Optimal load sharing in a binary multicomponent system

Arne Bang Huseby and Olav Breivik
Department of Mathematics, University of Oslo
P.O.Box 1053 Blindern, N-0316 Oslo, Norway

Abstract

In the present paper we consider a system consisting of n components that is exposed to the load of supplying a certain amount of utility, e.g., electrical power. The load on the system is distributed among the components. When functioning each component is capable of handling a certain amount of load. The load capacity of a component is assumed to be constant throughout its lifetime. The main objective of the present paper is developing methods for *optimal load sharing* among the components subject to the constraints imposed by the load capacities and demand on the system. In the paper we show how to solve the problem in several special cases, and outline a greedy algorithm for handling the general case.

1. Introduction

In the present paper we consider a system consisting of n components that is exposed to the load of supplying a certain amount of utility. For simplicity we assume that this load is constant over time, and denote it by K. The load on the system is distributed among the components.

A bivariate load sharing model was introduced by Freund (1961) and generalized by Mendel and Huseby (1996). Similar models were also studied by Noortwijk et. al. (1994) and Singpurwalla and Youngren (1993). The present paper extends the principles from Mendel and Huseby (1996).

The components of the system have two possible states: *functioning* or *failed*. The state of the ith component at time t is denoted by $X_i(t)$, where $X_i(t)$ is one if the ith component is functioning at time t, and zero otherwise, $i = 1, \ldots , n$. As long as the ith component is functioning, it is capable of handling a certain maximum load denoted by $\kappa_i > 0$, which will be referred to as the *capacity* of the component, $i = 1, \ldots , n$. When a component fails, its capacity is immediately reduced to zero. The total capacity of the system at time t, denoted by $Y(t)$, can then be expressed as:

$$Y(t) = \sum_{i=1}^{n} \kappa_i \cdot X_i(t) \tag{1}$$

We do not consider the possibility of repairing components. The system is said to be *functioning* at time t if it is capable of supplying the demanded load, i.e., if $Y(t) \geq K$, otherwise it is said to be *failed*. The lifetime of the system, denoted T, is given by:

$$T = \sup\left\{t > 0 : Y\left(t\right) \geq K\right\} \tag{2}$$

The load on the ith component at time $t > 0$ is denoted by $q_i(t)$, $i = 1, \dots, n$. We will refer to these functions as the *load functions* of the system. These functions are assumed to be controllable for all $0 < t < T$ subject to the following restrictions:

$$\sum_{i=1}^{n} q_i(t) = K, \tag{3}$$

and:

$$0 \leq q_i(t) \leq \kappa_i, \quad i = 1, \dots, n, \tag{4}$$

where (3) ensures that the system load satisfies the demand, while (4) ensures that the component loads do not exceed their respective capacities. We also introduce *cumulative load functions* $Q_1(t), \dots, Q_n(t)$, defined as:

$$Q_i(t) = \int_0^t q_i(u)\,du, \quad i = 1, \dots, n. \tag{5}$$

We assume that each component is capable of handling a certain stochastic *load volume* during its lifetime. The load volume provided by the ith component is denoted by V_i, $i = 1, \dots, n$. As long as $Q_i(t) < V_i$ the component is considered to be functioning. However, as soon as $Q_i(t)$ reaches V_i, the component fails. Thus, we have the following relation:

$$X_i(t) = \begin{cases} 1 & \text{if } Q_i(t) < V_i \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, \dots, n. \tag{6}$$

Moreover, the lifetimes of the components, denoted $T_1, \dots, T_n$ are given by:

$$T_i = \sup\{t > 0 : Q_i(t) < V_i\}, \quad i = 1, \dots, n. \tag{7}$$

Note that the lifetime of a component depends on the load it is exposed to. Hence, before introducing the component lifetime distributions, it is convenient to start out with the distributions for the load volumes $V_1, \dots, V_n$. Thus, we introduce the function:

$$P_i(v) = P(V_i > v), \quad v > 0 \text{ and } i = 1, \dots, n, \tag{8}$$

and assume that these functions satisfy the following separable differential equations:

$$P_i'(v) = -\lambda_i(v) P(v), \quad v > 0 \text{ and } i = 1, \dots, n, \tag{9}$$

for suitable nonnegative functions $\lambda_1, \dots, \lambda_n$. With a slight abuse of terminology we refer to the λ_n-functions as the *failure rate functions* of the components. With boundary conditions $P_i(0) = 1$, $i = 1, \dots, n$, we get the familiar representation:

$$P_i(v) = \exp\left(-\int_0^v \lambda_i(u)\,du\right), \quad v > 0 \text{ and } i = 1, \dots, n. \tag{10}$$

