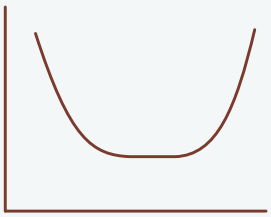
PROBABILISTIC GRAPH MODELING
FOR ENGINEERING DISCIPLINES · VOLUME I



Reliability Engineering by Derivation

From engineering questions to probabilistic inference and decisions

672 pages

Ricardo Marquez, Ph.D.

ricardo-marquez.com

Reliability Engineering by Derivation

Volume I

From engineering questions to probabilistic inference and decisions

A lesson sequence, written up as a book

Ricardo Marquez

October 31, 2026 (work in progress: one chapter per lesson, added as the lessons proceed;
Volume II holds Parts V–VII, Volume III Parts VIII–IX)

How to read this book

This book teaches one skill: to take a reliability question stated in ordinary engineering language, write it as a probability or an expectation over the unknowns of a physical model, derive what that expression requires, compute it, check it, and act on it. It is a companion to the two-volume *Reliability Engineering*, which states the discipline's concepts and vocabulary; this book goes deeper into the formulation and derivation of the probabilistic questions that discipline raises, and it does not repeat that book's introductory material on failure modes, maintenance organization or standards.

Every chapter follows the same progression, and the progression is the method.

Step 0: the plant graph. The physical system as it is, before any question is asked: its nodes, paths and sources; which quantities are pinned and which are free; the physical relations that must hold among them. The free quantities form the state vector $\mathbf{z}$. The graph is built once and serves every question put to it.

Step 1: the question. A probability or an expectation of a function of the state, with the unknown on the left of the conditioning bar and everything treated as known on the right.

Step 2: the graph extended for the question. Scenario nodes, a second drawing when two structures compete, and the evidence attached as rows: readings, counts, handbook values as priors, each with the modeling assumption it rests on, what that assumption buys and how it can fail.

Step 3: the posterior of the state. The rows combined into one objective; its minimum, the MAP state $\hat{\mathbf{z}}$; its curvature, the information $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; and the covariance $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$, derived and then computed for the plant, with the exact posterior beside them where the model admits one. It is built once and serves every question.

Step 4: the answers. Each question read from that posterior, from the engineering question through its general and specialized forms down to its number: an entry of the state or the gradient of a derived quantity, the quantity at a time asked included; when drawings compete, a ratio of evidences; when actions compete, an expected consequence.

Step 5: evaluation and checks. The closed forms evaluated, and their approximations tested against the full calculation.

Step 6: the report. The probability, the Bayes factor or the expected consequence with its band, or a crossing time with its band by inversion.

Step 7: the next measurement. If the band is too wide, the candidate measurements ranked by the width each would remove, and a return to step 3 with the new row.

Five kinds of question cover everything the book asks, and they are distinguished by the mathematical object that answers them rather than by the domain word in the question: *estimation* (a marginal of the posterior: a mean and a spread), *prediction under a threshold* (a tail probability of that marginal, or of the trend when the quantity moves in time), *hypothesis comparison* (a ratio of evidences between two complete models), *diagnosis* (hypothesis comparison

run over a family of candidate faults generated from the graph, returned as a ranking), and *design* (a preposterior: the expected posterior after a measurement not yet taken; or an expected consequence minimized over the actions). Five computations sit underneath: the posterior of the state; the trend, the posterior over a state indexed by time, read as the sign of a trend, a crossing time or a forecast; the evidence of a model and its ratios; the preposterior of a measurement not yet taken; and the expected consequence of an action, minimized over the actions. The notation chapter that follows fixes the symbols, and the appendix collects the questions in an atlas that maps each to its mathematical form and its derivation.

The register is that of a textbook written for a reader who builds or uses physical models of engineered systems and wants every formula derived rather than quoted. Every new quantity is introduced with its name, meaning, units and equation; every derivation proceeds one relation per line from the probability rules; every numerical example is worked with units. Where the field uses a shorthand term, a *Jargon* box expands it once. Three systems recur with growing complexity: an equipment hall with one heat-removal path, a set of valves and pumps whose test and failure records are the evidence, and a redundant cooling plant that later grows into a data center.

The framework the book teaches is physics-constrained probabilistic modeling: a plant written as a graph, its unknowns as a state vector, every piece of evidence as a row, and every question as a probability over that state. TERRA is the author's own implementation of this framework; the book names the framework and the computations, not the software, and every number in it is reproducible from the derivations as written.

Each chapter closes with a statement of what to retain, a table of its terms with their meaning, units and the section that defines each, and exercises at three levels, formulation, derivation and application, each followed by its answer.

Contents

How to read this book **ii**

Notation and vocabulary **viii**

I Thinking in Probabilities **1**

1 Engineering systems as graphs **2**
1.1 One system and several questions . . . 2
1.2 The posterior of the state (step 3) . . . 15
1.3 The answers (step 4) . . . 21
1.4 The checks (step 5) . . . 26
1.5 The report (step 6) . . . 28
1.6 The next measurement (step 7) . . . 29
1.7 What to retain . . . 32
1.8 Exercises . . . 34

2 The language of probabilistic questions **37**
2.1 Engineering questions and their mathematical form . . . 37
2.2 The grammar of probability statements . . . 40
2.3 The general forms and the five computations . . . 43
2.4 Estimation, prediction and joint events . . . 51
2.5 Model comparison, diagnosis and decisions . . . 62
2.6 Measurements, classification and workflow . . . 66
2.7 What to retain . . . 72
2.8 Exercises . . . 74

3 Uncertainty and conditional probability **76**
3.1 Uncertainty and evidence in reliability questions . . . 76
3.2 Conditional probability and Bayes' theorem . . . 80
3.3 Bayesian diagnosis of a cooling fault . . . 82
3.4 Marginalization and dependence . . . 85
3.5 Conditioning in a physical model . . . 89
3.6 Predictive variance and the posterior covariance . . . 97
3.7 What to retain . . . 103
3.8 Exercises . . . 105

4 The anatomy of probabilistic inference **107**
4.1 From a plant graph to an engineering probability . . . 107
4.2 The posterior of the state (step 3) . . . 115
4.3 The answers (step 4) . . . 124

4.4 The checks (step 5) . . . 130
4.5 The report (step 6) . . . 135
4.6 The next measurement (step 7) . . . 136
4.7 What to retain . . . 140
4.8 Exercises . . . 144

II Estimating Reliability from Data 146

5 Estimating failure probability from data 147
5.1 Failure probability from a record of tests . . . 147
5.2 The posterior of the state (step 3) . . . 155
5.3 The answers (step 4) . . . 159
5.4 The checks (step 5) . . . 173
5.5 The report (step 6) . . . 176
5.6 The next measurement (step 7) . . . 178
5.7 What to retain . . . 181
5.8 Exercises . . . 183

6 Estimating failure rates from records 185
6.1 Failure rates from operating records . . . 185
6.2 The posterior of the state (step 3) . . . 197
6.3 The answers (step 4) . . . 202
6.4 The checks (step 5) . . . 221
6.5 The report (step 6) . . . 226
6.6 The next measurement (step 7) . . . 228
6.7 What to retain . . . 231
6.8 Exercises . . . 234

7 Failure-time distributions and survival analysis 236
7.1 Lifetimes from a record with survivors . . . 236
7.2 The posterior of the state (step 3) . . . 250
7.3 The answers (step 4) . . . 255
7.4 The checks (step 5) . . . 264
7.5 The report (step 6) . . . 272
7.6 The next measurement (step 7) . . . 274
7.7 What to retain . . . 277
7.8 Exercises . . . 279

8 Repair rates and component availability 281
8.1 Repair rates and the availability of a repairable component . . . 281
8.2 The posterior of the state (step 3) . . . 289
8.3 The answers (step 4) . . . 293
8.4 The checks (step 5) . . . 305
8.5 The report (step 6) . . . 310
8.6 The next measurement (step 7) . . . 312
8.7 What to retain . . . 316
8.8 Exercises . . . 317

III Predicting Future Reliability 320

9 Mission reliability 321
9.1 Mission reliability under an uncertain failure rate . . . 321
9.2 The posterior of the state (step 3) . . . 327
9.3 The answers (step 4) . . . 331
9.4 The checks (step 5) . . . 344
9.5 The report (step 6) . . . 348
9.6 The next measurement (step 7) . . . 349
9.7 What to retain . . . 354
9.8 Exercises . . . 355

10 Remaining life and conditional survival 358
10.1 Remaining life of units that have already survived . . . 358
10.2 The posterior of the state (step 3) . . . 365
10.3 The answers (step 4) . . . 368
10.4 The checks (step 5) . . . 385
10.5 The report (step 6) . . . 389
10.6 The next measurement (step 7) . . . 391
10.7 What to retain . . . 395
10.8 Exercises . . . 398

11 Predicting future failure counts 400
11.1 Failure counts over a future period . . . 400
11.2 The posterior of the state (step 3) . . . 406
11.3 The answers (step 4) . . . 409
11.4 The checks (step 5) . . . 425
11.5 The report (step 6) . . . 428
11.6 The next measurement (step 7) . . . 430
11.7 What to retain . . . 435
11.8 Exercises . . . 437

