Kalman Filtering with State Equality Constraints

Daniel J. Simon  
*Cleveland State University*, d.j.simon@csuohio.edu

Tien Li Chia  
*ControlSoft, Inc.*, tchia@controlsoftinc.com

Follow this and additional works at: https://engagedscholarship.csuohio.edu/enece_facpub

Part of the Electrical and Computer Engineering Commons

**How does access to this work benefit you? Let us know!**

**Original Citation**

**Repository Citation**
Simon, Daniel J. and Chia, Tien Li, "Kalman Filtering with State Equality Constraints" (2002). Electrical Engineering & Computer Science Faculty Publications. 158. https://engagedscholarship.csuohio.edu/enece_facpub/158

This Article is brought to you for free and open access by the Electrical Engineering & Computer Science Department at EngagedScholarship@CSU. It has been accepted for inclusion in Electrical Engineering & Computer Science Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.
Kalman filtering with inequality constraints for turbofan engine health estimation

D. Simon and D.L. Simon

Abstract: Kalman filters are often used to estimate the state variables of a dynamic system. However, in the application of Kalman filters some known signal information is often either ignored or dealt with heuristically. For instance, state-variable constraints (which may be based on physical considerations) are often neglected because they do not fit easily into the structure of the Kalman filter. Thus, two analytical methods to incorporate state-variable inequality constraints into the Kalman filter are now derived. The first method is a general technique that uses hard constraints to enforce inequalities on the state-variable estimates. The resultant filter is a combination of a standard Kalman filter and a quadratic programming problem. The second method uses soft constraints to estimate those state variables that are known to vary slowly with time. (Soft constraints are constraints that are required to be approximately satisfied rather than exactly satisfied.) The incorporation of state-variable constraints increases the computational effort of the filter but significantly improves its estimation accuracy. The improvement is proven theoretically and simulations are used to show that the proposed algorithms can provide an improved performance over unconstrained Kalman filtering.

1 Introduction

For linear dynamic systems with white process and measurement noises, the Kalman filter is known to be an optimal estimator. However, in the application of Kalman filters there is often available model or signal information that is either ignored or dealt with heuristically [1]. We intend to derive ways to modify the Kalman filter state estimate such that known inequality constraints are satisfied by the state-variable estimates.

The first method presented to enforce inequality constraints on the state-variable estimates uses hard constraints. It is based on a generalisation of the approach presented in [2], which dealt with the incorporation of state-variable equality constraints in the Kalman filter. Inequality constraints are inherently more complicated than equality constraints, but standard quadratic programming results can be used to solve the Kalman filter problem with inequality constraints. At each time step of the constrained Kalman filter, we solve a quadratic programming problem to obtain the constrained state estimate. A family of constrained state estimates is obtained, where the weighting matrix of the quadratic programming problem determines which family member forms the desired solution. We state, on the basis of [2], that the constrained estimate has several important properties. The constrained state estimate is unbiased and has a smaller error covariance than the unconstrained estimate. We show which member of all possible constrained solutions has the smallest error covariance. We also show the one particular member that is always (i.e. at each time step) closer to the true state than the unconstrained estimate.

The second method to enforce inequality constraints uses soft constraints via a penalty term in an optimisation problem. This prevents the state estimate from changing too rapidly. It essentially smooths the unconstrained Kalman filter estimate when the state variables are known to vary slowly with time. It is shown that the constrained state estimate is unbiased, approaches the unconstrained estimate as the time approaches infinity, and (under certain special conditions) is equal to the running average of the unconstrained estimate.

The application considered in this study is turbofan engine health parameter estimation [3]. The performance of gas turbine engines deteriorates over time. This deterioration reduces the fuel economy of the engine. Airlines periodically collect engine data in order to evaluate the health of the engine and its components. The health evaluation is then used to determine maintenance schedules. Reliable health evaluations are used to anticipate future maintenance needs. This offers the benefits of improved safety and reduced operating costs. The money-saving potential of such health evaluations is substantial, but only if the evaluations are reliable. The data used to perform health evaluations are typically collected during flight and later transferred to ground-based computers for post-flight analysis. Data are collected each flight at the same engine operating points and corrected to account for variability in ambient conditions. Typically, data are collected for a period of about 3 s at a rate of about 10 or 20 Hz. Various algorithms have been proposed to estimate engine health parameters, such as weighted least squares [4], expert systems [5], Kalman filters [6], neural networks [6], and genetic algorithms [7].

