 다루기 위해 벡터와 행렬이라는 선형대수의 강력한 도구를 도입하고, 이를 통해 칼만 필터의 전체 알고리즘을 구체적으로 살펴보겠습니다.

섹션 4: 현실을 방정식으로 담다: 상태 공간 모델

칼만 필터를 적용하기 위한 첫 단계는 우리가 다루고자 하는 물리 시스템을 표준화된 행렬 방정식의 집합, 즉 **상태 공간 모델(State-Space Model)**로 표현하는 것입니다.¹

상태 벡터(State Vector, x) 정의하기

우리의 목표는 카트의 위치뿐만 아니라, 그와 연관된 속도와 가속도까지 함께 추적하는 것입니다. 이 세 가지 변수는 시스템의 동적 상태를 완벽하게 기술합니다. 따라서 우리는 다음과 같은 3×1 크기의 상태 벡터 x 를 정의합니다.

$$x = [\text{위치}(p) \text{ 속도}(v) \text{ 가속도}(a)]$$

이 벡터 x 는 특정 시점에서의 시스템 상태에 대한 완전한 요약입니다.¹⁶ 여기서 우리는 센서가 위치만 측정할 수 있고, 속도와 가속도는 직접 측정할 수 없다고 가정합니다. 이처럼 직접 관측되지 않는 변수들을 '숨겨진 상태(hidden states)'라고 부르며, 칼만 필터는 이러한 숨겨진 상태들까지도 훌륭하게 추정해 냅니다.¹⁷

프로세스 모델 (상태 방정식)

다음으로, 현재 상태 x_{k-1} 을 기반으로 다음 시점의 상태 x_k 를 예측하는 방정식이 필요합니다. 이는 시스템의 물리 법칙을 나타냅니다. 고등학교 물리 시간에 배운 등가속도 운동 공식을 사용합니다.

- $p_k = p_{k-1} + v_{k-1} \cdot \Delta t + \frac{1}{2} a_{k-1} \cdot (\Delta t)^2$
- $v_k = v_{k-1} + a_{k-1} \cdot \Delta t$
- $a_k = a_{k-1}$ (이것이 바로 '등가속도(constant acceleration)' 모델의 핵심 가정입니다.)

상태 전이 행렬(State Transition Matrix, F) 유도하기

위의 세 스칼라 방정식을 단 하나의 행렬 방정식 $\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1}$ 로 표현할 수 있습니다. 각 행을 주의 깊게 살펴보면 다음과 같이 3x3 크기의 **상태 전이 행렬 F**를 유도할 수 있습니다.¹⁸

$$\begin{bmatrix} p_k & v_k & a_k \end{bmatrix} = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}(\Delta t)^2 & 0 & 1 & \Delta t & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{k-1} & v_{k-1} & a_{k-1} \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}(\Delta t)^2 & 0 & 1 & \Delta t & 0 & 0 & 1 \end{bmatrix}$$

이 행렬 \mathbf{F} 는 우리가 가정한 시스템의 운동 모델(물리 법칙)을 우아하게 압축하여 담고 있습니다.

측정 모델 (측정 방정식)

마지막으로, 우리가 직접 관측할 수 없는 '상태' \mathbf{x}_k 와 실제 센서 '측정값' \mathbf{z}_k 를 연결하는 방정식이 필요합니다.

우리의 센서는 위치만 측정할 수 있다고 가정했으므로, 측정값 z_k 는 실제 위치 p_k 에 측정 노이즈 v_k 가 더해진 값입니다. 이를 행렬 방정식 $\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$ 로 표현합니다.

측정 행렬(Measurement Matrix, H)

측정 행렬 H는 상태 공간(3차원)을 측정 공간(1차원)으로 변환(mapping)하는 역할을 합니다. 우리는 상태 벡터의 세 성분 중 첫 번째인 위치만 측정하므로, \mathbf{H} 는 다음과 같은 1x3 행렬이 됩니다.

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

이 행렬 \mathbf{H} 는 상태 벡터 \mathbf{x}_k 에서 위치 성분 p_k 만을 '추출'하여 측정값 z_k 와 비교할 수 있도록 만들어주는 역할을 합니다.

섹션 5: 알고리즘의 전모: 행렬 형태의 예측-갱신 사이클

이제 칼만 필터를 구성하는 5개의 핵심 방정식을 소개합니다. 이 방정식들은 '예측(Predict)'과 '갱신(Correct/Update)'이라는 두 단계의 재귀적 루프를 통해 작동합니다.¹⁰

칼만 필터는 항상 두 가지를 추적합니다.