12 Aging and time-varying failure rates 439
12.1 A failure rate that changes with time . . . 439
12.2 The posterior of the state (step 3) . . . 446
12.3 The answers (step 4) . . . 450
12.4 The checks (step 5) . . . 465
12.5 The report (step 6) . . . 470
12.6 The next measurement (step 7) . . . 472
12.7 What to retain . . . 476
12.8 Exercises . . . 477

IV System Reliability and Dependencies 480

13 Series, parallel and k-out-of-n systems 481
13.1 Services provided by several units . . . 481
13.2 The posterior of the state (step 3) . . . 488
13.3 The answers (step 4) . . . 490
13.4 The checks (step 5) . . . 512
13.5 The report (step 6) . . . 516
13.6 The next measurement (step 7) . . . 518

13.7 What to retain . . . 522
13.8 Exercises . . . 524

14 Fault trees and minimal cut sets **527**
14.1 The combinations that take a service down . . . 527
14.2 The posterior of the state (step 3) . . . 533
14.3 The answers (step 4) . . . 535
14.4 The checks (step 5) . . . 553
14.5 The report (step 6) . . . 557
14.6 The next measurement (step 7) . . . 559
14.7 What to retain . . . 563
14.8 Exercises . . . 564

15 Dependent failures and common causes **567**
15.1 Dependent failures and the common-cause share . . . 567
15.2 The posterior of the state (step 3) . . . 576
15.3 The answers (step 4) . . . 581
15.4 The checks (step 5) . . . 593
15.5 The report (step 6) . . . 600
15.6 The next measurement (step 7) . . . 602
15.7 What to retain . . . 607
15.8 Exercises . . . 609

16 Repairable and multistate systems **612**
16.1 Repair capacity and the states of a redundant pair . . . 612
16.2 The posterior of the state (step 3) . . . 619
16.3 The answers (step 4) . . . 621
16.4 The checks (step 5) . . . 640
16.5 The report (step 6) . . . 646
16.6 The next measurement (step 7) . . . 648
16.7 What to retain . . . 651
16.8 Exercises . . . 652

A Atlas of probabilistic questions **655**

Notation and vocabulary

Every symbol below is used with the same meaning in every chapter. The unknown always stands on the left of a conditioning bar and the evidence on the right. Vectors and matrices are bold.

Symbol	Unit	Meaning
The plant graph and its state		
$\mathbf{z}$, z_i	varies	State vector of the plant graph: every free quantity at once (temperatures, flows, conductances, log-rates); its i-th entry.
$\mathbf{y}$, y	varies	The evidence: all readings, counts and prior information treated as known; a single reading.
$r_k(\mathbf{z})$	varies	Row k of the model: a physical relation that must hold ($r_k = 0$), a reading compared with the state ($r_k = z_i - y$, in the units of the reading), or a prior ($r_k = z_i - z_i^0$).
w_k, σ_k	varies	Weight (precision) of row k and the standard deviation it corresponds to, $w_k = 1/\sigma_k^2$. A row may equivalently be written normalized, r_k/σ_k, with unit weight; both put the same term $\frac{1}{2}w_k r_k^2$ in the objective.
T, T_a	°C	Temperature of the equipment hall; ambient temperature (pinned).
Q, G	kW, kW/K	Heat load supplied to the hall; effective thermal conductance of the heat-removal path.
g, G_{ref}	–, kW/K	Log-conductance $\ln(G/G_{\text{ref}})$, the entry of the state that carries the conductance; its reference value.
y_T, σ_T	°C, K	A temperature reading of the hall and the standard deviation of its error.
$T_{\max}$, f^*	varies	A limit on a temperature; a requirement level on a state entry or on a derived quantity $f(\mathbf{z})$: the event $f \geq f^*$ is *the requirement holds*, and an upper limit is the complementary event (Chapter 2). From Part V a starred plant quantity (Q_1^*, T_1^*) is a scenario value: the quantity at a load not yet applied.
ω, ξ	–	Log-rate $\ln(\lambda/\lambda_{\text{ref}})$ and logit of a probability, $\ln(p/(1-p))$: the unbounded entries that carry a rate or a probability in the state.
u, v	–	Log-scale $\ln\eta$ and log-shape $\ln\beta$ of a Weibull lifetime, the two entries of the state in a lifetime fit.
Probability		
$\Pr(A)$, $\Pr(A \mid B)$	–	Probability of event A; conditional on event B.

Symbol	Unit	Meaning
$p(\cdot)$	varies	A probability density; $p(\mathbf{z} \mid \mathbf{y})$ is the posterior of the state, $p(\mathbf{y} \mid \mathbf{z})$ the likelihood, $p(\mathbf{z})$ the prior.
$\mathbb{E}$, Var, Cov	varies	Expectation, variance, covariance.
$\mathcal{N}(\mu, \sigma^2)$	–	Normal distribution with mean μ and variance σ^2; Φ its standard cumulative distribution function.
Beta(a, b), Gamma(a, b)	–	Beta distribution with shape parameters a, b; gamma distribution with shape a and rate b.
Binomial(n, p), Poisson(μ)	–	Number of failures in n trials with failure probability p; count with mean μ.
Inference objects		
$\mathcal{L}(\mathbf{z})$, $\mathcal{L}_k(\mathbf{z})$	–	The objective, minus the log posterior up to a constant, and the term row k puts in it: $\mathcal{L} = \sum_k \mathcal{L}_k$, with $\mathcal{L}_k = \frac{1}{2} w_k r_k(\mathbf{z})^2$ for a Gaussian row and $\mathcal{L}_k = -\ln p(y_k \mid \mathbf{z})$ for a count, lifetime or demand row (Section 4.2.7).
$\mathcal{L}'$, $\mathcal{L}''$, $\mathcal{L}'_k$, $\mathcal{L}''_k$	varies	For a state of one entry, say ω: the slope and the curvature of the objective, $\mathrm{d}\mathcal{L}/\mathrm{d}\omega$ and $\mathrm{d}^2\mathcal{L}/\mathrm{d}\omega^2$, and of row k's term; they add over the rows, and the Newton step is $\Delta\omega = -\mathcal{L}'/\mathcal{L}''$ (Chapter 5).
Working form	varies	The way a count row joins the Jacobian beside Gaussian rows (equation (4.43)): the line $\sqrt{\mu}\,\partial \ln \mu / \partial \mathbf{z}$ with the working residual $(\mu - k)/\sqrt{\mu}$ for a Poisson count k of mean μ, and $\sqrt{np(1-p)}\,\partial \xi / \partial \mathbf{z}$ with $(np - k)/\sqrt{np(1-p)}$ for a binomial count carried on the logit ξ. The pair reproduces the row's gradient and the $\mathbf{J}^\mathsf{T}\mathbf{J}$ part of its Hessian; its half square is not the row's term.
$\hat{\mathbf{z}}$	varies	The maximum a posteriori (MAP) state, the minimizer of $\mathcal{L}$.
$\mathbf{J}$	varies	Jacobian of the rows with respect to the state, evaluated at $\hat{\mathbf{z}}$: for a Gaussian row the line $\partial r_k / \partial z_i$, for a count row the line of its working form (Section 4.2.7).
$\mathbf{\Lambda}$, $\mathbf{H}$, $\mathbf{C}$	varies	Information matrix $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$ with $\mathbf{W} = \mathrm{diag}(w_k)$ ($\mathbf{J}^\mathsf{T}\mathbf{J}$ for normalized rows); the Hessian of $\mathcal{L}$, $\mathbf{H} = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{C}$, with $\mathbf{C}$ the curvature term the rows add beyond the information, zero for linear Gaussian rows and for count rows on their canonical parameter (equation (4.44)).
Λ	varies	For a state of one entry, the information, the curvature of the objective at the MAP, $\Lambda = \mathcal{L}''(\hat{\omega})$ (Chapter 5).
$\mathbf{\Sigma}$, Σ_{ij}	varies	Posterior covariance of the state under the Laplace approximation, $\mathbf{H}^{-1}$ ($\mathbf{\Lambda}^{-1}$ when $\mathbf{C} = \mathbf{0}$); its entries. The posterior spread of entry i is written $\sqrt{\Sigma_{ii}}$; σ_k always names a row's standard deviation and σ_p a physics row's tolerance.
σ^2_ω, σ_ω	varies	For a state of one entry, say ω or ξ: its Laplace variance $1/\Lambda$ and standard deviation (σ^2_ξ, σ_ξ for the logit). A derived quantity with derivative $g = \mathrm{d}f/\mathrm{d}\omega$ at the MAP has variance $g^2\sigma^2_\omega$.