We then turn to the lifetimes $T_1, \dots, T_n$ and note that for $t > 0$, $T_i > t$ if and only if $Q_i(t) < V_i$, $i = 1, \dots, n$. Hence,

$$P(T_i > t) = P(V_i > Q_i(t)) = \exp\left(-\int_0^{Q_i(t)} \lambda_i(u)\,du\right), \quad t > 0 \text{ and } i = 1, \dots, n. \tag{11}$$

By substituting $u = Q_i(s)$ the expression for $P(T_i > t)$ can be written in the following form:

$$P(T_i > t) = \exp\left(-\int_0^t \lambda_i(Q_i(s))q_i(s)\,ds\right), \quad t > 0 \text{ and } i = 1, \dots, n. \tag{12}$$

Throughout this paper we assume that:

$$\sum_{i=1}^n \kappa_i \geq K \tag{13}$$

$$\sum_{i=1}^n \kappa_i - \kappa_j < K, \quad j = 1, \dots, n. \tag{14}$$

Thus, the system is a simple *series system* which is functioning if and only if all the components are functioning. Hence, the lifetime distribution of the system is given by:

$$P(T > t) = \exp\left(-\int_0^t \sum_{i=1}^n \lambda_i(Q_i(s))q_i(s)\,ds\right), \quad t > 0. \tag{15}$$

Moreover, the expected system lifetime is given by:

$$E[T] = \int_0^\infty \exp\left(-\int_0^t \sum_{i=1}^n \lambda_i(Q_i(s))q_i(s)\,ds\right)dt. \tag{16}$$

The main objective is to choose the functions $q_1, \dots, q_n$ so that $E[T]$ is maximized subject to the constraints (3) and (4).

2 Optimal load sharing under constant failure rates

In this section we assume that the failure rates of the components are constant, i.e.:

$$\lambda_i(t) = \lambda_i, \quad t > 0 \quad \text{and } i = 1, \dots, n. \tag{17}$$

By inserting this into (15) and (16), it is easy to see that these formulas can be simplified to:

$$P(T > t) = \exp\left(-\int_0^t \sum_{i=1}^n \lambda_i(Q_i(s))q_i(s)\,ds\right) = \exp\left(-\sum_{i=1}^n \lambda_i\, Q_i(t)\right), \quad t > 0. \tag{18}$$

$$E[T] = \int_0^\infty \exp\left(-\sum_{i=1}^n \lambda_i\, Q_i(t)\right)dt. \tag{19}$$

In order to proceed it is convenient to express the constraints (3) and (4) in terms of the cumulative load functions:

$$\sum_{i=1}^{n} Q_i(t) = Kt, \tag{20}$$

$$0 \le Q_i(t) \le \kappa_i t, \quad i = 1, \dots, n. \tag{21}$$

Considering $P(T > t)$, we see that for any given $t > 0$ this is maximized by putting as much load as possible on components with low failure rates, and as little as possible on components with high failure rates. Thus, we order the failure rates letting $\lambda_{(i)}$ denote the ith smallest rate, $i = 1, \dots, n$:

$$\lambda_{(1)} \le \lambda_{(2)} \le \dots \le \lambda_{(n)}. \tag{22}$$

Then it is easy to see that for a given t the optimal values of $Q_1(t), \dots, Q_n(t)$ are given by:

$$Q_{(1)}(t) = \min\left\{\kappa_{(1)}t,\, Kt\right\},$$

$$Q_{(2)}(t) = \min\left\{\kappa_{(2)}t,\, Kt - Q_{(1)}(t)\right\},$$

$$\dots$$

$$Q_{(n)}(t) = \min\left\{\kappa_{(n)}t,\, Kt - \sum_{i=1}^{n-1} Q_{(i)}(t)\right\}. \tag{23}$$