We apply constrained Kalman filtering to estimate engine component efficiencies and flow capacities, which are
referred to as health parameters. We can use our knowledge of the physics of the turbofan engine in order to obtain a
dynamic model [8, 9]. The health parameters that we try
to estimate can be modelled as slowly-varying biases. The
state vector of the dynamic model is augmented to
include the health parameters, which are then estimated
with a Kalman filter [10]. The model formulation in this
study is similar to previous work [11]. However, [11] was
limited to a three-state dynamic model and two health
parameters, whereas this present work includes a more
complete 16-state model and eight health parameters. In
addition, we have some a priori knowledge of the engine’s
health parameters: we know that they never improve.
Engine health always degrades over time, and we can
incorporate this information into state constraints to
improve our health parameter estimation. (This is assuming
that no maintenance or engine overhaul is performed.)
This is similar to the probabilistic approach to turbofan
prognostics proposed in [12]. It should be emphasised
that in this study we are confining the problem to the
estimation of engine health parameters in the sole presence
degradation. There are specific engine cases that can
result in abrupt shifts in filter estimates, possibly even
indicating an apparent improvement in some engine com-
ponents. An actual engine performance monitoring system
would need to include additional logic to detect and
isolate such faults.

2 Kalman filtering

This Section reviews standard (unconstrained) state
estimation via the Kalman filter and some important properties
of the filter that will be used later in this study. The results
and notation are taken from [13]. Consider the discrete
linear time-invariant system given by:

\[ x_{k+1} = Ax_k + Bu_k + w_k \]
\[ y_k = Cx_k + e_k \]

where \( k \) is the time index, \( x \) is the state vector, \( u \) is the
known control input, \( y \) is the measurement, and \( w_k \) and
\( e_k \) are noise input sequences. The problem is to find an
estimate \( \hat{x}_{k+1} \) of \( x_{k+1} \) given the measurements \( y_0, y_1, \ldots, y_k \).
We will use the symbol \( Y_k \) to denote the column
vector that contains the measurements \( \{y_0, y_1, \ldots, y_k\} \).
We assume that the following standard conditions are satisfied

\[ E[x_0] = \bar{x}_0 \]
\[ E[w_k] = E[e_k] = 0 \]
\[ E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_0 \]
\[ E[w_k w_m^T] = Q \delta_{km} \]
\[ E[e_k e_m^T] = R \delta_{km} \]
\[ E[w_k e_m^T] = E[e_k w_m^T] = 0 \]
\[ E[x_k w_m^T] = 0 \quad (m \geq k) \]

where \( E[\cdot] \) is the expectation operator, \( \bar{x} \) is the expected
value of \( x \), and \( \delta_{km} \) is the Kronecker delta function
(\( \delta_{km} = 1 \) if \( k = m, 0 \) otherwise). \( Q \) and \( R \) are positive semi-
definite covariance matrices. The Kalman filter equations
are given by:

\[ K_k = A \Sigma_k C^T (C \Sigma_k C^T + R)^{-1} \]
\[ \hat{x}_{k+1} = A \hat{x}_k + Bu_k + K_k (y_k - C \hat{x}_k) \]
\[ \Sigma_{k+1} = (A \Sigma_k - K_k C \Sigma_k) A^T + Q \]

where the filter is initialised with \( \hat{x}_0 = \bar{x}_0 \), and \( \Sigma_0 \) given
above. It can be shown [13] that the Kalman filter has
several attractive properties. For instance, if \( x_0, [w_k], \) and
\([e_k]\) are jointly Gaussian, the Kalman filter estimate \( \hat{x}_{k+1} \)
is the conditional mean of \( x_{k+1} \) given the measurements
\( Y_k \); i.e. \( \hat{x}_{k+1} = E[x_{k+1} | Y_k] \). Even if \( x_0, [w_k], \) and \([e_k]\)
are not jointly Gaussian, the Kalman filter estimate is the best
affine estimator given the measurements \( Y_k \); i.e. of all esti-
mates of \( x_{k+1} \) that are of the form \( FY_k + g \) (where \( F \) is
a fixed matrix and \( g \) is a fixed vector), the Kalman filter
estimate is the one that minimise the covariance of the
estimation error. It can be shown [13, pp. 92 ff.] that the
Kalman filter estimate (i.e. the minimum variance estimate)
can be given by:

\[ \hat{x}_{k+1} = \bar{x}_{k+1} = \bar{x}_{k+1} + \Sigma_{k+1} \Sigma_{yy}^{-1} (Y_k - \bar{Y}_k) \]

where \( \bar{x}_{k+1} \) is the mean of \( x_{k+1} \), \( \Sigma_{k+1} \) is the covariance matrix
of \( x_{k+1} \) and \( Y_k \), \( \Sigma_{yy} \) is the covariance matrix of \( Y_k \), and \( \bar{x}_{k+1} \)
is the conditional mean of \( x_{k+1} \) given the measurements \( Y_k \).
In addition, from [13, p. 93] we know that the Kalman filter
estimate \( \hat{x}_{k+1} \) and \( Y_k \) are jointly Gaussian, in which case
\( \hat{x}_{k+1} \) is conditionally Gaussian given \( Y_k \). The conditional
probability density function of \( x_{k+1} \) given \( Y_k \) is:

\[ P(x|Y) = \frac{\exp[-(x - \bar{x})^T \Sigma^{-1} (x - \bar{x})/2]}{(2\pi)^{n/2} \Sigma_x^{1/2}} \]

where \( n \) is the dimension of \( x \) and

\[ \Sigma = \Sigma_x - \Sigma_{yx} \Sigma_{yy}^{-1} \Sigma_{yx} \]