Symbol	Unit	Meaning
ρ	K/kW	Thermal resistance of a path, $1/G$ (Chapter 4); not a correlation.
f_L, f_U	varies	The ends of the credible interval of a derived quantity $f(\mathbf{z})$ (Chapter 2).
y_{new}, $\tilde{y}_a$	varies	A future observation (the predictive) and the outcome of a candidate measurement a before it is taken (the preposterior).
M, M_j, B_{AB}	–	A drawing (model) of the plant; the members of a family; the Bayes factor $p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$, written B_{10} for M_1 against M_0.
$f_t(\mathbf{z})$, $\mathbf{g}(t)$, $\bar{t}$	varies	A quantity read at time t from a state indexed by time, $f_t = \mathbf{g}(t)^{\mathsf{T}}\mathbf{z}$ with the gradient $\mathbf{g}(t) = (1, t - \bar{t})^{\mathsf{T}}$ about the center $\bar{t}$ of the record, the state holding a level entry α and a trend entry γ (the trend form, Chapter 2).
$U(a)$, u	varies	The expected value of a candidate action a over the outcomes it may produce, and the utility u of the updated posterior it is built from (the design form, Chapter 2).
$\mathbf{J}_a$, $\mathbf{\Lambda}_a$	varies	The Jacobian rows of a candidate measurement and its information $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ (expected counts at $\hat{\mathbf{z}}$ for a count row).
κ	–	Assurance level of an acceptance rule: the posterior probability a requirement must reach, $\Pr(R \geq R_{\min} \mid \mathbf{y}) \geq \kappa$ (Chapter 5).
π_0	–	The prior row's share of the information, $\mathcal{L}_0''/\mathcal{L}''$ for a state of one entry (Chapter 5); in Chapter 16 the stationary probability that no unit is up.
$f(\mathbf{z})$, $\mathbf{g}$	varies	A derived quantity of the state and its gradient $\partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$.
μ_f, σ_f^2	varies	Mean and variance of a derived quantity under the delta method: $f(\hat{\mathbf{z}})$ and $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. From Part V the spread is written $\text{sd}(f \mid \mathbf{y}) = \sqrt{\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}}$.
$p(\mathbf{y} \mid M)$	–	Evidence (marginal likelihood) of model M.
B_{10}	–	Bayes factor of model M_1 against M_0, $p(\mathbf{y} \mid M_1)/p(\mathbf{y} \mid M_0)$; usually reported as a log in nats.
$\mathbf{\Upsilon}$, υ	varies	The gain of the correction form, $\mathbf{\Sigma}_0\mathbf{A}^{\mathsf{T}}(\mathbf{A}\mathbf{\Sigma}_0\mathbf{A}^{\mathsf{T}} + \mathbf{R})^{-1}$, which moves the prior mean toward the readings; its scalar value for one reading, the weight $\sigma_0^2/(\sigma_0^2 + \sigma_T^2)$ (Chapter 3).
$\mathbf{\Sigma}_a$	varies	Preposterior covariance after a candidate measurement a with information $\mathbf{\Lambda}_a$: $(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a)^{-1}$.
EIG	nat or bit	Expected information gain of a candidate measurement.
Reliability quantities (as in the companion book)		
T, $R(t)$, $F(t)$	h, –, –	Time to failure; reliability $\Pr(T > t)$; cumulative failure probability.
$f(t)$, $\lambda(t)$, $\Lambda(t)$	h^{-1}, h^{-1}, –	Failure density; hazard function; cumulative hazard.
λ, μ	h^{-1}	Constant failure rate; constant repair rate.
p, k, n	–	Failure probability per demand; number of failures observed; number of trials.

Symbol	Unit	Meaning
T_{tot}	h	Exposure (total operating time on test) behind a failure count.
β, η	–, h	Weibull shape and scale (characteristic life).
t_i, δ_i, c_i	h, –, h	Observed time of unit i; its event indicator (1 failed, 0 censored); its censoring time.
t_{50}, t_0, T_{new}	h	Median lifetime; an age already reached; the lifetime of a new unit.
ℓ, ℓ_p	–	Log-likelihood; profile log-likelihood (maximized over the other entries).
r, D	–, h	Number of repairs observed; their total duration.
U, A_{req}	–	Unavailability $1 - A$; a required availability.
α, γ	–, yr^{-1}	The log-rate case of the level and trend entries of the trend form: log-rate at the record's center and its trend per year in $\ln \lambda(t) = \alpha + \gamma\,(t - \bar{t})$; t^* the time the rate crosses a limit.
K_f, τ	–, h or yr	A future failure count; the future period or horizon it is counted over.
q_i, $\mathbf{s}$, $\phi(\mathbf{s})$	–	Unavailability of unit i; a configuration of units (up or down); the structure function, 1 when the configuration delivers the service.
R_{sys}, U_{sys}	–	Service availability and unavailability of several units under a drawing, by the 2^n enumeration.
β_{cc}	–	Common-cause share: the fraction of a unit's failures that fail all units sharing the cause (β is the Weibull shape).
A, MTTF, MTTR	–, h, h	Steady-state availability; mean time to failure; mean time to repair.
$L(a, \mathbf{z})$, $\mathcal{R}(a)$, C, C_e, c_a	R$	The loss of action a when the state is $\mathbf{z}$, its consequence; its expected value over the posterior, the risk $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$; the consequence of an event, of event e; the fixed cost of the action (the expected-consequence form, Chapter 2).

Vocabulary

Plant graph The representation of the physical system as nodes (places where a conserved quantity balances), edges (paths along which it flows) and sources, with the relations that must hold among them. Built before any question is asked.

State vector The collection of free quantities the inference estimates at once; outputs that are entries of the state need no further transformation.

Pinned, free A quantity treated as known (boundary condition) against one carried in the state.

Residual, row A relation written so that it equals zero when satisfied; physical rows constrain the state, reading rows compare it with evidence, prior rows carry handbook knowledge.

Prior, likelihood, posterior Knowledge before the evidence; the probability of the evidence given a state; knowledge after the evidence, by Bayes' theorem.

Marginalization Integrating a joint distribution over the unknowns not asked about; the law of

total probability.

Aleatory, epistemic Variability in outcomes the model treats as random; incomplete knowledge of quantities that have one true value.

Maximum a posteriori (MAP) estimate The state at which the posterior density is highest; the minimizer of the objective.

Laplace approximation Replacing the posterior by the Gaussian whose covariance is the inverse curvature of the objective at the MAP state.

Information matrix The matrix $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, the Hessian of the objective without the curvature term $\mathbf{C}$ of $\mathbf{H} = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{C}$ (equation (4.44)); precisions of independent rows add. For a state of one entry it is the number Λ.

Delta method Linearizing a derived quantity around the MAP state so that its variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$.

Change of variables Carrying a density to a new coordinate $u = h(x)$:

$$p_u(u) = p_x(h^{-1}(u))\,|\mathrm{d}x/\mathrm{d}u|,$$

the Jacobian factor keeping the probability of each interval; for several entries the factor is $|\det \partial\mathbf{z}/\partial\mathbf{u}|$. It gives the shape of a derived quantity, where the delta method gives only its spread (Section 4.3.3).

Variance budget The expansion of $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ into one term per state entry plus the cross terms from their covariance.

Threshold probability The probability that a state entry or derived quantity is below (or above) a limit; a normal tail under the Laplace approximation.

Posterior predictive The distribution of a future observation obtained by averaging its model over the posterior.

Credible interval An interval that contains the unknown with a stated posterior probability.

Evidence, Bayes factor The normalizing constant of one model's posterior; the ratio of two models' evidences, which compares the models rather than their parameters.

Preposterior The posterior that would result from a measurement not yet taken; the object that values a measurement before buying it.

Identifiability Whether the evidence determines the state entries separately or only in combination; read from the shape of the posterior, not from marginal spreads.

The five question categories Estimation, prediction under a threshold, hypothesis comparison, diagnosis and design, answered by a marginal, a tail of the posterior or of the trend, an evidence ratio, a ranking of evidences, and a preposterior or an expected consequence.

The five computations The objects that answer every category (Chapter 2): the *posterior* $p(\mathbf{z} \mid \mathbf{y})$, read for an entry or a derived quantity; the *trend*, the posterior over a state indexed by time or window with a trend entry, read as the probability that the trend is positive, as a crossing time with its band, or as a forecast; the *evidence* $p(\mathbf{y} \mid M)$ and its ratios; the *preposterior*, the expected posterior under a measurement not yet taken; and the *expected consequence* $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ of an action over the posterior, with the choice of the action that minimizes it (risk is probability times consequence).

Factor graph A graph of the state's entries and the factors (the rows) that bear on them; the plant graph with its rows is one, and the posterior is the normalized product of its factors (Chapters 1 and 4).

Variable elimination Integrating the entries a question does not name out of the joint distribution one at a time; for a Gaussian posterior it is the Schur complement, and its cost on a large sparse graph is set by the order of elimination (Chapter 3).

Belief propagation Computing the marginals of a graphical model by passing messages between neighboring factors and entries (the sum-product algorithm); exact on a graph without loops, where on a Gaussian graph it returns the Schur complements. Not used in this book, whose posteriors are factored jointly.

Kalman filter and smoother Variable elimination on a state indexed by time, taken in time order: the filter forward, the smoother back.

Importance sampling (reweighted draws) Drawing from an approximation, here the Laplace normal, and weighting each draw by the ratio of the exact posterior to it, so that the draws check the approximation (Chapter 20, in Volume II).