Differentiating $Q_1(t), \dots, Q_n(t)$ with respect to t we get the following optimal load functions:

$$q_{(1)}(t) = \min\left\{\kappa_{(1)},\, K\right\},$$

$$q_{(2)}(t) = \min\left\{\kappa_{(2)},\, K - q_{(1)}(t)\right\},$$

$$\dots$$

$$q_{(n)}(t) = \min\left\{\kappa_{(n)},\, K - \sum_{i=1}^{n-1} q_{(i)}(t)\right\}. \tag{24}$$

Note that since obviously $q_{(1)}(t)$ is constant with respect to t, then so is $q_{(2)}(t)$ etc. Hence, all these load functions are constant with respect to t. Moreover, the system survival probability is maximized for all $t > 0$. Hence, this solution also maximizes the expected system lifetime.

Finally, note that since we have assumed that we have a series system, it is easy to see, using (13) and (14), that we in fact have:

$$\kappa_{(j)} < K - \sum_{i=1}^{j-1} q_{(i)}(t), \quad j = 1, \dots, n-1. \tag{25}$$

Hence, (24) can be simplified to:

$$q_{(1)}(t) = \kappa_{(1)}, \dots, q_{(n-1)}(t) = \kappa_{(n-1)}, \quad \text{and} \quad q_{(n)}(t) = K - \sum_{i=1}^{n-1} \kappa_{(i)}. \tag{26}$$

3 Optimal load sharing under constant loads

We proceed by optimizing load sharing assuming that the load functions are constant. That is, we let:

$$q_i(t) = q_i, \quad t > 0 \text{ and } i = 1, \dots, n. \tag{27}$$

The cumulative load functions can be written as:

$$Q_i(t) = q_i t, \quad t > 0 \text{ and } i = 1, \dots, n. \tag{28}$$

By inserting this into (15) and (16), we obtain:

$$P(T > t) = \exp\left(-\int_0^t \sum_{i=1}^n \lambda_i(q_i s) q_i ds\right), \quad t > 0. \tag{29}$$

$$E[T] = \int_0^\infty \exp\left(-\int_0^t \sum_{i=1}^n \lambda_i(q_i s) q_i ds\right) dt. \tag{30}$$

Maximizing $P(T > t)$ and $E[T]$ in this case is difficult, so we simplify the problem by assuming that the failure rate functions are of the following form:

$$\lambda_i(Q_i(t)) = \lambda(Q_i(t)/\alpha_i), \quad t > 0 \quad \text{and} \quad i = 1, \dots, n, \tag{31}$$

where $\alpha_1, \dots, \alpha_n$ are suitable positive numbers, and λ is a positive differentiable function. By considering (29), we see that in order to maximize the system's survival probability for a given t, $\mathbf{q} = (q_1, \dots, q_n)$ should be chosen so that the integrand is as small as possible for each s. We denote this integrand by $\phi = \phi(\mathbf{q})$.

By combining (28) and (31), we see that ϕ can be written as:

$$\phi(\mathbf{q}) = \sum_{i=1}^n \lambda(q_i s/\alpha_i) q_i. \tag{32}$$

We also compute the partial derivatives of ϕ with respect to $q_1, \dots, q_n$:

$$\frac{\partial \phi(\mathbf{q})}{\partial q_i} = \lambda'(q_i s/\alpha_i) q_i s/\alpha_i + \lambda(q_i s/\alpha_i), \quad i = 1, \dots, n. \tag{33}$$