1. $\hat{\mathbf{x}}$ (**상태 추정치**): 시스템 상태에 대한 최선의 추정값 (벡터).
2. \mathbf{P} (**오차 공분산 행렬**): 상태 추정치의 불확실성을 나타내는 행렬. 대각 원소들은 각 상태 변수(위치, 속도, 가속도)의 분산(불확실성)을, 비대각 원소들은 각 변수 오차 간의 상관관계를 나타냅니다.¹⁴

1단계: 예측 (시간 갱신, Time Update)

이 단계에서는 이전 시점의 추정치를 바탕으로 현재 시점의 상태를 '예측'합니다.

- **방정식 1: 상태 추정치 예측**

$$\hat{\mathbf{x}}_k^- = \mathbf{F}\hat{\mathbf{x}}_{k-1}$$

왜? 이전 시점의 최종 추정치($\hat{\mathbf{x}}_{k-1}$)에 우리의 물리 모델(\mathbf{F})을 적용하여 현재 상태를 예측합니다. 위치자 '-'는 측정을 반영하기 '전(prior)'의 예측값임을 의미합니다.²³

- **방정식 2: 오차 공분산 예측**

$$\mathbf{P}_k^- = \mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^T + \mathbf{Q}$$

왜? 불확실성도 함께 예측합니다. $\mathbf{F}\mathbf{P}_{k-1}\mathbf{F}^T$ 항은 이전의 불확실성(\mathbf{P}_{k-1})이 물리 모델(\mathbf{F})을 통과하면서 어떻게 변하는지를 나타냅니다 (분산을 변환하는 기본적인 통계적 성질).²⁴ 여기에 **프로세스 노이즈 공분산** \mathbf{Q} 를 더해줍니다. 이는 우리의 모델이 완벽하지 않다는 사실을 인정하고, 그로 인해 발생할 불확실성을 시스템에 주입하는 과정입니다. 이 과정을 통해 필터가 자신의 예측을 과신하는 것을 막고, 새로운 측정에 귀를 기울이도록 만듭니다.²³

2단계: 갱신 (측정 갱신, Measurement Update)

이 단계에서는 실제 측정값을 사용하여 '예측'을 '보정'합니다.

- **방정식 3: 칼만 이득 계산**

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})^{-1}$$

왜? 이것이 바로 섹션 3에서 보았던 1차원 이득의 다차원 버전입니다. 측정값의 '놀라움'에 얼마만큼의 가중치를 부여할지 최적으로 계산합니다. 분모에 해당하는 $(\mathbf{H} \mathbf{P}_k^- \mathbf{H}^T + \mathbf{R})$ 항은 측정 공간에서의 총 불확실성(예측된 불확실성 + 측정 자체의 불확실성)을 의미합니다. 칼만 이득 \mathbf{K}_k 는 예측의 불확실성(\mathbf{P}_k^-)과 측정의 총 불확실성 사이의 비율을 통해 최적의 균형점을 찾습니다.¹⁰

- **방정식 4: 상태 추정치 갱신**

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H} \hat{\mathbf{x}}_k^-)$$

왜? 이것이 핵심적인 보정 단계입니다. 예측된 상태($\hat{\mathbf{x}}_k^-$)를 가져와, 측정값과 예측된 측정값의 차이인 혁신($\mathbf{z}_k - \mathbf{H} \hat{\mathbf{x}}_k^-$)에 '신뢰도 가중치'인 칼만 이득 \mathbf{K}_k 를 곱한 값을 더해줍니다.⁶

- **방정식 5: 오차 공분산 갱신**

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^-$$

왜? 이것이 바로 불확실성이 줄어드는 마법이 일어나는 순간입니다. 새로운 측정을 반영하고 나면 우리의 불확실성은 감소해야 합니다. $(\mathbf{I} - \mathbf{K}_k \mathbf{H})$ 항은 항상 1보다 작은 값을 가지는 행렬이므로(정확히는 그 행렬의 크기가 작아짐), 새로운 공분산 \mathbf{P}_k 는 예측된 공분산 \mathbf{P}_k^- 보다 항상 작아집니다. 즉, 측정은 우리를 더 '확신'에 차게 만듭니다.⁶

섹션 6: 노이즈의 해부: Q와 R 행렬 심층 분석

\mathbf{Q} 와 \mathbf{R} 은 단순히 필터 성능을 조절하기 위한 임의의 파라미터가 아닙니다. 이들은 우리 시스템에 존재하는 노이즈와 불확실성에 대한 물리적 기술이며, 이들을 올바르게 설정하는 것은 필터 성능에 지대한 영향을 미칩니다.²⁴