The Kalman filter estimate is that value of \( x \) that maximises
the conditional probability density function \( P(x|Y) \), and \( \Sigma \)
is the covariance of the Kalman filter estimation error.

3 Kalman filtering with hard inequality
constraints

This Section modifies the well known state estimate of the
previous Section so that the estimate satisfies linear
inequality constraints that are known to exist among the
state components. Also, several important properties of
the constrained filter are discussed. Consider the dynamic
system of (1) where we are given the additional constraint

\[ Dx_k \leq d_k \]

where \( D \) is a known \( s \times n \) constant matrix, \( s \) is the number
of constraints, \( n \) is the number of state variables, and \( s \leq n \).
It is assumed in this study that \( D \) is full rank, i.e. that \( D \) has
rank \( s \). This is an easily satisfied assumption. If \( D \) is not full
rank that means we have redundant state constraints. In that
case we can simply remove linearly-dependent rows from \( D \)
(i.e. remove redundant state constraints) until \( D \) is full rank.
Three different approaches to the constrained state esti-
mation problem are given in this Section. The time index
\( k \) is omitted in the remainder of this Section for ease of
notation.
3.1 The maximum probability method

In this Section we derive the constrained Kalman filtering problem by using a maximum probability method. From [13, pp. 93 ff.] we know that the Kalman filter estimate is that value of \(x\) that maximises the conditional probability density function \(P(x|y)\), which is given in (13). The constrained Kalman filter can be derived by finding an estimate \(\hat{x}\) such that the conditional probability \(P(\hat{x}|y)\) is maximised and \(\hat{x}\) satisfies the constraint (15). Maximising \(P(\hat{x}|y)\) is the same as maximising its natural logarithm. So the problem we want to solve can be given by:

\[
\max \ln P(\hat{x}|y) \iff \min(\bar{x} - \hat{x})^T \Sigma^{-1}(\bar{x} - \hat{x})
\]

such that \(D\hat{x} = d\) \hspace{1cm} (16)

Using the fact that the unconstrained state estimate \(\bar{x} = \bar{x}\) (the conditional mean of \(x\)), we rewrite the above equation as:

\[
\min(\bar{x}^T \Sigma^{-1}\bar{x} - 2\bar{x}^T \Sigma^{-1}\hat{x}) \quad \text{such that} \quad D\hat{x} = d
\]

Note that this problem statement depends on the conditional Gaussian nature of \(\bar{x}\), which in turn depends on the Gaussian nature of \(x_0\), \(\{w_k\}\), and \(\{e_k\}\) in (1).

3.2 The mean-square method

In this Section we derive the constrained Kalman filtering problem by using a mean-square minimisation method. We seek to minimise the conditional mean-square error subject to the state constraints:

\[
\min E(\|x - \hat{x}\|^2|y) \quad \text{such that} \quad D\hat{x} = d
\]

where \(\|\cdot\|\) denotes the vector two-norm. If we assume that \(x\) and \(y\) are jointly Gaussian, the mean-square error can be written as

\[
E(\|x - \hat{x}\|^2|y) = \int (x - \hat{x})^T(x - \hat{x})P(x|y)dx
\]

\[
= x^T xP(x|y)dx - 2\bar{x}^T \int xP(x|y)dx + \bar{x}^T \bar{x}
\]

Noting that the Kalman filter estimate is the conditional mean of \(x\), i.e.:

\[
\hat{x} = \int xP(x|y)dx
\]

we formulate the first-order conditions necessary for a minimum as:

\[
\min(\hat{x}^T \hat{x} - 2\bar{x}^T \hat{x}) \quad \text{such that} \quad D\hat{x} = d
\]

Again, this problem statement depends on the conditional Gaussian nature of \(\hat{x}\), which in turn depends on the Gaussian nature of \(x_0\), \(\{w_k\}\), and \(\{e_k\}\) in (1).