In order to minimize ϕ subject to the constraints (3) and (4) we introduce the function $\psi(\mathbf{q}) = q_1 + \cdots + q_n$, and reformulate (3) in terms of ψ as:

$$\psi(\mathbf{q}) = K. \tag{34}$$

Ignoring for the moment the other constraints, the optimization problem can then be expressed in terms of the Lagrange function:

$$\Lambda(\mathbf{q}, z) = \phi(\mathbf{q}) + z(\psi(\mathbf{q}) - K), \tag{35}$$

where z denotes the Lagrange multiplier. A stationary point for $\Lambda(\mathbf{q}, z)$ is found by solving the equation:

$$\nabla\phi(\mathbf{q}) = -z\nabla\psi(\mathbf{q}), \tag{36}$$

combined with the constraint (34). Noting that $\nabla\psi(\mathbf{q}) = (1, \dots, 1)$, it follows that the solution must satisfy:

$$\frac{\partial\phi(\mathbf{q})}{\partial q_1} = \dots = \frac{\partial\phi(\mathbf{q})}{\partial q_n}. \tag{37}$$

Using (33) we see that this will hold provided that $q_i = q\,\alpha_i$, $i = 1, \dots, n$, where q is chosen so that $\psi(\mathbf{q}) = q\alpha_1 + \cdots + q\alpha_n = K$. Thus, we get that:

$$q_i = \frac{\alpha_i K}{\sum_{j=1}^{n} \alpha_j}, \quad i = 1, \dots, n. \tag{38}$$

Assuming that this stationary point is a *minimum* value of ϕ and that this solution also satisfies (4), we have found the optimal loads. Note that this solution does not depend on time, so these loads will in fact minimize the integrand of (29) for all s, and hence also maximize the survival probability for all t. From this we finally get that this solution maximizes the expected system lifetime.

Unfortunately, the above procedure may sometimes produce a *maximum* value of ϕ. The following result based on the concept of *quasi-convexity*, provides a criterion for when the stationary point corresponds to an optimal solution. Quasi-convexity is defined in Appendix A (Definition A.1). See also Boyd and Vandenberghe (2004).

Theorem 1. Assume that of ϕ is *quasi-convex* and that the loads given in (38) also satisfy (4). Then the solution maximizes the survival probability for all t and the expected system lifetime.

Proof: We assume that ϕ is *quasi-convex* and that $\mathbf{q}$ satisfies (34) and (36). We then assume that $\mathbf{q}' = (q'_1, \dots, q'_n)$ is a vector such that $\phi(\mathbf{q}') < \phi(\mathbf{q})$. By Proposition A.4 it follows that $\nabla\phi(\mathbf{q})^T(\mathbf{q}' - \mathbf{q}) < 0$. Hence, by (36) we also have $-z\nabla\psi(\mathbf{q})^T(\mathbf{q}' - \mathbf{q}) < 0$, and since $\nabla\psi(\mathbf{q}) = (1, \dots, 1)$, this implies that:

$$-z\nabla\psi(\mathbf{q})^T(\mathbf{q}' - \mathbf{q}) = -z\left(\sum_{i=1}^{n} q'_i - \sum_{i=1}^{n} q_i\right) = -z\left(\sum_{i=1}^{n} q'_i - K\right) < 0 \tag{39}$$

Hence, $\psi(\mathbf{q}') = (q'_1 + \dots + q'_n) \neq K$, so $\mathbf{q}'$ does not satisfy (34). Thus, we conclude that $\mathbf{q}$ is indeed an optimal solution ■

If the loads given in (38) do not satisfy (4), we let $I = \{i : q_i > \kappa_i\}$. For all $i \in I$, we replace q_i by κ_i. This adjustment causes the sum of the loads to be smaller than K. We compensate for this by increasing the loads currently not exceeding their respective capacities. Let $S = \{\mathbf{q} : q_i = \kappa_i, i \in I\}$, and let $\phi|_S$ denote the function ϕ restricted to the set S. We then proceed by minimizing $\phi|_S$ subject to:

$$\sum_{i \notin I} q_i = K - \sum_{i \in I} \kappa_i \quad (40)$$

$$0 \leq q_i \leq \kappa_i, \quad i \notin I. \quad (41)$$

Assuming that ϕ is quasi-convex, it follows by Proposition A.2 that $\phi|_S$ is quasi-convex as well. Thus, we may apply the same method as above, and obtain the solution:

$$q_i = \frac{\alpha_i \left(K - \sum_{i \in I} \kappa_i \right)}{\sum_{i \notin I} \alpha_i}, \quad i \notin I. \quad (42)$$

As above it may happen that this solution does not satisfy (41). If so, we adjust the loads once again and repeat the process on the remaining loads until the capacity constraint is satisfied.