Markov chain Monte Carlo Drawing from a posterior by a Markov chain whose stationary distribution it is (Gibbs sampling, the Metropolis algorithm); the route when neither the Laplace normal nor reweighting is adequate, as for a hierarchical model (Chapter 33, in Volume III).

Explaining away Two causes independent before an effect is seen become dependent once it is seen: evidence for one lowers the probability of the other, as a biased sensor and a changed conductance compete for one disagreeing reading (Chapters 22 and 23, in Volume II).

Part I

Thinking in Probabilities

Chapter 1

Engineering Systems as Graphs

1.1 One system and several questions

The owner of a data center equipment hall wants to double its electrical load, from 30 to 60 kW, and the hall must then stay below 27.5°C. The record is thin: one temperature sensor, with an error of about half a kelvin, reads 23.2°C at the present load, and an engineering handbook gives the conductance of the hall's cooling path to the 18°C ambient as 7.5 kW/K, right to within a factor of about 1.35. The handbook value gives the first answer. Heat leaves the hall in proportion to its excess over the ambient, so at 7.5 kW/K the hall sits at 22°C under 30 kW, comfortably below its present limit of 24°C, and at 26°C under 60 kW, well below 27.5°C: the full load passes with room to spare.

The sensor tells another story. Taken as the truth, 23.2°C at 30 kW means a conductance of 5.77 kW/K, and at that conductance the full load puts the hall at 28.4°C, over the limit. One source says the hall passes and the other that it fails; each treats its own number as exact, neither is weighed against the other, and neither says how sure the owner can be. The decision needs a probability, that the hall stays below 27.5°C at a load at which it has never run, and that probability can only come from holding the handbook value, the reading and the energy balance that ties them together in one description of the hall, each with its own uncertainty. The question of this chapter is therefore

how a physical engineering system is represented so that the same representation serves every probabilistic question asked of it, and what must be fixed before any inference is performed.

Jargon: plant graph, node, path, source, pinned and free quantities, state vector, residual, row, observation, prior, posterior, model discrepancy, scenario node, Jacobian, normal equations, information matrix, covariance, Laplace approximation, factor graph

A *plant graph* represents a physical system as *nodes*, places where a conserved quantity such as energy or mass must balance; *paths* (edges), along which the quantity flows; and *sources*, which supply it. A quantity is *pinned* when it is treated as known, as a boundary condition, and *free* when it is unknown and must be inferred. The free quantities, collected in one vector, form the *state vector* $\mathbf{z}$. A *residual* is a relation written so that it equals zero when it is satisfied; each residual is a *row* of the model. A physical row states a relation that must hold among entries of the state; an *observation* (a reading) is a measured value related to an entry of the state with an error, and its row compares the two. The *prior* expresses what is known about the state before the readings, for example a handbook value of a conductance; the *posterior* $p(\mathbf{z} \mid \mathbf{y})$ expresses what is known after the evidence $\mathbf{y}$. A *model discrepancy* is the difference between the physical system and the model of it that would remain if every entry of the state were known exactly. A *scenario node* is a node added to the graph for an operating condition that has not occurred, so that a question can be asked about it. The *Jacobian* $\mathbf{J}$ is the matrix of derivatives of every row with respect to every entry of the state, one row per row of the model and one column per entry. The *normal equations* $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ give the step that best reduces all the rows at once when each is replaced by its tangent. The *information matrix* $\mathbf{J}^\mathsf{T}\mathbf{J}$ states how sharply the rows together fix the state, direction by direction, and its inverse is the posterior *covariance* $\boldsymbol{\Sigma}$, whose diagonal holds the variances of the entries and whose off-diagonal entries hold how they vary together. The *Laplace approximation* replaces the posterior by the normal distribution centered on its most probable state with that covariance. A *factor graph* draws the entries of the state and the rows that act on them, one node each.

Key idea

A graph describes the engineering system; a probabilistic question states what is to be learned about it. The plant graph, its pinned and free quantities and its physical rows are built once, and each question is a probability over the state of that one graph, with the unknown on the left of the conditioning bar and the evidence on the right. The posterior of the state is carried by three objects, the most probable state $\hat{\mathbf{z}}$, the covariance $\boldsymbol{\Sigma}$ and the Jacobian $\mathbf{J}$ of the rows, and every answer is an entry of the state read off them or a function of the state carried through them.

1.1.1 Several questions about one equipment hall

An engineer responsible for the hall might ask any of the following:

1. What is the actual temperature inside the hall?
2. What is the effective conductance of the cooling path?
3. Will the temperature exceed its limit if the electrical load doubles?
4. Is the temperature sensor biased?
5. Would installing another sensor reduce the uncertainty?
6. What happens if the cooling system partially fails?

The questions are different. The first two ask for the value of an unknown quantity; the third asks for the probability that a future temperature stays within a limit; the fourth asks which of two descriptions of the sensor is true; the fifth asks what a measurement not yet taken would be worth; the sixth asks about an operating condition that has not occurred. All six, however, concern the same physical system: the same hall, the same cooling path, the same sensor.

Constructing a separate mathematical model for every question would repeat the same physics six times, and the six models could disagree about it. The method of this book begins instead with one common representation of the equipment and its physical relations, the *plant graph*. It is step 0 of the method stated in the preface: the system as it is, before any question is asked, built once and kept for every question that follows. Section 1.1.6 writes the questions that this chapter answers over the graph.

1.1.2 The system, the question and the evidence

Three things must be fixed before any probabilistic inference is performed, and the plant graph is where they are fixed. The first is the division of the quantities: which are known and pinned (the ambient temperature, the heat load), which are unknown and free (the hall temperature, the conductance), and which are observed (the sensor reading). The field often separates the unknowns into *states*, such as temperatures and flows, and *parameters*, such as conductances and rates; in this book both kinds are entries of one state vector, and the inference treats them alike. The second is the set of physical relations that must hold among the quantities, such as conservation of energy at the hall. The third is the evidence: the readings, the records and the prior knowledge, attached to the graph as rows of their own.

The question is a separate object. It is stated over the graph after the graph is built, and it decides what is computed, not what the system is. The separation makes complex reliability assessments systematic: when a new question arrives, the graph is already in hand, and the work is to write the question over its state.

> **Principle 1.1.** *The physical model is reusable; the probabilistic question determines what is calculated from it. A plant graph is built once from the equipment and its physics, and every question is a probability or an expectation over the state of that graph.*

1.1.3 Larger plant graphs and factor graphs

The equipment hall of this chapter has one thermal path and two unknown quantities, the smallest graph with every kind of element. The same construction serves a plant of many components, and the graph on which an inference operates is a second object, drawn from the plant graph and the evidence attached to it.

A cooling network. Real engineering systems may include hundreds or thousands of components, each with its own equations and uncertainties. Consider a cooling network with two chillers, two pumps, a common electrical supply and several temperature sensors. Its plant graph may contain temperatures, pressures, flows and electrical states at its nodes; physical components represented by paths or by component relations; heat, mass and power conservation equations; sensor measurements and historical observations; failure and repair states of equipment; and shared parameters and common-cause dependencies.

Figure 1.1 draws a reduced version. The green boxes are free nodes: the two chillers and the two pumps, each carrying its status and capacity or flow, the supply header with its water temperature T_s and flow, and the hall with its temperature T. The gray box on the left is the pinned ambient, to which the chillers reject heat; the white box in the middle is the common electrical supply, which feeds all four machines; the white box at the upper right is the heat load of the hall. Solid green arrows are physical paths of heat and water; thin gray arrows are electrical feeds; dashed arrows are observations, two temperature readings and the maintenance record of a chiller. The common supply is a single node on which four components depend: if it fails, all four stop together, and a graph that omitted it would treat their failures as independent.

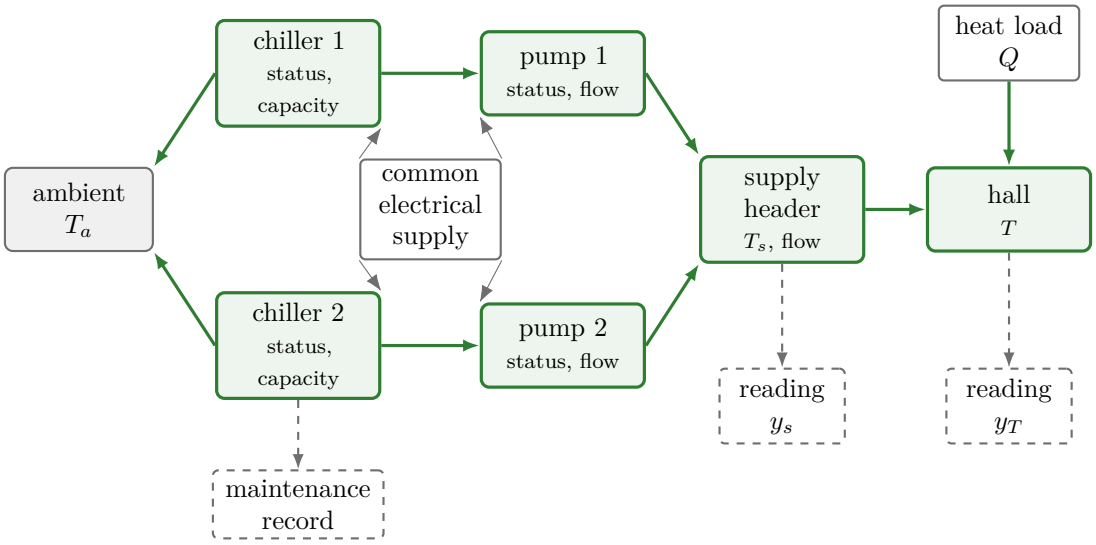


Figure 1.1: Plant graph of a reduced cooling network (illustrative; return piping omitted). Free: the status and capacity of each chiller, the status and flow of each pump, the header water temperature T_s and flow, and the hall temperature T. Pinned: the ambient temperature T_a and the heat load Q; the common electrical supply feeds all four machines and is the shared dependency. Observed: T_s and T through two readings, and chiller 2 through its maintenance record.