3.3 The projection method

In this Section we derive the constrained Kalman filtering problem by directly projecting the unconstrained state estimate \(\hat{x}\) onto the constraint surface. That is, we solve the problem:

\[
\min(\bar{x} - \hat{x})^T W(\bar{x} - \hat{x}) \quad \text{such that} \quad D\hat{x} = d
\]

where \(W\) is any symmetric positive definite weighting matrix. This problem can be rewritten as:

\[
\min(\bar{x}^T W\bar{x} - 2\bar{x}^T W\hat{x}) \quad \text{such that} \quad D\hat{x} = d
\]

The constrained estimation problems derived by the maximum probability method (17) and the mean-square method (22) can be obtained from this equation by setting \(W = \Sigma^{-1}\) and \(W = I\) respectively. Note that this derivation of the constrained estimation problem does not depend on the conditional Gaussian nature of \(\bar{x}\); i.e. \(x_0\), \(\{w_k\}\), and \(\{e_k\}\) in (1) are not assumed to be Gaussian.

3.4 The solution of the constrained state estimation problem

The problem defined by (24) is known as a quadratic programming problem [14, 15]. There are many algorithms to solve quadratic programming problems, almost all of which fall into the category known as active set methods. An active set method uses the fact that it is only those constraints that are active at the solution of the problem that are significant in the optimality conditions. Assume that \(t\) of the \(s\) inequality constraints are active at the solution of (24), and denote by \(D\) and \(d\) the \(t\) rows of \(D\) and \(t\) elements of \(d\) corresponding to the active constraints. If the correct set of active constraints was known \textit{a priori} then the solution of (24) would also be a solution of the equality-constrained problem:

\[
\min(\bar{x}^T W\bar{x} - 2\bar{x}^T W\hat{x}) \quad \text{such that} \quad D\hat{x} = d
\]

This shows that the inequality-constrained problem defined by (24) is equivalent to the equality-constrained problem defined by (25). The equality-constrained problem was discussed in [2], and so those results can be used to investigate the properties of the inequality-constrained problem.

3.5 Properties of the constrained state estimate

In this Section we examine some of the statistical properties of the constrained Kalman filter. We use \(\bar{x}\) to denote the state estimate of the unconstrained Kalman filter, and \(x\) to denote the state estimate of the constrained Kalman filter as given by (24), recalling that (17) and (22) are special cases of (24).

\textbf{Theorem 1:} The solution \(\hat{x}\) of the constrained state estimation problem given by (24) is an unbiased state estimator for the system (1) for any symmetric positive definite weighting matrix \(W\). That is:

\[
E(\hat{x}) = E(x)
\]

\textbf{Theorem 2:} The solution \(\bar{x}\) of the constrained state estimation problem given by (24) with \(W = \Sigma^{-1}\), where \(\Sigma\) is the covariance of the unconstrained estimate given in (11) and (14), has an error covariance that is less than or equal to that of the unconstrained state estimate. That is:

\[
\text{cov}(x - \bar{x}) \leq \text{cov}(x - \hat{x})
\]
state-variable constraints. Therefore, the standard Kalman filter is no longer the minimum variance filter, and we can do better with the constrained Kalman filter.

**Theorem 3:** Among all the constrained Kalman filters resulting from the solution of (24), the filter that uses $W = \Sigma^{-1}$ has the smallest estimation error covariance. That is:
\[
\text{cov}(\tilde{x}_{\Sigma^{-1}}) \leq \text{cov}(\tilde{x}_{W}) \quad \text{for all } W
\] (28)

**Theorem 4:** The solution $\tilde{x}$ of the constrained state estimation problem given by (24) with $W = I$ satisfies the inequality:
\[
\|x_k - \tilde{x}_k\| \leq \|x_k - \hat{x}_k\| \quad \text{for all } k
\] (29)
where $\|\cdot\|$ is the vector two-norm and $\hat{x}$ is the unconstrained Kalman filter estimate.

**Theorem 5:** The error of the solution $\tilde{x}$ of the constrained state estimation problem given by (24) with $W = I$ is smaller than the unconstrained estimation error in the sense that:
\[
\text{tr}[\text{cov}(\tilde{x})] \leq \text{tr}[\text{cov}(\hat{x})]
\] (30)
where tr[·] indicates the trace of a matrix, and cov(·) indicates the covariance matrix of a random vector.

The above theorems all follow from the equivalence of (24) and (25), and the proofs presented in [2]. We note that if any of the $s$ constraints are active at the solution of (24), then strict inequalities hold in the statements of theorems 2–5. The only time that equalities hold in the theorems is if there are no active constraints at the solution of (24); that is, if the unconstrained Kalman filter satisfies the inequality constraints.