4 Optimal load sharing in the general case

We now consider the general problem of maximizing (15) and (16) subject to the constraints (3) and (4). We denote the system failure rate at time t by $\rho(t)$. This is given by:

$$\rho(t) = \sum_{i=1}^{n} \lambda_i \left(Q_i(t) \right) q_i(t) \quad (43)$$

The main idea now is to apply a *greedy* algorithm where at each point of time we choose the locally optimal solution hoping that this would lead to a globally optimal load sharing strategy as well. That is, for each t we share the loads so that $\rho(t)$ is minimized subject to the constraints (3) and (4). While this principle often produces optimal or close to optimal solutions, it is not guaranteed to work. Thus, the proposed method should be used with caution.

It is easy to see that $\rho(t)$ given in (43) is minimized by putting as much load as possible on the components with low failure rates, and as little as possible on components with high failure rates. By ordering the failure rates such that:

$$\lambda_{(1)} \left(Q_{(1)}(t) \right) \leq \lambda_{(2)} \left(Q_{(2)}(t) \right) \leq \ldots \leq \lambda_{(n)} \left(Q_{(n)}(t) \right). \quad (44)$$

we end up with a solution like the one given in (24). Note, however, that the optimal solution may change over time. If the optimization problem is solved numerically by discretizing the time, one may experience that the order of the failure rates changes very rapidly. As a result the load functions become very unstable. However, this problem can often be avoided by balancing the loads so that the failure rate functions are equal at each point of time. Thus, we look for equilibrium solutions where the cumulative load functions satisfy:

$$\lambda_1 \left(Q_1(t) \right) = \ldots = \lambda_n \left(Q_n(t) \right) \quad (45)$$

$$\sum_{i=1}^{n} Q_i(t) = Kt \quad (46)$$

In order to solve these equations, we introduce $L(t)$ as the common failure rate value:

$$L(t) = \lambda_i\left(Q_i(t)\right), \quad t > 0 \quad \text{and} \quad i = 1, \dots, n. \tag{47}$$

Assuming that all the λ_i-functions have well-defined inverses, we can solve (47) with respect to the cumulative load functions:

$$Q_i(t) = \lambda_i^{-1}\left(L(t)\right), \quad t > 0 \quad \text{and} \quad i = 1, \dots, n. \tag{48}$$

Combining (48) and (46) we then get:

$$\sum_{i=1}^{n} \lambda_i^{-1}\left(L(t)\right) = Kt \tag{49}$$

from which $L(t)$ may be determined. Finally, the load functions $q_1(t), \dots, q_n(t)$ are determined by inserting $L(t)$ into (48) and differentiating:

$$q_i(t) = \tfrac{d}{dt}\lambda_i^{-1}\left(L(t)\right), \quad t > 0 \quad \text{and} \quad i = 1, \dots, n. \tag{50}$$

Assuming that these load functions also satisfies the capacity constraint (4), we have a feasible solution. If this is *not* the case, it is not possible to obtain a perfectly balanced solution. In such cases one may have to resort to a numerical solution using discretized time.

We close this section by presenting two examples. In the first example we consider failure rate functions of the form given in (31), but where we also assume that the function λ has a well-defined inverse. In this case (47) can be written as:

$$L(t) = \lambda\left(Q_i(t)/\alpha_i\right), \quad t > 0 \quad \text{and} \quad i = 1, \dots, n. \tag{51}$$

Hence, it follows that:

$$Q_i(t) = \alpha_i \lambda^{-1}\left(L(t)\right), \quad t > 0 \quad \text{and} \quad i = 1, \dots, n. \tag{52}$$

By inserting this into (46), we get the solution:

$$q_i(t) = \frac{\alpha_i K}{\sum_{j=1}^{n} \alpha_j}, \quad t > 0 \quad \text{and} \quad i = 1, \dots, n. \tag{54}$$

Thus, we get exactly the same solution as we did assuming constant loads.