The mathematical principle is the one the single hall shows below. The network is a collection of physical relations,

$$r_1(\mathbf{z}) = 0, \quad r_2(\mathbf{z}) = 0, \quad \ldots, \quad r_m(\mathbf{z}) = 0,$$

one per balance and per component relation, over a state vector $\mathbf{z}$ that now has many entries. Observations and probability models for the uncertain quantities are attached as rows of their own. What is ultimately sought is the joint posterior of the state,

$$p(\mathbf{z} \mid \mathbf{y}, M),$$

where M represents the assumed system model, including its structure: which components exist, how they are connected and which dependencies are represented. Writing M explicitly matters because the structure itself may be uncertain, and a question about the structure is answered by comparing two such models, not by the posterior inside either.

Physical graphs and factor graphs. A physical graph and a probabilistic factor graph are related, but they are not the same object. The physical graph represents equipment and connections: halls, paths, chillers, pumps, sensors. A *factor graph* represents the variables of the inference and the mathematical relations among them, including the observations, the priors and the dependencies: each entry of the state is one node, each row of the model is one factor node, and each factor is linked to the entries its row involves. It is constructed from the physical graph and the evidence attached to it, and it shows which row acts on which unknown. The distinction becomes important when competing system structures and diagnosis are studied in later chapters. Two physical drawings of the same plant, one with a single cooling path and one with two, give two factor graphs with different variables and rows, and comparing them is comparing two models M. A sensor bias adds a variable to the factor graph without changing the equipment at all.

1.1.4 The plant graph (step 0)

The simplest plant graph with every kind of element has one free node, one pinned node, one path, one source and one observation, and the equipment hall is that graph. It is built before any question is asked and before any number is fitted.

Nodes, a path, a source and an observation. Consider a simplified equipment hall with one heat-removal path to the ambient air. Figure 1.2 draws it as a plant graph. The green box on the left is the hall, a free node whose temperature T is unknown; the gray box on the right is the ambient air, a pinned node whose temperature $T_a = 18°\text{C}$ is known. The solid green arrow between them is the heat-removal path, characterized by its effective thermal conductance G. The box above the hall is the heat load Q supplied by the equipment, a source. The dashed box below is the temperature sensor, and the dashed arrow from the hall to it is an observation: the sensor reports a value related to T, not T itself. Solid connections are physical relations; the dashed connection is evidence.

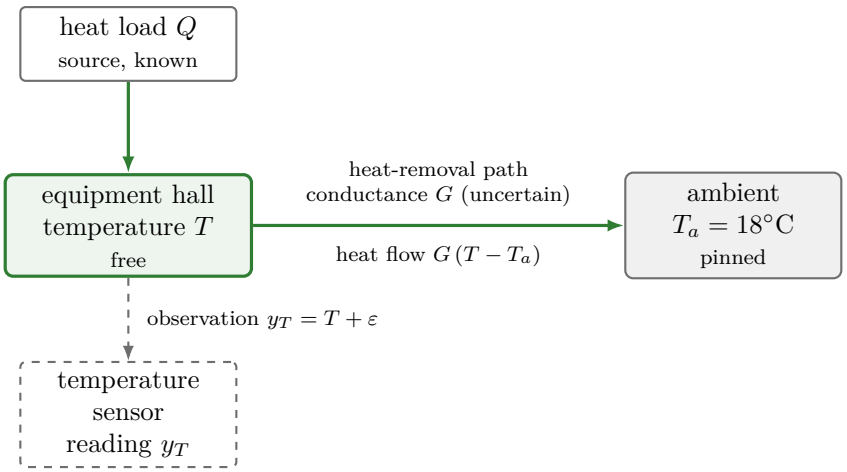


Figure 1.2: Plant graph of the equipment hall in steady state. Free: the hall temperature T and the conductance G of the heat-removal path. Pinned: the ambient temperature $T_a = 18°\text{C}$ and the heat load Q. Observed: T, through the reading y_T with error ε. Solid arrows are physical relations; the dashed arrow is an observation.

Table 1.1 lists the elements of the graph. The first column names each element, the second its role in the graph, the third its symbol and the fourth its unit; the last column states how the element enters this first example.

Table 1.1: The elements of the plant graph of Figure 1.2.

Element	Role	Symbol	Unit	In the example
Equipment hall	free node	T	°C	temperature unknown
Ambient air	pinned node	T_a	°C	known, 18°C
Heat load	source	Q	kW	thermal power supplied by the equipment, known
Thermal path	path (edge)	G	kW/K	effective conductance, uncertain
Temperature sensor	observation	y_T	°C	a reading related to the hall temperature

For now the ambient temperature and the heat load are assumed known, and the effective conductance is uncertain. The conductance lumps every mechanism by which heat leaves the

hall, the cooling equipment and the envelope alike, into one number; how well that lumping represents the hall is a modeling assumption, examined in Section 1.6.5. This is the starting physical model.

Conservation of energy at the hall node. A graph becomes useful for computation when the mathematical relations between its elements are specified. Assume steady state: the hall stores no more and no less heat from one moment to the next, so that its temperature is constant. Conservation of energy at the hall node then requires the heat entering to equal the heat leaving,

$$Q_{\text{in}} = Q_{\text{out}},$$

where Q_{in} and Q_{out} are thermal powers in kilowatts. The heat entering is the load, $Q_{\text{in}} = Q$. The heat leaving is modeled as proportional to the temperature difference across the path,

$$Q_{\text{out}} = G\,(T - T_a),$$

the defining relation of the effective conductance: G is the heat flow per kelvin of difference between the hall and the ambient. Combining the two,

$$Q = G\,(T - T_a), \tag{1.1}$$

where Q is the heat generated by the equipment in kW, G the effective thermal conductance in kW/K, T the hall temperature in °C and T_a the ambient temperature in °C. A temperature difference has the same value in kelvin and in degrees Celsius, so the units balance: kW/K times K is kW.

Equation (1.1) is a *physical relation*, not a probability statement. It states what must hold if the simplified thermal model is correct, whatever is or is not known about its quantities. Since $G > 0$, dividing by G and adding T_a gives

$$T = T_a + \frac{Q}{G}, \tag{1.2}$$

which gives the temperature when the other three quantities are specified. The hall is warmer than the ambient by Q/G: more load or less conductance raises it.

The forward calculation. The hall is operated at the load $Q = 30\,\text{kW}$, and its cooling path was designed for a conductance of $G = 6\,\text{kW/K}$. With $T_a = 18°\text{C}$, equation (1.2) gives

$$T = 18 + \frac{30\,\text{kW}}{6\,\text{kW/K}} = 18 + 5 = 23°\text{C},$$

the steady-state temperature of the hall at its design conductance. The calculation used no probability, estimated no uncertainty and used no observational evidence. It is a deterministic *forward calculation*: from pinned inputs and a specified conductance to the temperature they imply. Every later chapter needs it, because every inference compares temperatures computed this way with what the sensor reports. The temperature falls as the conductance grows, with the slope $\partial T/\partial G = -Q/G^2$, which at the design point is $-30/36 = -0.833\,\text{K}$ per kW/K: an error of 1 kW/K in the conductance moves the computed temperature by almost a kelvin. Doubling the load doubles the excess $T - T_a$ at every conductance: at $G = 6\,\text{kW/K}$ the hall at 60 kW would be at 28°C.

The state vector. The forward calculation required the conductance to be known. In the inference problems of this book it is not, and neither is the hall temperature. The *state vector* is the collection of the unknown quantities that the inference estimates at once. In the equipment hall the unknowns are the hall temperature and the thermal conductance, and one possible representation is

$$\mathbf{z} = \begin{bmatrix} T \\ G \end{bmatrix}.$$

The state holds everything the inference treats as free; the pinned quantities T_a and Q are not in it, and neither is the reading y_T, which is evidence. The concept recurs throughout the book: the graph fixes the state, the evidence constrains it, and each question is a statement about it. Quantities that are entries of the state, such as T here, are answered directly by the inference; quantities that are functions of the state, such as the heat flow through the path, require one further transformation.