### 4 Kalman filtering with soft inequality constraints

In this section we are interested in obtaining a Kalman-filtrbased state estimate for state variables which we know a priori vary slowly with time. Since we are concerned with using the Kalman filter as a parameter estimator, we will assume for this problem that the $A$ matrix in (1) is the identity matrix and the $B$ matrix is zero. With this in mind, we can use the results of the previous section, especially (23), to formulate a Kalman-filter-based estimate as follows:
\[
\min_{\tilde{x}_k} (\tilde{x}_k - \hat{x}_k)^T W (\tilde{x}_k - \hat{x}_k)
\]
where, as before, $W$ is a constant symmetric positive definite weighting matrix. This is a type of regularisation; that is, some additional structure is incorporated into the Kalman filter estimate [16–18]. The above problem can be formulated as
\[
\min_{\tilde{x}_k} (\tilde{x}_k - \hat{x}_k)^T W (\tilde{x}_k - \hat{x}_k) + (\tilde{x}_k - \hat{x}_k-1)^T V_k (\hat{x}_k - \hat{x}_k-1)
\] (32)
where $V_k$ is a (possibly time-varying) symmetric positive definite weighting matrix that balances the desire for a close approximation to $\tilde{x}$ and smooth estimate $\hat{x}$. The solution to the above problem is:
\[
\tilde{x}_0 = E[x_0]
\]
\[
\tilde{x}_k = (W + V_k)^{-1} (W \hat{x}_k + V_k \hat{x}_k-1)
\] (33)
Since $W$ and $V_k$ are both positive definite, we know that $(W + V_k)^{-1}$ exists.

**Theorem 6:** Assume (as stated above) that $A = I$ and $B = 0$ in (1). Then the solution $\tilde{x}$ of the constrained state estimation problem given by (33) is an unbiased state estimator for the system (1) for any symmetric positive definite weighting matrices $W$ and $V_k$. That is:
\[
E(\tilde{x}) = E(x)
\] (34)

**Proof:** The theorem can be proven by induction. Since $A = I$ and $B = 0$ we know that $E[x_k] = x_0$ for all $k$. We therefore know from (33) that $\tilde{x}_0 = \hat{x}_0$. From (33) with $k = 0$ we see that $E[\tilde{x}_1] = x_0$. We repeat this process to show that $E[\tilde{x}_k] = E[x_k] = \hat{x}_0$ for all $k$.

**Theorem 7:** Assume (as stated above) that $A = I$ and $B = 0$ in (1). Further assume that $w_k = 0$ in (1) (since we are trying to estimate constant parameters). Then the constrained state estimate $\tilde{x}$ approaches the unconstrained estimate $\hat{x}$ in the limit as time goes to infinity. That is:
\[
\lim_{k \to \infty} \tilde{x}_k = \lim_{k \to \infty} \hat{x}_k
\] (35)

**Proof:** We see from (9)–(11) that, under the conditions stated here, $K_k \to 0$ as $k \to \infty$. Therefore, $\hat{x}_k$ approaches a constant value as $k \to \infty$. From (33) we see that, in steady-state conditions:
\[
\tilde{x} = (W + V_k)^{-1} (W \hat{x} + V_k \hat{x})
\]
\[
\implies \tilde{x} = (I - (W + V_k)^{-1} V_k)^{-1} (W + V_k)^{-1} W \hat{x}
\]
\[
= (I + W^{-1} V_k)(W + V_k)^{-1} W \hat{x}
\] (36)
where the last equality follows from the matrix inversion lemma. Premultiplying both sides of the above equation by $W$ we obtain $W \tilde{x} = W \hat{x}$, so if $W$ is invertible (which it is, since we are assuming in this Section that $W$ is positive definite), we obtain $\tilde{x} = \hat{x}$ (in steady-state conditions). Note that the theorem is true even if $V_k$ does not approach a steady-state value as $k \to \infty$.

**Theorem 8:** If $V_k = (k - 1)W$ in (33) then $\tilde{x}_k$ is the running average of $\hat{x}_k$.

**Proof:** The running average of $\hat{x}_k$ is defined as:
\[
X_k = \frac{1}{k} \sum_{i=1}^{k} \hat{x}_i
\] (37)
which implies that
\[
X_{k+1} = \frac{1}{k + 1} (\hat{x}_{k+1} + k \hat{x}_k)
\] (38)
Now if $V_k = (k - 1)W$ then (33) shows that:
\[
\tilde{x}_{k+1} = [(k + 1)W]^{-1} (W \hat{x}_{k+1} + kW \hat{x}_k)
\]
\[
= \frac{1}{k + 1} (\hat{x}_{k+1} + k \hat{x}_k)
\] (39)
which is exactly the running average shown in (38).
5 Turbofan engine health monitoring

Figure 1 shows a schematic representation of a turbofan engine. A single inlet supplies airflow to the fan. Air leaving the fan separates into two streams: one stream passes through the engine core, and the other stream passes through the annular bypass duct. The fan is driven by the low-pressure turbine. The air passing through the engine core moves through the compressor, which is driven by the high-pressure turbine. Fuel is injected into the main combuster and burned to produce hot gas to drive the turbines. The two air streams combine in the augmentor duct, where additional fuel is added to further increase the air temperature. The air leaves the augmentor through the nozzle, which has a variable cross-sectional area.