측정 노이즈 공분산 (R)

R 행렬은 우리 센서의 불확실성을 나타냅니다. 위치를 측정하는 단일 센서의 경우, R은 1×1 행렬, 즉 측정 노이즈의 분산 값($\sigma_{\text{measurement}}^2$)인 스칼라가 됩니다. 이 값은 보통 센서의 데이터시트에서 찾거나, 정지해 있는 물체를 반복적으로 측정하여 그 값들의 분산을 계산함으로써 실험적으로 결정할 수 있습니다.²⁴

프로세스 노이즈 공분산 (Q)

Q 행렬은 종종 가장 이해하기 어려운 부분으로 여겨집니다. Q는 우리 **모델의 불확실성**을 나타냅니다. 우리가 사용한 '등가속도' 모델은 현실에 대한 가정일 뿐입니다. 실제 카트는 도로의 요철, 바람의 영향 등으로 인해 가속도가 미세하게 변할 수 있습니다. Q는 이처럼 모델이 설명하지 못하는 무작위적인 변화의 크기를 정량화합니다.²⁸

Q 행렬을 단순히 추측하는 대신, 물리적 원리로부터 직접 유도해 보겠습니다. 이는 물리적 세계와 필터의 수학을 연결하는 매우 중요한 과정입니다.

1. 실제 운동은 우리의 '등가속도' 모델에 더해, 모델링되지 않은 미세하고 불규칙적인 **가속도의 변화율(jerk, j)**을 포함한다고 가정합니다. 이 무작위적인 jerk가 우리 프로세스 노이즈의 근원이며, 그 분산을 σ_j^2 라고 합니다.
2. 시간 간격 Δt 동안, 이 무작위 jerk j_k 는 가속도에 $j_k \cdot \Delta t$, 속도에 $\frac{1}{2} j_k \cdot (\Delta t)^2$, 위치에 $\frac{1}{6} j_k \cdot (\Delta t)^3$ 만큼의 무작위적인 변화를 일으킵니다.
3. 이 변화를 노이즈 벡터 $\mathbf{w}_k = \mathbf{G} j_k$ 로 표현할 수 있습니다. 여기서 G는 스칼라 값인 jerk를 상태 공간(위치, 속도, 가속도)에 대한 영향으로 변환하는 '노이즈 입력 행렬'입니다:

$$\mathbf{G} = \begin{bmatrix} \frac{1}{6}(\Delta t)^3 & \frac{1}{2}(\Delta t)^2 & \Delta t \end{bmatrix}$$

4. 변환된 확률 변수의 공분산은 $\mathbf{Q} = E[\mathbf{w}_k \mathbf{w}_k^T] = E[\mathbf{G} j_k j_k^T \mathbf{G}^T] = \mathbf{G} E[j_k^2] \mathbf{G}^T$ 로 계산됩니다.²⁹
5. $E[j_k^2]$ 는 jerk의 분산인 σ_j^2 이므로, 최종적으로 $\mathbf{Q} = \mathbf{G} \mathbf{G}^T \sigma_j^2$ 라는 관계식을 얻습니다.
6. 이 행렬 곱셈을 수행하면, Δt 와 근원적인 물리적 노이즈 σ_j^2 로만 구성된 완전한 3×3 Q 행렬이 유도됩니다. 이 엄밀한 유도 과정은 Q가 마법이 아니라, 물리적 불확실성이 시스템 동역학을 통해 전파되는 과정을 수학적으로 표현한 것임을 보여줍니다.

Q/R 비율: 필터의 '성격'

Q와 R의 상대적인 크기는 필터의 행동 양식을 결정합니다.²⁸

- **Q가 R에 비해 크다면:** 이는 (모델의 불확실성) >> (측정의 불확실성)을 의미합니다. 이 경우 칼만 이득 K가 커져서, 필터는 새로운 측정값을 더 신뢰하게 됩니다. 필터는 실제 변화에 민첩하게 반응하지만, 측정 노이즈에도 민감해질 수 있습니다.
- **R이 Q에 비해 크다면:** 이는 (측정의 불확실성) >> (모델의 불확실성)을 의미합니다. 이 경우 칼만 이득 K가 작아져서, 필터는 자신의 예측을 더 신뢰하게 됩니다. 필터의 추정치는 매우 부드러워지

지만, 실제 상태의 급격한 변화를 놓치거나 반응이 느려질 수 있습니다(소위 '게으른' 필터).³⁰ 따라서 '튜닝'이란 특정 응용 분야에 가장 적합한 균형점을 찾는 과정이라 할 수 있습니다.