The log-conductance. A conductance is positive, and an estimate of it must be positive too. In practice the conductance is therefore often represented by a logarithmic entry,

$$g = \ln\!\left(\frac{G}{G_{\mathrm{ref}}}\right),$$

where $G_{\mathrm{ref}} > 0$ is a reference conductance in kW/K and g is dimensionless. Inverting,

$$G = G_{\mathrm{ref}}\, \mathrm{e}^{g},$$

which is positive for every real g: an inference free to move g anywhere on the real line can never produce a negative or zero conductance. A change in g is a relative change in G, since $\mathrm{d}G = G_{\mathrm{ref}}\, \mathrm{e}^{g}\, \mathrm{d}g = G\, \mathrm{d}g$, so that $\mathrm{d}g = \mathrm{d}G/G$; uncertainty in g reads directly as a fractional uncertainty in G. The reference value only sets the origin of g. In this chapter it is taken as $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the design conductance of the forward calculation, so that $g = 0$ there, $G = 7.5\,\mathrm{kW/K}$ corresponds to $g = \ln 1.25 = 0.2231$, and a conductance 10 % above the reference to $g = \ln 1.1 = 0.0953$. The state vector becomes

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} T \\ g \end{bmatrix}, \tag{1.3}$$

with entries $z_1 = T$ in °C and $z_2 = g$, dimensionless. It is the one state of this chapter: every question below is a statement about these two entries.

The physics row. The thermal relation can now be written in terms of the state. Substituting $G = G_{\mathrm{ref}}\, \mathrm{e}^{g}$ into equation (1.1),

$$Q = G_{\mathrm{ref}}\, \mathrm{e}^{g}\, (T - T_a),$$

or, moving everything to one side,

$$r_{\mathrm{phys}}(\mathbf{z}) = G_{\mathrm{ref}}\, \mathrm{e}^{g}\, (T - T_a) - Q = 0. \tag{1.4}$$

The function $r_{\mathrm{phys}}(\mathbf{z})$ is a *residual*, in kW. A residual measures how far a candidate state is from satisfying a relation: it is zero when the relation holds exactly and grows as the candidate departs from it. With $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, $T_a = 18°\mathrm{C}$ and $Q = 30\,\mathrm{kW}$, the candidate $\mathbf{z} = (23°\mathrm{C},\ 0)$ gives $r_{\mathrm{phys}} = 6 \times 1 \times 5 - 30 = 0$: it is the forward calculation. The candidate $\mathbf{z} = (23.2°\mathrm{C},\ 0)$ gives $6 \times 5.2 - 30 = 1.2\,\mathrm{kW}$: at that temperature and conductance the path would remove 1.2 kW more than the equipment supplies, so the hall could not be in steady state. The candidate $(23.2°\mathrm{C},\ \ln(30/31.2)) = (23.2°\mathrm{C},\ -0.0392)$ again gives zero: the temperature 23.2°C

is consistent with the physics if the conductance is $30/5.2 = 5.769\,\mathrm{kW/K}$. The zero set of r_{phys} is the curve $T = T_a + Q/(G_{\mathrm{ref}}\,\mathrm{e}^{g})$ in the plane of (g, T); the physics alone allows every point of that curve and no point off it.

The same row has a second form, in the units of the quantity it fixes. Dividing equation (1.4) by $G_{\mathrm{ref}}\,\mathrm{e}^{g} = G$, which is positive, leaves the zero set unchanged and gives the row in kelvin, the departure of a candidate temperature from the temperature the physics implies. Divided further by a *model tolerance* σ_p, the temperature error in kelvin that the physics is allowed, it becomes a dimensionless row,

$$r_1(\mathbf{z}) = \frac{T - T_a - (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}}{\sigma_p}, \tag{1.5}$$

which counts the violation of the physics in multiples of its tolerance. This is the form in which the physics row enters every computation of the book. The formulation extends to systems of any size. A larger physical system is represented by a collection of residuals, one per relation: an electrical network has rows for Kirchhoff's current and voltage laws, a fluid network rows for mass conservation at its junctions and for the pressure losses along its pipes, and a cooling plant rows for the energy balance at each of its nodes. Each row has the form of equation (1.4), and the inference treats them alike.

Exact constraints and model discrepancy. Physical rows and reading rows should not automatically be treated as equivalent kinds of noisy measurement. Conservation of energy is a law: if the hall is in steady state and the model has captured every path, equation (1.4) holds exactly, and it can be imposed as an *exact constraint*, $r_{\mathrm{phys}}(\mathbf{z}) = 0$, which restricts the state to the zero set of the row. In a computation the exact constraint is carried as the physics row (1.5) with a model tolerance σ_p far smaller than every other spread of the problem, and exactness is the limit $\sigma_p \to 0$; the limit is checked by showing that the results stop moving as σ_p shrinks, never by using the row to eliminate an entry of the state. A reading row, by contrast, is meant to be violated by an amount of the order of the sensor's error: the reading is not the temperature.

Approximate physics is a third case. The single effective conductance of equation (1.1) is an idealization; the real hall has several paths, a conductance that varies with airflow, and stored heat. When the relation is only approximately true, the honest model is

$$G_{\mathrm{ref}}\,\mathrm{e}^{g}\,(T - T_a) - Q = \delta,$$

with an explicit *model discrepancy* δ in kW, an entry of its own with a stated prior, for example $\delta \sim \mathcal{N}(0, \sigma_\delta^2)$. Algebraically this is a physics row with the weight $1/\sigma_\delta^2$, and in the objective it looks like a reading row. It behaves differently. A measurement error is new at each reading and averages out over many readings; a discrepancy is a property of the model and repeats wherever the same operating condition repeats, so that many readings at one condition do not reduce it. Treating a discrepancy as measurement noise therefore makes the inference overconfident, and treating a law as noisy lets the inference trade physics against data. The choice between an exact constraint and an explicit discrepancy is made when the graph is built, and it is stated with the model. In this chapter conservation of energy is exact.

1.1.5 The record and the handbook value

The record of the hall is one reading. A temperature sensor installed in the hall reports, at the load of 30 kW,

$$y_T = 23.2^{\circ}\mathrm{C}.$$

The sensor is not perfectly accurate. Its measurement model is

$$y_T = T + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, \sigma_T^2),$$

where ε is the measurement error in kelvin, assumed normal with zero mean, and σ_T its standard deviation, which the sensor's specification gives as

$$\sigma_T = 0.5\,\mathrm{K}.$$

The reading is the only observation: the ambient temperature and the load are pinned, and nothing measures the conductance.

Knowledge of the conductance that does not come from the reading is also in hand. An engineering handbook, the design documents or earlier tests of similar halls give an estimate of the effective conductance with an uncertainty; for this hall (illustrative) the handbook gives 7.5 kW/K, with the true conductance within a factor of about 1.35 of it with probability of about 0.68. The handbook value is not a measurement of this hall. It is information held before the reading, and it enters the model as a prior.

1.1.6 The questions (step 1)

The engineer's questions can now be written over the graph, each as a probability or an expectation over the state, with the unknown on the left of the bar and the evidence on the right. Table 1.2 lists the five that this chapter answers: the first column is the question in engineering language; the second the general form that every question of its category takes, written in the book's notation for a quantity $f(\mathbf{z})$ derived from the state, a threshold f^* on it and the posterior density $p(\mathbf{z} \mid \mathbf{y})$; the third its mathematical form for the hall; the fourth its category among the five of the preface, which Chapter 2 develops; and the last the subsection where it is answered. The first two are estimation, of an entry of the state and of a quantity derived from one, the conductance $G = G_{\mathrm{ref}}\,\mathrm{e}^{g}$. The third and fourth are predictions under a threshold, tail probabilities of the posterior: the third at the present load, the fourth at a load of 60 kW at which the hall has never run, whose temperature T_2 is a quantity the graph acquires in step 2. The fifth is a design question about a measurement not yet taken; it asks nothing new of the present evidence, and it is answered as the next measurement for the first four.

Table 1.2: The questions asked of the equipment hall, the general form of their category, their mathematical form for the hall, their category and where each is answered. T_2 is the hall temperature at the load $Q_2 = 60\,\mathrm{kW}$ (Section 1.1.7); $\boldsymbol{\Sigma}_a$ the covariance expected after a candidate row with Jacobian row $\mathbf{J}_a$.

Engineering question	General form	Mathematical form	Category	Answered in
What is the hall temperature now?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$p(T \mid \mathbf{y})$, $T = z_1$	estimation	1.3.1
What is the effective conductance of the path?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$p(G \mid \mathbf{y})$, $G = G_{\mathrm{ref}}\,\mathrm{e}^{z_2}$	estimation of a derived quantity	1.3.1
Will the hall stay below 24°C?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$	$\Pr(T < 24^\circ\mathrm{C} \mid \mathbf{y})$	prediction under a threshold	1.3.2
Will the hall stay below 27.5°C at 60 kW?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$	$\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, $T_2 = T_a + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2}$	prediction under a threshold	1.3.2
Would another sensor narrow the answers?	preposterior covariance $(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$	$\boldsymbol{\Sigma}_a$ for each candidate row	design	1.6.1

The questions differ in what they compute and agree in what they are computed over: the same two entries of the state, under the same posterior. Two of the engineer's six questions

are not in the table. Whether the sensor is biased compares two complete models of the same reading, the plant with an unbiased sensor and the plant with a bias entry on the reading row, and its answer is a ratio of their evidences rather than a property of either posterior; when every sensor of a larger plant is a candidate, the same comparison run over the candidates is a diagnosis. Chapter 22, in Volume II, answers it, and the residual check of Section 1.4.3 takes a first, coarse look at it. What happens if the cooling partially fails is a question about a transient that the steady-state graph cannot represent, and Section 1.6.5 states what the graph would need for it.