In the second example we consider a case where $n = 2$, and where:

$$\lambda_1\left(Q_1(t)\right) = \sqrt{Q_1(t)/\alpha_1}, \quad \lambda_2\left(Q_2(t)\right) = Q_2(t)/\alpha_2, \quad t > 0. \tag{55}$$

Thus, we get:

$$Q_1(t) = \alpha_1 L^2(t), \quad Q_2(t) = \alpha_2 L(t), \quad t > 0. \tag{56}$$

By inserting this into (46) we get the following equation:

$$\alpha_1 L^2(t) + \alpha_2 L(t) = Kt, \quad t > 0. \tag{57}$$

Solving this with respect to $L(t)$, eliminating the negative root and inserting the result into (56) yields:

$$Q_1(t) = \alpha_1 \left[\frac{\sqrt{\alpha_2^2 + 4\alpha_1 Kt} - \alpha_2}{2\alpha_1} \right]^2, \quad Q_2(t) = \alpha_2 \frac{\sqrt{\alpha_2^2 + 4\alpha_1 Kt} - \alpha_2}{2\alpha_1}, \quad t > 0. \qquad (58)$$

By differentiating we get the following load functions:

$$q_1(t) = \left[1 - \frac{\alpha_2}{\sqrt{\alpha_2^2 + 4\alpha_1 Kt}} \right] K, \quad q_2(t) = \left[\frac{\alpha_2}{\sqrt{\alpha_2^2 + 4\alpha_1 Kt}} \right] K, \quad t > 0. \qquad (59)$$

We observe that in this case the load functions are *not* constant. More specifically, $q_1(0) = 0$ while $q_2(0) = K$. As t increases, however, component 1 gets an increasing share of the total load. When t goes to infinity, $q_1(0) = K$ while $q_2(0) = 0$. Thus, unless $\kappa_1 = \kappa_2 = K$, the perfectly balanced solution is *not* feasible.

5. Conclusions and further work

In the present paper we have briefly considered the problem of optimizing load sharing between components in a simple series system, and shown how to solve this problem in several cases. In Section 2 we considered the situation where all the failure rates were constant, in which case it is easy to derive exact analytical expressions for the survival probability and expected lifetime of the system. In Section 3 we described how to solve the optimization problem assuming constant loads and that all the failure rate functions were equal except for a scaling factor. Under these assumptions a stationary solution can be found using standard Lagrange methods. Moreover, if the objective function is quasi-convex, the stationary solution is an optimal solution as well. Finally, in Section 4 we consider the general case, and describe how this can be handled using a greedy approach aiming at an equilibrium solution. In a forthcoming paper these results will be extended to general threshold systems. In that paper we will also provide further details about how the optimization can be carried out when the optimal solution does not satisfy the capacity constraint or when the objective function is not quasi-convex.

References

1. Boyd, S. and Vandenberghe, L. *Convex Optimization*, Cambridge University Press. (2004).
2. Freund J. E., A bivariate extension of the exponential distribution *J. Am. Stat. Assoc.*, **56**, pp 971-977 (1961).
3. Mendel, M. B. and Huseby, A. B., The influence of load on the lifetime distribution. *Lifetime Data Analysis*, **2**, pp 261-276 (1996).
4. Noortwijk, J. M., Cook, R. M. and Kok, M., A Bayesian failure model based on isotropic deterioration. *European Journal of Operational Research*, (1994).
5. Singpurwalla, N. and Youngren, M., Multivariate distributions induced by dynamic environments. *Scand. Jour. Statist.*, **20**, pp 251-261 (1993)

Appendix. Quasi-convex functions

According to Boyd and Vandenberghe (2004) a *quasi-convex function* is defined as follows (**R** denotes the set of real numbers).

Definition A.1. Let $S \subseteq \mathbf{R}^n$ be a convex set. We say that a function $g : S \to \mathbf{R}$ is *quasi-convex* if for any pair of vectors $\mathbf{x}_1, \mathbf{x}_2 \in S$ and $\lambda \in [0, 1]$ we have:

$$g\left(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2\right) \leq \max\left\{g(\mathbf{x}_1), g(\mathbf{x}_2)\right\}. \qquad \text{(A.1)}$$

It is easy to see that quasi-convexity is a weaker property than ordinary convexity. In particular, any convex function is quasi-convex as well.