Various turbofan simulation packages have been proposed over the years [19–21]. The model used in this study is based on a gas turbine engine simulation software package called DIGTEM (digital turbofan engine model) [8, 22]. DIGTEM is written in Fortran and includes 16 state variables. It uses a backward difference integration scheme because the turbofan model contains time constants that differ by up to four-orders of magnitude.

The nonlinear equations used in DIGTEM can be found in [8, 9]. The time-invariant equations can be summarised as follows:

\[
\dot{x} = f(x, u, p) + w_1(t)
\]
\[
y = g(x, u, p) + e(t)
\]

(40)

Table 1: Turbofan states and nominal values

<table>
<thead>
<tr>
<th>State</th>
<th>Nominal value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-pressure turbine rotor speed</td>
<td>6140</td>
</tr>
<tr>
<td>High-pressure turbine rotor speed</td>
<td>9395</td>
</tr>
<tr>
<td>Compressor mass flow, kg s⁻¹</td>
<td>0.457</td>
</tr>
<tr>
<td>Combustor inlet temperature, K</td>
<td>965</td>
</tr>
<tr>
<td>Combustor mass flow, kg s⁻¹</td>
<td>0.264</td>
</tr>
<tr>
<td>High-pressure turbine inlet temperaure, K</td>
<td>1593</td>
</tr>
<tr>
<td>High-pressure turbine mass flow, kg s⁻¹</td>
<td>1.48</td>
</tr>
<tr>
<td>Low-pressure turbine inlet temperaure, K</td>
<td>1129</td>
</tr>
<tr>
<td>Low-pressure turbine mass flow, kg s⁻¹</td>
<td>1.79</td>
</tr>
<tr>
<td>Augmentor inlet temperature, K</td>
<td>790</td>
</tr>
<tr>
<td>Augmentor mass flow, kg s⁻¹</td>
<td>1.46</td>
</tr>
<tr>
<td>Nozzle inlet temperature, K</td>
<td>790</td>
</tr>
<tr>
<td>Duct fluid momentum, kg s⁻²</td>
<td>53.6</td>
</tr>
<tr>
<td>Augmentor fluid momentum, kg s⁻²</td>
<td>103</td>
</tr>
<tr>
<td>Duct mass flow, kg s⁻¹</td>
<td>4.52</td>
</tr>
<tr>
<td>Duct temperature, K</td>
<td>571</td>
</tr>
</tbody>
</table>

where x is the 16-element state vector, u is the six-element control vector, p is the eight-element vector of health parameters, and y is the 12-element vector of measurements. The noise term w₁(t) represents inaccuracies in the model, and e(t) represents measurement noise. The elements in these vectors are summarised in Tables 1–4, along with their values at the nominal operating point (x₀, u₀, p₀, y₀) considered in this study. Table 4 also shows typical signal-to-noise (SNR) ratios for the measurements, based on NASA experience and previously published data [23]. Sensor dynamics are assumed to be high enough bandwidth that they can be ignored in the dynamic equations [23]. Equation (40) can be linearised about the nominal operating point by using the first-order approximation of the Taylor series expansion. Therefore, a linear small-signal system model can be defined for small excursions from the nominal operating point.

We obtained numerical approximations to the linearised system matrices by varying x and p from their nominal values (one element at a time) and recording the new \( \dot{x} \) and y vectors in DIGTEM.

Turbofan engine health monitoring is typically a two-step process [3]. In the first step, engine data is collected each flight at the same engine operating points and corrected to account for variability in ambient conditions. Data are typically collected for a period of about 3 s per flight at a rate of about 10 or 20 Hz. In the second step, the data are transferred to ground-based computers for post-flight analysis to determine engine health.

The goal of our turbofan engine health monitoring problem is to obtain an accurate estimate of \( \hat{\alpha} \) (the change in the health parameter vector), which varies slowly with time. We therefore assume that \( \hat{\alpha} \) is constant between measurement times. We also assume that the control input is perfectly known, so \( \delta u = 0 \). We augment the state vector with the health parameter vector [11] to obtain an augmented system equation. Then we can use a Kalman filter to estimate \( \hat{\alpha} \) and \( \hat{p} \). Actually, we are only interested in estimating \( \hat{p} \) (the health parameter

Table 2: Turbofan controls and nominal values

<table>
<thead>
<tr>
<th>Control</th>
<th>Nominal value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combustor fuel flow, kg s⁻¹</td>
<td>0.37</td>
</tr>
<tr>
<td>Augmentor fuel flow, kg s⁻¹</td>
<td>0</td>
</tr>
<tr>
<td>Nozzle throat area, cm²</td>
<td>430</td>
</tr>
<tr>
<td>Nozzle exit area, cm²</td>
<td>492</td>
</tr>
<tr>
<td>Fan vane angle, deg</td>
<td>-25</td>
</tr>
<tr>
<td>Compressor vane angle, deg</td>
<td>-20</td>
</tr>
</tbody>
</table>