1.1.7 The evidence as rows (step 2)

Step 2 attaches the reading and the handbook value to the graph as rows on the state $\mathbf{z} = (T, g)^{\mathsf{T}}$, and extends the graph with the operating condition the fourth question names. The three rows, the physics row of Section 1.1.4 among them, are Gaussian, and each is written normalized, with a residual and a Jacobian row.

The modeling choices. One choice makes the physics row, and it buys something and can fail. The energy balance is taken as exact, the limit $\sigma_p \to 0$ of the physics row: this is what confines the posterior to the physics curve, so that a reading of the temperature informs the conductance, and it fails if heat leaves the hall by a mechanism the single conductance does not lump or if the hall is not in steady state (Section 1.6.5); Section 1.4.1 checks that the tolerance used has reached the limit.

The reading row. The reading of Section 1.1.5 becomes the normalized observation residual

$$r_{\text{obs}}(\mathbf{z}) = \frac{T - y_T}{\sigma_T}, \tag{1.6}$$

which is dimensionless: the difference between the candidate temperature and the reading, in units of the sensor's standard deviation. At $T = 23.2°\text{C}$ it is zero; at the design temperature $T = 23°\text{C}$ it is $(23 - 23.2)/0.5 = -0.4$. The row carries the probability of the reading. Given the state, the reading differs from T only by the error, so its density is the normal density of ε evaluated at $y_T - T$:

$$p(y_T \mid \mathbf{z}) = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\frac{(y_T - T)^2}{2\sigma_T^2}\right] = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\tfrac{1}{2}\, r_{\text{obs}}(\mathbf{z})^2\right].$$

Read as a function of the state for the fixed reading, this is the *likelihood* of the state. A candidate state with a small $|r_{\text{obs}}|$ makes the reading probable; the ratio of the likelihoods of $T = 23°\text{C}$ and $T = 23.2°\text{C}$ is $\mathrm{e}^{-0.4^2/2} = \mathrm{e}^{-0.08} = 0.923$, so the reading alone hardly distinguishes the two. The likelihood does not depend on g: the sensor sees the temperature, and it informs the conductance only through the physics row that ties the two together.

The prior row. The handbook value becomes a row of the same form. Written as a normal distribution on the log-conductance, $g \sim \mathcal{N}(g_0, \sigma_g^2)$, it is the row $(g - g_0)/\sigma_g$: a handbook value is treated as a reading of g taken before the test. The handbook conductance of $7.5\,\text{kW/K}$ gives $g_0 = \ln(7.5/6) = 0.2231$, and $\sigma_g = 0.3$ states that the true conductance is within a factor $\mathrm{e}^{0.3} = 1.35$ of it with probability of about 0.68. The prior is written $p(g)$ in general; its form is a modeling choice that must be stated and justified, and Section 1.4.4 tests how much this one weighs.

Table 1.3: The three kinds of row in the equipment hall model, with $G = G_{\text{ref}}\,\mathrm{e}^{g}$.

Row	Equation	Residual	Role
Physics	$G\,(T - T_a) - Q = 0$	$G_{\text{ref}}\,\mathrm{e}^{g}(T - T_a) - Q$	constrains what is physically possible
Reading	$y_T = T + \varepsilon$	$(T - y_T)/\sigma_T$	provides imperfect information about the state
Prior	$g \sim \mathcal{N}(g_0, \sigma_g^2)$	$(g - g_0)/\sigma_g$	carries knowledge held before the readings

Three kinds of row. The model now holds three very different kinds of relation. Table 1.3 sets them side by side. Each row of the table is one kind of relation; the columns give its equation for the hall, its residual and its role.

The distinction between the first two rows is foundational. The physical equation constrains the system: it says which combinations of temperature and conductance can occur at all. The reading provides evidence about the unknown quantities: it says which of the possible combinations are plausible in light of what was measured. The prior row has the form of a reading row and the role of knowledge held in advance. A probabilistic inference method combines the three kinds of relation to obtain a distribution over the plausible states.

Rows and weights. Chapter 3 shows that, for normal errors, the negative logarithm of the posterior is, up to a constant, the objective $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\sum_k w_k\, r_k(\mathbf{z})^2$, a weighted sum of the squared rows with weights $w_k = 1/\sigma_k^2$. A reading row may be written either normalized, as in equation (1.6), with unit weight, or in kelvin, as $T - y_T$ with weight $1/\sigma_T^2$; both put the same term $(T - y_T)^2/\sigma_T^2$ into the objective. Every row of this chapter is written normalized: the physics row in the form (1.5) with the tolerance σ_p, the reading row (1.6) and the prior row, so that every row counts in standard deviations and none dominates because of its unit.

The scenario node. The fourth question doubles the electrical load from 30 kW to 60 kW, an operating condition at which the hall has never run. A second operating condition is added to the graph, sharing the same unknown conductance. Figure 1.3 draws it. The upper row is the test condition, at which the hall has been operated and observed: the free node at $Q_1 = Q = 30\,\text{kW}$ with temperature T, its path to the pinned ambient, and its sensor. The lower row is the full-load condition, a scenario node at $Q_2 = 60\,\text{kW}$ with temperature T_2 and its own path to the same ambient. The vertical dashed line between the two paths marks that they carry one and the same conductance. The scenario node has no sensor: it is in the graph only because the question asks about it.

The two operating conditions satisfy the energy balance of equation (1.1) each at its own load,

$$Q_1 = G\,(T - T_a), \tag{1.7}$$

$$Q_2 = G\,(T_2 - T_a), \tag{1.8}$$

with $Q_1 = Q$. The two temperatures are different, but the conductance is shared, and that connection is what makes the second condition predictable from the first. The scenario node adds no evidence and no free quantity: by equation (1.8) its temperature is fixed by the conductance, $T_2 = T_a + (Q_2/G_{\text{ref}})\,\mathrm{e}^{-g}$, a function of the state $\mathbf{z} = (T, g)^{\mathsf{T}}$, which keeps its two entries and its three rows. Chapter 4 carries T_2 instead as a third entry with a second physics row, the same formulation; with the same rows it gives the same answers.

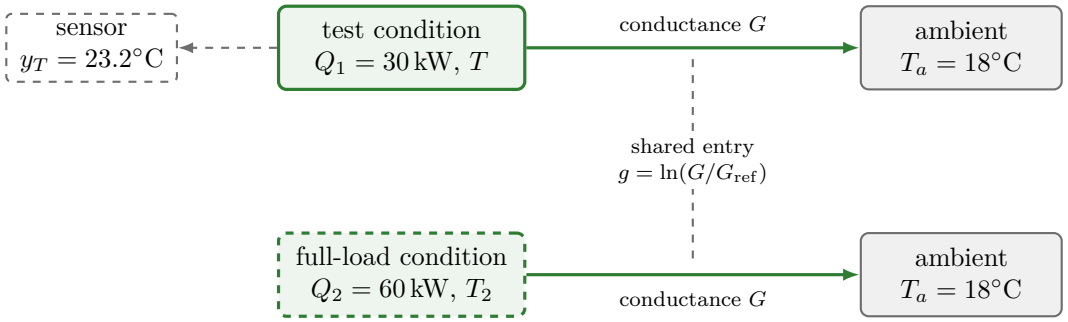


Figure 1.3: Two operating conditions of the equipment hall linked through the same physical conductance. Free: the test-condition temperature T and the shared log-conductance g, the two entries of the state; the full-load temperature T_2 (dashed node) is a function of g. Pinned: the ambient temperature and the two loads. Observed: T only, through the reading y_T; the scenario node is unobserved and enters the graph because the question asks about it.

Which row acts on which entry. The three rows act on the two entries of the state in a pattern fixed by the graph, before any number is computed. The physics row ties the temperature to the conductance and involves both; the sensor sees the temperature and not the conductance; the handbook speaks about the conductance and not the temperature; the scenario node adds no row. The Jacobian of the rows, the matrix of their derivatives with respect to the entries, can therefore be nonzero only where marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & * \\ * & 0 \\ 0 & * \end{bmatrix} \begin{matrix} \text{physics row} \\ \text{reading row,} \\ \text{prior row} \end{matrix} \qquad \text{columns } z_1 = T,\ z_2 = g. \tag{1.9}$$

The zeros are structural: they hold at every state. Only the physics row has entries in both columns, and it is the one row through which a reading of the temperature reaches the conductance.