The following result states that quasi-convexity is preserved under restrictions.

Proposition A.2. Let $S \subseteq \mathbf{R}^n$ be a convex set, and let $g : S \to \mathbf{R}$ be a quasi-convex function. Moreover, let T be a convex subset of S then the restriction of g to T, denoted $g|_T$ is a quasi-convex function.

Proof: Consider a pair of vectors $\mathbf{x}_1, \mathbf{x}_2 \in T$ and let $\lambda \in [0, 1]$. Since T is a subset of S, obviously $\mathbf{x}_1, \mathbf{x}_2 \in S$ as well. Since g is assumed to be quasi-convex, it follows by Definition A.1 that g satisfies (A.1) for the chosen pair of vectors. Moreover, since T is a convex set, it follows that $\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2 \in T$. Thus, $g|_T$ is defined for this vector, and we have $g|_T(\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) = g(\lambda\mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2)$. Finally, since $g|_T(\mathbf{x}_1) = g(\mathbf{x}_1)$ and $g|_T(\mathbf{x}_2) = g(\mathbf{x}_2)$, we get that $g|_T$ satisfies (A.1) as well ■

If g is differentiable, we have the following result:

Proposition A.3. Let $S \subseteq \mathbf{R}^n$ be a convex set, and let $g : S \to \mathbf{R}$ be a differentiable function. Then g is quasi-convex if and only if for any pair of vectors $\mathbf{x}_1, \mathbf{x}_2 \in S$ we have:

$$g(\mathbf{x}_2) \leq g(\mathbf{x}_1) \quad \Rightarrow \quad \nabla g(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) \leq 0 \qquad \text{(A.2)}$$

Proof: See Boyd and Vandenberghe (2004).

Proposition A.4. Let $S \subseteq \mathbf{R}^n$ be a convex set, and let $g : S \to \mathbf{R}$ be a differentiable quasi-convex function. Then let $\mathbf{x}_1, \mathbf{x}_2$ be chosen in the interior of S such that $g(\mathbf{x}_2) < g(\mathbf{x}_1)$. Then $\nabla g(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) < 0$.

Proof: Let $\mathbf{x}_1, \mathbf{x}_2 \in S$ be so that $\varepsilon = g(\mathbf{x}_1) - g(\mathbf{x}_2) > 0$. By Proposition A.3 we immediately get that $\nabla g(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) \leq 0$. In order to prove the result, we need to show that this inequality is strict. Assume instead that

$\nabla g(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) = 0$. Since g is differentiable, g is continuous as well. Thus, since $\mathbf{x}_2$ is in the interior of S, there exists a neighbourhood $O(\mathbf{x}_2)$ in S around $\mathbf{x}_2$ such that for any $\mathbf{x} \in O(\mathbf{x}_2)$ we have $|g(\mathbf{x}) - g(\mathbf{x}_2)| < \varepsilon/2$. Hence, it follows that for any $\mathbf{x} \in O(\mathbf{x}_2)$ we also have:

$$g(\mathbf{x}_1) - g(\mathbf{x}) \geq g(\mathbf{x}_1) - g(\mathbf{x}_2) - |g(\mathbf{x}) - g(\mathbf{x}_2)| = \varepsilon - |g(\mathbf{x}) - g(\mathbf{x}_2)| > \frac{\varepsilon}{2} \qquad \text{(A.3)}$$

At the same time, it is easy to see that we can also find a vector $\mathbf{x}_3$ within the neighbourhood $O(\mathbf{x}_2)$ such that $\nabla g(\mathbf{x}_1)^T (\mathbf{x}_3 - \mathbf{x}_1) > 0$. However, since by (A.3) we know that $g(\mathbf{x}_3) < g(\mathbf{x}_1)$, this contradicts that g is quasi-convex. Hence, we conclude that $\nabla g(\mathbf{x}_1)^T (\mathbf{x}_2 - \mathbf{x}_1) < 0$ as stated ■