Table 3: Turbofan health parameters and nominal values

<table>
<thead>
<tr>
<th>Health parameter</th>
<th>Nominal value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan airflow, kg s⁻¹</td>
<td>102</td>
</tr>
<tr>
<td>Fan efficiency</td>
<td>0.82</td>
</tr>
<tr>
<td>Compressor airflow, kg s⁻¹</td>
<td>48.7</td>
</tr>
<tr>
<td>Compressor efficiency</td>
<td>0.83</td>
</tr>
<tr>
<td>High-pressure turbine airflow, kg s⁻¹</td>
<td>41.0</td>
</tr>
<tr>
<td>High-pressure turbine enthalpy change, J kg⁻¹</td>
<td>101</td>
</tr>
<tr>
<td>Low-pressure turbine airflow, kg s⁻¹</td>
<td>48.3</td>
</tr>
<tr>
<td>Low-pressure turbine enthalpy change, J kg⁻¹</td>
<td>27.1</td>
</tr>
</tbody>
</table>
deviations, but the Kalman filter gives us the bonus of also estimating $\hat{\Delta}x$, the excursions of the original turbofan state variables.

It is known that health parameters do not improve over time. That is, $\delta p(1), \delta p(3), \delta p(4), \delta p(6)$ and $\delta p(8)$ are always less than or equal to zero and always decrease with time. Similarly, $\delta p(5)$ and $\delta p(7)$ are always greater than or equal to zero and always increase with time. In addition, it is known that the health parameters vary slowly with time. As an example, since $\delta p(1)$ is the constrained estimate of $\delta p(1)$, we can enforce the following constraints on $\delta p(1)$:

$$\hat{\delta}p(1) \leq 0$$
$$\hat{\delta}p_{k+1}(1) \leq \hat{\delta}p_k(1) + \gamma_1^-$$
$$\hat{\delta}p_{k+1}(1) \geq \hat{\delta}p_k(1) - \gamma_1^+$$

where $\gamma_1^+$ and $\gamma_1^-$ are non-negative factors chosen by the user that allow the state estimate to vary only within prescribed limits. Typically we choose $\gamma_1^+ > \gamma_1^-$ so that the state estimate can change more in the negative direction than in the positive direction. This is in keeping with our a priori knowledge that this particular state variable never increases with time. Ideally we would have $\gamma_1^+ = 0$ since $\delta p(1)$ never increases. However, since the state-variable estimate varies around the true value of the state-variable, we choose $\gamma_1^+ > 0$. This allows some time-varying increase in the state variable estimate to compensate for a state-variable estimate that is smaller than the true state-variable value.

These constraints are linear and can therefore easily be incorporated into the form required in the constrained filtering problem statement (15). Note that this does not take into account the possibility of abrupt changes in health parameters due to discrete damage events. That possibility must be addressed by other means (e.g., residual checking [3]) in conjunction with the methods presented in this study.

## 6 Simulation results

We simulated the discussed methods using Matlab. We simulated a steady-state 3 s burst of engine data measured at 10 Hz during each flight. Each of these routine services was performed at the single operating point shown in

![Fig. 2 Unconstrained Kalman filter estimates of the health parameters. The true health parameter changes were various values between $-3\%$ and $+3\%$. The true health parameter changes are shown as heavy lines, and the filter estimates are shown as lighter lines.](image)
The improved performance of the constrained filters comes at a price, and that price is computational effort. The filter with soft constraints requires only slightly (14%) more computational effort than the unconstrained filter, but the filter with hard constraints requires about four-times the computational effort of the unconstrained filter. This is because of the additional quadratic programming problem that is required for hard constraints. However, computational effort is not a critical issue for the particular application of turbofan health estimation since the filtering is performed on ground-based computers after each flight.

7 Discussion and conclusions

We have presented two methods to incorporate linear state inequality constraints into a Kalman filter. The first method incorporated hard constraints into the Kalman filter to maintain the state-variable estimates within a user-defined envelope. The second method incorporated soft constraints into the Kalman filter to ensure that the state-variable estimates vary slowly with time. The simulation results demonstrate the effectiveness of these methods, particularly for turbofan engine health estimation.

If the system whose state variables are being estimated has known state-variable constraints, then those constraints can be incorporated into the Kalman filter as shown in this study. However, in practice, the constraints enforced in the filter might be more relaxed than the true constraints. This allows the filter to correct state-variable estimates in a direction that the true state variables might never change. This is a departure from strict adherence to theory, but in practice this improves the performance of the filter. This is an implementation issue that is conceptually similar to tuning a standard Kalman filter.