The factor graph of the hall. The pattern (1.9) is what a factor graph (Section 1.1.3) draws. Figure 1.4 draws it for the hall with its scenario node. The two solid circles are the entries of the state, T and g. The black squares are the rows: the physics row of equation (1.5), linked to both entries; the reading row, linked to T alone; and the prior row, linked to g alone. The dashed circle is the full-load temperature T_2, joined to g by a dashed line: it is a function of g through equation (1.8), not an entry, and no row acts on it. The figure makes visible what the physical graph only implies: T_2 has no reading of its own and is connected to the evidence only through g, and g is connected to the reading only through the physics row. The physical drawing has two ambient boxes and two paths; the factor graph has no node for the ambient at all, since T_a is pinned and appears only as a constant inside the physics row.

The residual vector. The state is the column vector of equation (1.3),

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} T \\ g \end{bmatrix},$$

with T in °C and g dimensionless. Three rows act on it, the physics row (1.5), the reading row (1.6) and the prior row, each divided by its standard deviation, so that every row counts in standard deviations and none dominates because of its unit. Stacked into one residual vector,

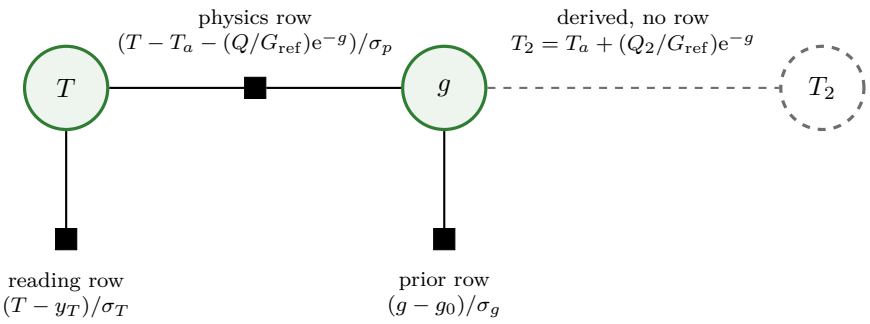


Figure 1.4: Factor graph of the equipment hall with its scenario node. Solid circles: the entries of the state, T and g. Squares: the rows of the model, each linked to the entries it involves, the pattern (1.9) of the Jacobian. Dashed: the full-load temperature T_2, a function of g with no row of its own, which reaches the evidence only through the shared g.

the three rows are

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} r_1(\mathbf{z}) \\ r_2(\mathbf{z}) \\ r_3(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} (T - T_a - (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (T - y_T)/\sigma_T \\ (g - g_0)/\sigma_g \end{bmatrix}, \tag{1.10}$$

with $T_a = 18°\mathrm{C}$, $Q = 30\,\mathrm{kW}$, $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the reading $y_T = 23.2°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$, and the handbook prior $g_0 = \ln(7.5/6) = 0.2231$ with $\sigma_g = 0.3$. Each entry of $\mathbf{r}$ is one row: a dimensionless number stating by how many standard deviations a candidate state violates one relation. The objective is half the sum of their squares,

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\,\mathbf{r}(\mathbf{z}) = \tfrac{1}{2}\sum_{k=1}^{3} r_k(\mathbf{z})^2, \tag{1.11}$$

which Chapter 3 shows to be minus the logarithm of the posterior up to a constant, so that the most probable state $\hat{\mathbf{z}}$, the maximum a posteriori (MAP) state, is the minimizer of $\mathcal{L}$. Conservation of energy is exact in this model, and an exact constraint is the limit $\sigma_p \to 0$ of the physics row (Section 1.1.4). The computation takes $\sigma_p = 0.01\,\mathrm{K}$, fifty times tighter than the sensor, and Section 1.4.1 checks that the limit has been reached.

The Jacobian. The Jacobian $\mathbf{J}$ collects the derivative of every row with respect to every entry of the state: its entry in row k and column i is $\partial r_k/\partial z_i$, so it has one row per row of the model and one column per entry of the state, here three by two. Differentiating equation (1.10) entry by entry,

$$\mathbf{J}(\mathbf{z}) = \frac{\partial \mathbf{r}}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_p & (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 1/\sigma_T & 0 \\ 0 & 1/\sigma_g \end{bmatrix}, \tag{1.12}$$

the pattern (1.9) with its entries filled in. Each entry is a sensitivity in standard deviations: how far one row moves when one entry of the state moves by one unit. Only the entry in the first row and second column depends on the state; the others are constants. The first row is the physics row's own sensitivities in its physical unit, rescaled. Differentiating the kilowatt form of equation (1.4),

$$\frac{\partial r_{\mathrm{phys}}}{\partial T} = G_{\mathrm{ref}}\,\mathrm{e}^{g} = G, \qquad \frac{\partial r_{\mathrm{phys}}}{\partial g} = G_{\mathrm{ref}}\,\mathrm{e}^{g}\,(T - T_a) = G\,(T - T_a),$$

in kW/K and kW; on the zero set the second equals the load Q by equation (1.1). A one-kelvin error in the temperature unbalances the hall by G kilowatts, and a unit change in g, a factor e

in the conductance, by about Q kilowatts to first order. Divided by G, to reach the temperature form, and by σ_p, the pair becomes $(1/\sigma_p,\ (Q/G)/\sigma_p)$, the first row of equation (1.12) on the zero set.

1.2 The posterior of the state (step 3)

All the questions of Table 1.2 are read from one posterior: the posterior of the state that the physics row, the reading row and the prior row of step 2 give, approximately the Laplace normal of (T, g) and exactly a density along the physics curve. This section derives it and computes it for the hall; Section 1.3 reads every answer from it.

Four objects of the fit recur in the specialized forms, and they are named here by their roles. The *MAP state* $\hat{\mathbf{z}} = (\hat{T}, \hat{g})^{\mathsf{T}}$, the maximum a posteriori state, is the minimum of the objective that the three rows put on the state, the most probable state; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ is the curvature that the rows give that objective at the minimum, with $\mathbf{J}$ the 3×2 Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma}$, a 2×2 matrix, is the inverse of that curvature after a term from the physics row's own curvature is added to it, and it makes the posterior of the state approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$; and a derived quantity f is carried by its *gradient* $\mathbf{g} = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$ (bold, to tell it from the entry g), with the variance $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. Section 1.2.1 derives the four from the rows. Beside them stands the exact posterior that this problem admits in the limit $\sigma_p \to 0$: on the physics curve the state has one free direction, and the posterior is the density of one entry, evaluated by quadrature. The answers are reported by the Laplace route, a normal distribution centered on the MAP state: this is what lets every answer be read from $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ by operations that serve a plant of any size, and it misleads where the posterior is skewed along the curved ridge, which Section 1.4.2 measures against the exact posterior. The chapter reports the Laplace values, the route that carries over unchanged to plants of many entries, sets the exact value beside each answer with the gap, and collects the comparison in Section 1.4.

1.2.1 The posterior from the rows

The answers use the objects of one posterior: the MAP state $\hat{\mathbf{z}}$, the covariance $\mathbf{\Sigma}$ and the Jacobian $\mathbf{J}$ of the rows. The Jacobian is written out in Section 1.1.7; the other two are derived here for the hall. Chapter 3 derives these objects from the probability rules for Gaussian rows over any state, and Chapter 4 dissects the same objects in full on the hall with its scenario node carried as an entry, including the Laplace approximation and its curvature term. Here each is written out once for a reader meeting it for the first time.

The posterior. Four ingredients are in hand: the unknown state $\mathbf{z} = (T, g)^{\mathsf{T}}$; the observation $y_T = 23.2°\mathrm{C}$; the physical relation $Q = G\,(T - T_a)$; and the prior information $p(g)$. What is ultimately wanted is

$$p(\mathbf{z} \mid \mathbf{y}), \tag{1.13}$$

the *posterior distribution* of the state, conditional on the evidence $\mathbf{y}$, here the reading y_T and the handbook value, and on the assumed physical model. It states which values of the state remain plausible after the physics, the measurement and the prior information have been combined. Chapter 3 derives Bayes' theorem, by which the posterior is proportional to the likelihood times the prior, $p(\mathbf{z} \mid \mathbf{y}) \propto p(y_T \mid \mathbf{z})\, p(\mathbf{z})$; here the reading row plays the part of the likelihood, and the physics row and the prior row on g together play the part of the prior $p(\mathbf{z})$.

The normal equations and the Gauss–Newton step. Near a state $\mathbf{z}$, each row is replaced by its tangent, $\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}) \approx \mathbf{r}(\mathbf{z}) + \mathbf{J}\,\Delta\mathbf{z}$, and the objective by $\frac{1}{2}\,\|\mathbf{r} + \mathbf{J}\,\Delta\mathbf{z}\|^2$, a quadratic in the

END OF THE PREVIEW

Reliability Engineering by Derivation

These are the opening pages. The complete book runs 672 pages.

The full PDF opens with a token. Write to **rmarquez123@gmail.com**, name the book, and I will send you one.

One token opens every book on the site.

Ricardo Marquez, Ph.D.

ricardo-marquez.com