It was seen in theorem 2 that the filter with hard constraints has a smaller estimation error covariance than the unconstrained Kalman filter. At first this seems counter-intuitive, since the standard Kalman filter is by definition the minimum variance filter. However, we have changed the problem by introducing state-variable constraints. Therefore, the standard Kalman filter is not the minimum-variance filter for the turbofan engine health estimation cases) but they do show the improvement that is possible with constrained Kalman filters.

![Figure 3](image1.png)

**Fig. 3** Soft-constrained Kalman filter estimates of the health parameters. The true health parameter changes were various values between −3% and +3%. The true health parameter changes are shown as heavy lines, and the filter estimates are shown as lighter lines.

![Figure 4](image2.png)

**Fig. 4** Hard-constrained Kalman filter estimates of the health parameters. The true health parameter changes were various values between −3% and +3%. The true health parameter changes are shown as heavy lines, and the filter estimates are shown as lighter lines.

<table>
<thead>
<tr>
<th>Health parameter</th>
<th>Estimation error, %</th>
<th>Unconstrained filter</th>
<th>Soft-constrained filter</th>
<th>Hard-constrained filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan airflow</td>
<td>0.129</td>
<td>0.113</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>Fan efficiency</td>
<td>0.163</td>
<td>0.149</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td>Compressor airflow</td>
<td>0.152</td>
<td>0.146</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>Compressor efficiency</td>
<td>0.101</td>
<td>0.087</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td>HPT airflow</td>
<td>0.119</td>
<td>0.114</td>
<td>0.076</td>
<td></td>
</tr>
<tr>
<td>HPT enthalpy change</td>
<td>0.092</td>
<td>0.078</td>
<td>0.050</td>
<td></td>
</tr>
<tr>
<td>LPT airflow</td>
<td>0.104</td>
<td>0.091</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td>LPT enthalpy change</td>
<td>0.168</td>
<td>0.155</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.128</td>
<td>0.116</td>
<td>0.080</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Kalman filter estimation errors. HPT = high-pressure turbine, and LPT = low-pressure turbine. The numbers shown are root-mean-square estimation errors (percent) averaged over 16 simulations, where each simulation had a linear degradation of all eight health parameters.
problem, and we can do better with the constrained Kalman filter.

We saw that the filter with hard constraints required a much larger computational effort than the standard Kalman filter. This is due to the addition of the quadratic programming problem that must be solved in the constrained Kalman filter. The engineer must therefore perform a trade of between computational effort and estimation accuracy. For real-time applications the improved estimation accuracy may not be worth the increase in computational effort.

It was seen in Figs. 2–4 that although the constrained filters improve the estimation accuracy, the general trend of the state-variable estimates does not change with the introduction of state constraints. This is because the constrained filters are based on the unconstrained Kalman filter. The constrained filter estimates therefore have the same shape as the unconstrained estimates until the constraints are violated, at which point the state-variable estimates are projected onto the edge of the constraint boundary. The constrained filters presented in this study are not qualitatively different to the standard Kalman filter; they are rather a quantitative improvement on the standard Kalman filter.

Note that the Kalman filter works well only if the assumed system model matches reality fairly closely. The method presented in this study, by itself, will not work well if there are large sensor biases or hard faults due to severe component failures. A mission-critical implementation of a Kalman filter should always include some sort of residual check to verify the validity of the Kalman filter results, particularly for the application of turbofan engine health estimation considered in this study [3, 24].

Although we have considered only linear state constraints, it is not conceptually difficult to extend this work to nonlinear constraints. If the state constraints are nonlinear they can be linearised as discussed in [2].

Finally we emphasise that the constrained filters presented here are not optimal filters; they are rather modifications of the optimal Kalman filter in the presence of constraints. An optimal constrained filter would have to take into account constrained probability density functions [25]. This is similar to saying that the Kalman filter itself is not optimal unless the noise is Gaussian. Particle filters could be used for truly optimal constrained filtering, although they tend to optimality only as the number of particles becomes very large [26] and therefore may not be computationally feasible.

Further work along the lines of this research could focus on combining our work with [27] in order to guarantee convergence in the presence of nonlinear constraints. Other efforts could explore the incorporation of state constraints for optimal smoothing, or the use of state constraints in $H_{\alpha}$ filtering [28]. Further work could also focus on integrating the nonlinear simulation logic in DIGTEM [8, 22] with the Kalman filter to obtain more complete results. This would also allow us to more easily test the Kalman filter at various operating points without translating data from DIGTEM to Matlab.

8 Acknowledgment

This work was supported in part by a NASA/ASEE Summer Faculty Fellowship.

9 References


